2003 Quality and Productivity Research Conference

May 21-23 IBM T.J. Watson Research Center

Plenary Presentation

Peter W.M. John

Thank you for inviting me to be with you today. I am delighted to be here.

Paul Tobias suggested that my talk should be somewhat autobiographical. So I am going to give you a rather personal talk about my fifty years or so of life with statistics in general and experimental design in particular. There are two themes: serendipity; growth and change in the field of design of experiments over the years.

In retrospect I am struck by the part played by serendipity: that is, a pattern of happy and unexpected discoveries made by accident. I now see my statistical life as having been heavily influenced by serendipity—fortunate accidents and the help of kind friends—essentially an Odyssey in the sense of a long uncharted voyage.

Such an odyssey does not fit the conventional model of pure mathematicians, who are supposed to have decided in the cradle to dedicate themselves to mathematics, pick an area of specialization in which to concentrate for the rest of their lives, producing papers on that specialty and turning out more students, who also pursue that specialty ad inf.

This is rarely a realistic model for applied statisticians. Why not? Because we are applications oriented. We are problem solvers. We research statisticians pursue "applications oriented theory". Either we find a problem, or it finds us, and we use a combination of statistical intuition and mathematical skills to develop a theoretical solution, if we can, or an ad hoc solution if we cannot. Our theoretical interests go where the problems are.

Hence, the subjects that we study, and the questions that we strive to answer, depend upon where we are at any given time. Where we are often depends upon forces beyond our control. It depends upon decisions made by us, or for us, for different reasons and on, at best, partial information.

I would fail the first of the criteria in the model for pure mathematicians mentioned earlier: my odyssey began with a fortunate accident. In Wales my parents sent me to a local private middle school—very good, except that they did not teach Greek—plenty of Latin but no Greek. So, when I won a scholarship to boarding high school in England in 1937, I did not have the Greek prerequisites to go with the best boys into the classical stream. I was, therefore, banished to the science side. There the mathematics master was outstanding, but he was called up for the army shortly before the war broke out. The chemistry master was an utter dud. I did not plan it that way, nor did my parents. The Fates did. Hence, I won my scholarships to Jesus College, Oxford as a mathematician. That gave me a grounding in old-fashioned mathematics. I emphasize old-fashioned. For example, in algebra I never heard about groups and vector spaces.

When I finished my B.A. after the war, in 1948, there were no jobs for mathematicians in U.K. except school teaching, which I did not want to do. But serendipity smiled again. One afternoon I saw on the notice board in the porter's lodge of my college an announcement of a new program: the post graduate diploma in statistics (now the M.Sc.). I had a year left on Britain's equivalent of the GI Bill, so I applied.

For the diploma I had three teachers: Pat Moran, who went home to Australia for a distinguished career as a probabilist; David Finney, who invented fractional factorials; and Mike Sampford, who died prematurely.

Fifty-eight years ago, in 1945, Finney, who was then in the Rothamsted Experiment Station, wrote the first paper on fractional factorials. It did not appear in an engineering journal. It appeared in a journal called The Annals of Eugenics.

I would love to continue by telling you how I sat at Finney's feet, heard him lecture on fractional factorials, and was inspired forevermore. It would have made a good story, and you would probably never have known the difference. But the truth is that David lectured that year on mathematical statistics—hypothesis testing and that kind of thing. Mike Sampford gave the design of experiments lectures: I remember that he included a couple of lectures on the 2^3 factorial. Unfortunately, and this is no reflection upon Mike, I did not have much of a clue of what he was talking about. I had, by then, never seen an experiment, except for those cookbook things that one does in high school and in undergraduate labs. I passed the diploma exams because I could do mathematics, not because I knew anything about data.

I did get one important thing out of the course and that was a copy of the original monograph by Yates: The Design and Analysis of Factorial Experiments. It was written in 1937, only twelve years before Mike's lectures, and published by the Imperial Bureau of Soil Science. All the examples are agricultural. I still have it in my bookcase. There are ninety-five pages. It cost five shillings, which in those days was about a dollar and a quarter. Amortized over fifty-four years, it has been very good value for money.

It is worthy of mention that this monograph was written before the days of the electronic computer. In the notes at the end Yates said, "When a computing machine is used (he really means mechanical adding and multiplying machines), working means are best avoided, especially if they are such as to introduce negative numbers." For the computation of standard errors, he said "a good 10 inch slide rule (three significant figures) will give all necessary accuracy, and is very convenient, since square roots may be read directly."

We sometimes forget that one of the main virtues of orthogonality, in addition to the part it plays in such modern ideas as D-optimality, was that, before modern computers became plentiful, statisticians could do the calculations for orthogonal designs with the equipment that they had and a bit of cunning, but the analysis of most nonorthogonal designs was a very time consuming and difficult job indeed and the amount of effort required was sometimes prohibitive.

The job situation in UK was no better in 1949. But again, by serendipity, a temporary job turned up in the USA as an instructor in mathematics at the University of Oklahoma, teaching calculus to veterans. I had no opportunity to do statistics for six years. I decided to stay and became a probabilist, writing a dissertation on Birth and Death Processes under a distinguished analyst, Casper Goffman. Like Paul, I am a reformed, or reconstructed, probabilist.

Upon graduation in 1955, I took the appropriate next step of an assistant professorship in mathematics at the University of New Mexico, teaching their statistics course. Two years later I went into industry for two reasons. First, I was learning nothing in isolation at UNM and I saw no opportunity for any personal growth in statistics. Secondly, I was making $4500 per annum; the move doubled my salary to the amazing sum of $9000.

I chose Chevron over another famous oil company in the San Francisco Bay area because they had Henry Scheffé as a consultant. It dawned upon me that, if I paid attention to what Scheffé said, I might really learn some statistics. I was right. Chevron had acquired a mainframe computer in 1955. It was across the Bay in the city. There was a smaller one in our lab at Richmond, but the research people really did not know what to do with it. However, Chevron has always had close ties with the Gordon Conferences. One of the newest conferences was about statistics, so they sent senior manager Ed Chiswell to that conference in 1955 to find out about statisticians, who they were and what they did. Ed talked with and listened to everyone in sight and came back with the recommendation that they ought to get one.

Hence, they hired Scheffé as a consultant and gave him the octane blending (mixtures) problem to solve. If we mix 50% cat-cracked gasoline with 30% light straight run and 10% platformate and 10% alkylate, what will the octane number of the resulting blend be?

You know about the mixtures problem. If you do not, you should have attended John Cornell's short course on Tuesday. Scheffé did so well that Chevron decided to get a statistician full-time. Quite by chance, it was then that I applied. I spent four years at Chevron; for the last three, I was also a visiting assistant professor of statistics at Berkeley. In effect, I had a four-year post-doc with Scheffé, thanks to Chevron. My innocent timing was correct. If I had applied a year earlier, Chevron would not have had such a vacancy. If I had applied a year later, Scheffé would not have been there because he was on sabbatical at Princeton—about which more later.

I had thought that I might be about to do a lot of analysis of variance. In the eight years between Sampford's lectures at Oxford and my move to Chevron, there had been significant growth in the field. Statisticians, led by Churchill Eisenhart at the old National Bureau of Standards, had wrestled with the anova model and clarified the concepts of fixed and random factors. Two important books had appeared, and one was in the wings.

In 1952, Wiley brought out Kempthorne's book on "The Design and Analysis of Experiments". It used matrices, which was to me an enormous step forward in understanding. Scheffé called my attention to the 1954 book by Bennett and Franklin, "Statistical Analysis in Chemistry and the Chemical Industry', with their rules for calculating the expected values of the mean squares in complex anova layouts. Carl Bennett was a chairman of the Gordon Conference back then.

Most important to me was Scheffé's own book "The Analysis of Variance". In 1957, it had just become available in a mimeograph edition for local consumption at Berkeley. Wiley published it in 1959, and it became the authoritative work on the subject.

Yes, indeed, Chevron would be my chance to learn and use analysis of variance.

But that was not at all what Chevron's engineers wanted. They were not interested in four-factor experiments with A and B fixed and C and D random and D nested in B. That was agriculture, not their kind of research.

What they wanted was regression, which they called doing correlations. The arrival of the computer had made it possible to do least squares regression with more than three predictors in a reasonable amount of time. We programmed our little Datatron in the lab to handle up to nine variables. This was amazing progress. Before that, one had to invert matrices by hand.

Can you imagine trying to invert a 10 x 10 matrix, albeit symmetric, on a desktop calculator, even with a register a yard wide? With the big computer in the city, we could handle twenty or thirty variables and, before long, do stepwise regression using Efroymson's algorithm. It does not seem like much now, but it was really exciting then. (Remember—the first edition of Draper and Smith's book on regression did not appear until ten years later.) Among our favorite applications of regression was the blending problem: we fitted Scheffé's quadratic model to blends of several constituent gasolines hoping to find a model that could help schedule the refinery's production.

I could perhaps have continued doing regression, and being applauded for it, for years except for one important thing. The real excitement amongst chemical engineers in those days was response surfaces. They were not yet using them at Chevron, but they had heard the news. George Box had joined John Tukey's Statistical Techniques Research Group in Princeton—the Gauss House. George had students using his ideas at American Cyanamid, Du Pont, and at whatever Exxon was called in those days (Esso, or Enco, or

Humble?). He was one of the attractions for Scheffé at Princeton in 1958-59. I was sent to a short conference there in December 1958; it was my first exposure to the subject.

I have already mentioned the Gordon Conference on Statistics in Chemistry and Chemical Engineering, held each summer for a week at New Hampton School in New Hampshire. It was the annual camp meeting for statisticians at oil companies. We lonely people got together and talked and talked, free to ask one another ignorant questions that we dared not ask at home. We became a little less paranoid when we discovered that everybody else's engineers also thought that statisticians were a bit weird. We also agreed that most senior engineers back home were mossbacks, impervious to any new ideas that their limited backgrounds could not very easily absorb.

In those days Gordon Conferences were dominated by Box and his response surface methodology. The majority of those present had little training in statistics. They were chemical engineers who, like Chevron's Ed Chiswell, had been sent either to find out what this stuff was about or else with the simple instruction: "You are going to be the house statistics expert starting next month--go and get a week's training!" They did not know much about statistics or mathematics, but they did know a lot about chemical engineering, and Box's approach to process optimization was an exciting revelation to them.

My greatest payoffs from three successive years of Gordon Conferences were the intoxication of real intellectual ferment, the discovery of 2^n factorials, and the concept of sequential experimentation. I remembered that Sampford had mentioned the 2^3 factorial back in 1949, but it had not made any impression at that time. Now I was able to use 2^n factorials right away at Chevron.

This brings me to another book that was vitally important to me. I learned more about experimental design from it than from all the others. It is "Design and Analysis of Industrial Experiments", written by the group at Imperial Chemicals Industries under the editorship of my fellow countryman, Owen Davies. Box had been a member of that group in the dyestuffs division at Macclesfield. They wrote two books: Little Davies and Big Davies. This was "Big Davies". It is far and away the least mathematical of the books that I have mentioned. But, with its help, I finally began to understand 2^n experiments from two points of view: that of the experimenter and also that of the old-fashioned mathematician with no knowledge of engineering, but comfortable with subgroups and cosets and hyperplanes.

Next, a tenured appointment turned up in the mathematics department at the University of California in Davis, arguably the premier agricultural campus in the country. The University hierarchy had decided to expand U.C. Davis, and there I was on the doorstep, fifty miles down the road at Berkeley! I now had a chance to learn about different applications--almost completely different from those of the engineers.

I came to realize a fundamental difference between agricultural experiments and engineering experiments. The agronomist has to plant his whole experiment in the spring and harvest all the data in the fall. He has one shot and plants are relatively cheap. He must cover all possibilities in that one shot. The engineer can run a small experiment in maybe a week, get the results back, and think about them. Then he can use that information to design and run a second small experiment, repeating the cycle again and again. That is what I mean by sequential experimentation.

Six years on the Davis faculty provided a rewarding experience of consultation with scientists in agronomy, genetics, food sciences....

Then 1967 brought a surprise offer of a joint appointment in mathematics and engineering at UT Austin. Having been in academia long enough to be wary of working under two deans, I rejected that, but accepted a subsequent offer of a professorship in mathematics, believing that they would build up statistics. But by the time I arrived, they had changed chairmen.

I had unwittingly moved to a mathematics department that was so famously hostile to applications that, years earlier, it had driven native Texan Sam Wilks to Iowa, where he became a statistician. From there he went on to distinguished national leadership at Princeton. My guardian serendipity napped for a few years. I finished my book on "Statistical Design and Analysis of Experiments" in isolation.

In 1970 I accepted the offer of a professorship in statistics at Kentucky, taking a year's leave from UT to test that situation. That was my address when the book was published, but the schools there proved so woefully inadequate for our two children, then on the verge of high school, that we returned to Texas.

That might have been the end of my Odyssey, because the rest of the U.S. has the great misfortune to be a long way from Austin, Texas. Although I turned to research on incomplete block designs, I remained in touch with the Gordon Conference, and served as chairman in 1976. Nevertheless, it seemed as if design of experiments had, rather like my serendipity, gone to sleep. A famous applied mathematician, writing a report about our mathematics department at Austin, dismissed me as "working in classical design of experiments, a topic that is no longer an active field of research."

Happily, after a few years rest, my serendipity woke up with a phone call from an old Chevron and Gordon Conference friend, Ed Jones, who is now a professor at Corpus Christi.

Almost unbeknownst to many academic statisticians, the turn towards quality assurance in the late nineteen eighties had sparked interest in experimental design. Manufacturing engineers in the semiconductor industry did not know about Yates or Fisher, but they had heard about Taguchi. Ed had a company providing short courses on experimental design to semiconductor companies. Could I help out when he was double-booked? So there I was again, back working with engineers--not in the petroleum industry, but in the brand new field of semiconductors. They were using 2^n factorials and, sometimes, response surface methodology. They had questions about moving beyond the simplistic Taguchi lattices, and I had new applications about which to learn. That led me to write a book, "Statistical Methods in Engineering and Quality Assurance," published by Wiley in 1990.

After I had become involved with manufacturing engineers, the Fates struck again. THEY (and I am still not quite sure who THEY were) decided to build a research consortium called Sematech in Austin, just twenty-five minutes drive from my house. Sematech brought Paul Tobias, Jack Reese, and my former student, Veronica Czitrom. Their statistics group asked me to spend part of my time working with them, and that was indeed a rewarding cooperation for me.

What was the first thing that they asked me to do? Investigate analysis of variance with mixed, fixed, and random models and nesting. Only they called it gauge studies, not agronomy!

There ends the geographical part of the odyssey. Serendipity all the way through. No Greek, just in time for the new Oxford statistics degree, a job at Oklahoma, the coincidence of timing with Chevron, Scheffé and Berkeley, the timely expansion of U.C. Davis, semiconductors, and finally the appearance, de novo, of Sematech and my friends there. Many fine people along the way.

How about serendipity in research?

One of the fundamental differences between the academic and industrial lives is that academics have much more free time. We do not have to account for each hour of the so-called 8 to 5 working day. We are expected to be doing research but, by and large, it is on topics of our own choosing. We have no responsibility to work with our colleagues unless we want to. I had used part of this free time to finish my book on analysis of variance. But then what? How do we find topics?

Henry Scheffé used to say that consulting gave him a wonderful source of problems. He worked a great deal with his good friend, Cuthbert Daniel, who knew more about industrial applications of experimental design than anyone else. Most of us in industrial positions are more than busy enough doing what are for us relatively routine things that are needed and appreciated by our friends whom we help. Henry argued that, if you are interested in research, there are new problems lurking everywhere. You can hardly avoid stumbling over one, or two, so long as you are inquisitive enough, persistent enough, or lucky enough, to spot them. Sometimes you can make a useful contribution, if you can find the time and energy. Sometimes not.

Let me give you two examples that stand out from my last year at Chevron. I was able to think more about them after I left. I had acquired a few clients who had faith in my witchcraft. One of them wanted to do a 2^3 factorial (eight points). But, to our surprise, when the raw material, which only cost about a dollar a run, arrived from Scotland, there was only enough for six runs. I knew what to do with eight runs. I knew what to do with four runs. But what could, or should, I do with six runs? There was a nice little problem.

Out of it came, as an extension, the resolution V design for four factors in twelve runs and finally, in 1962, the general theory of three-quarter replicates.

The other example occurred with an engineer who had been exposed to some statistics in graduate school. We agreed on a 2^4 factorial: sixteen runs in two weeks on the pilot plant (two runs a day with one day for set up and one day at the end for clean up). We both knew what we were supposed to do. We carefully randomized the order of the sixteen runs. However, the machinery broke down at the end of the first week, so he was only able to make eight runs. Eight runs chosen at random from sixteen are a mess. So why not design your experiment to give you insurance against early termination caused by Murphy or by an impatient vice-president? It would have been so much better if only we had not randomized the order. If only we had made the first eight points be one of the two resolution IV half replicates. We could also have chosen the next four points to make a twelve point fraction of resolution V, and, then, if all was going well, complete the full factorial.

Finally, two examples of serendipity from my days at UT. In the relative isolation there I had become interested in incomplete block designs—a mathematical subject of theoretical use in agriculture and, so far as I can tell, of little interest to engineers. I was suddenly invited to contribute a paper to a special volume to honor the eightieth birthday of Frank Yates. That invitation led me to combine Yates' two great contributions of 2^n designs and incomplete block designs into a study of how to run 2^n factorials in incomplete block designs. For example, how do you run a 2^3 factorial in eight blocks of three?

Second, Oscar Kempthorne asked me to give a talk about the growth of experimental design in engineering when the Statistical Laboratory at Iowa State celebrated its fiftieth anniversary in 1983. That led me to check through Technometrics, where I found a 1966 paper by two grand old men of industrial statistics, Cuthbert Daniel and Frank Wilcoxon, on running the points in 2^n factorials in an order that guarded against linear or quadratic time trends. The paper had come out just as I was winding up at Davis. It had been set aside to read later. But "later" did not come until Kempy phoned seventeen years afterwards and might never have come without that stroke of fortune. I have been working in that area ever since. It also took me to the work of two other Gordon conference alumni. Al Dickinson of Monsanto had worked on running the points in a 2^n factorial that minimized the number of times that factor levels had to be changed. It can be expensive to change the temperature on a plant because it takes a while for the plant to stabilize, so you do not want to do it too often.

Hubert Hill of Tennessee Eastman had been working on the problem longer than Daniel or Wilcoxon. He got the idea from a paper by Sir David Cox in 1952. Hubert's paper appeared in Technometrics in 1960.

Let me confess something to you. I heard Hubert talk to the Gordon Conference on that topic in about 1960. I got little or nothing out of it. It was good stuff, as the subsequent written version verifies, but I just was not ready for it at the time. I had no idea what he was driving at and no concept of why it might be important. Seventeen years later I realized what Hill and Daniel and Wilcoxon had been up to. How lucky I was to get a second chance!

So here, again, are my four examples:

- (i) three-quarter replicates;
- (ii) sequences for 2^n designs that enable the experimenter to salvage results in the case of early termination;
- (iii) 2^n factorials and incomplete block designs;
- (iv) trend-free sequences of 2^n and 3^n factorials.

The fourth, one might argue, is an obvious extension of the work of Daniel and Wilcoxon, but I would never have gotten to it without that phone call from Kempthorne.

The first two really illustrate more strongly one of the themes of today's talk. They came directly from industry. They were, as Henry Scheffé suggested, there waiting to be found. I found them with the help of some luck or by accident. They would not have come to mind if I had just been sitting in my study in my ivory tower.

So my odyssey has brought me around full circle, by way of agricultural experiments and incomplete block designs, back to my old engineering friends. As you can see, it has not been a charted course. On my way, I have picked up occasional gusts of wind that took me in one direction and friendly currents that took me in another direction. Sometimes my journey has taken me back to places at a later time when I had at last grown ready to appreciate them. I have been extremely fortunate.

Sometimes my students ask me such questions as "How will I ever find problems to work on? How will I find a chance to do research if I go into industry?" Some even go so far as to ask "if I go into industry, will I ever get to do any research?"

I cannot give a definitive answer to the last one. In the way it is phrased, a lot depends upon the philosophy of the company. Some are more research oriented than others.

I was lucky enough at Chevron to have a kind and inspiring mentor and also to have some time by day as well as by night to do the research. In some jobs there is no time by day.

Furthermore, it can be argued that, from a selfish perspective, nobody would have known any better if I had not spotted those two research problems.

But that is not the point. The real point is that those problems were there, a little below the surface, waiting to break through. So I tell the students what Scheffé told me.

There are all kinds of practical research problems lurking in industrial applications, just waiting to be found and solved. Don't trip over them. Keep your eyes open and spot them. Then pick them up and run with them. Some of them may be insoluble. But some of them can be a lot of fun—and useful too. And if you make it a practice, it comes easier.

I would remind those who are discouraged about not seeming to do any research that every time they help a colleague with a problem that is even slightly out of the ordinary and calls for a little head scratching on their parts, they are doing research (not very great perhaps) and making a contribution to knowledge. Maybe George Box or Cuthbert Daniel or Henry Scheffé could have solved the problem in a quarter of the time, but they were not there—just when you could have used them!

Have faith in luck and make it work for you.

Was it Gary Player who said, "The more I practice, the luckier I get"? Henry Scheffé was right. There are lots of problems out there waiting to be found. With a little bit of luck, you can find them—and then it is up to you. Remember

SERENDIPITY