Statistical Science
1994, Vol. 9, No. 3, 439–455

A Conversation with Sir David Cox

Nancy Reid

Abstract. David Roxbee Cox was born in Birmingham on July 15, 1924. He attended Handsworth Grammar School and St. John's College, Cambridge. From 1944 to 1946 he was employed at the Royal Aircraft Establishment, and from 1946 to 1950 he was employed at the Wool Industries Research Association in Leeds. He obtained his Ph.D. from the University of Leeds in 1949. He was an assistant lecturer at the University of Cambridge from 1950 to 1955, and then visited the United States for 15 months, mainly at the University of North Carolina. From 1956 to 1966 he was Reader and then Professor of Statistics at Birkbeck College, London, and from 1966 to 1988 was Professor of Statistics at Imperial College, London. In 1988 he moved to Oxford to become the Warden of Nuffield College, a post from which he retired on July 31, 1994. He is now an Honorary Fellow of Nuffield College and a member of the Department of Statistics at the University of Oxford. In 1947 he married Joyce Drummond. They have four children and two grandchildren.

Among his many honours, Sir David has received to date 10 honorary doctorates, an honorary fellowship from St. John's College, Cambridge, and honorary membership in four international academies. He has been awarded the Guy medals in Silver (1961) and Gold (1973) by the Royal Statistical Society. He was elected Fellow of the Royal Society of London in 1973 and was knighted in 1985. In 1990 he won the Kettering prize and gold medal for cancer research.

He has authored or coauthored over 200 papers and 15 books. A list of his publications through 1988 is included in Hinkley, Reid and Snell (1991). From 1966 through 1991 he was the editor of *Biometrika*. He has supervised, encouraged and collaborated with innumerable students, postdoctoral fellows and colleagues. He has served as president of the Bernoulli Society and the Royal Statistical Society, and he is president-elect of the International Statistical Institute.

This conversation took place in Sir David's office at Nuffield College on October 26 and 27, 1993.

WIRA AND CAMBRIDGE

Reid: I'd like to ask you about your work early in your career at the Wool Industries Research Association. What kind of a place was it, and what kind of position had you there?

Cox: Well, Henry Daniels has described it a bit in a recent interview (Whittle, 1993). It was a type of organization that was very common in the U.K. at that time, funded by government and by money obtained from a levy on the industry, to do basic research on problems to do with the industry; and it had at that time a remarkable director who simply had the idea that you get people and largely let them get on with it, with encouragement. I went there because with the previous job I had I'd worked on a problem to do with the strength of spot-welded joints and I just happened to read in the library a marvellous related paper by Henry Daniels. Then as I went out of the library there was an advertisement of a job to go and work with Henry. The war had just ended and I was free to move, so instead of going back to Cambridge to complete my BA, which I'd normally have done, I went to work with him.

In wool textiles, you see, you've got everything: from the biology and the nutrition of the sheep,

Nancy Reid is Professor of Statistics, Statistics Department, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 1A1.

through the chemistry and physics of various processes, to the operational research side, the engineering side and the economic side. So there was an enormous range of problems and extremely good people working there from whom I learned an enormous amount. I did some design of experiments and some analysis of data. People would come along with their split split plot experiments, or their analysis of covariance, and they'd say, "Oh, there's no hurry, tomorrow afternoon will do." I hadn't the remotest idea of how to analyse these things. There weren't really any books so one had to struggle to find out what to do. In addition I worked on more basic, longer term, problems. In fact I did a lot of work which was never published: more applied mathematics; theory of elasticity, large extension elasticity theory, things like that. I wasn't terribly good at it, but I did do a great deal of work.

Reid: Did you save all your papers?

Cox: Yes, they're around here somewhere. [Looking around vaguely.]

Reid: So the work in your paper "The theory of drafting wool slivers: I" [1], that was more or less part of your job.

Cox: Yes.

Reid: Plus no doubt a few evenings here and there.

Cox: [Laughs.]

Reid: Are parts 2 and 3 somewhere?

Cox: No; that's an important moral, you see, which I should have learned from Henry Daniels. His paper is "Theory of strength of bundles part 1," 1944, *Proceedings of the Royal Society*, and part 2 is still anxiously awaited.

Reid: Presumably you got your Ph.D. at the University of Leeds because it was nearby?

Cox: Right. At that time anybody who worked in the city of Leeds and had a degree from anywhere could register for a Ph.D.

Reid: Was there any particular advantage to you at that time to have a Ph.D.? Why did you go to the trouble?

Cox: I think Henry must have suggested it. I suspect (perhaps it's an arrogant thing to say but I don't mean it that way) that the director of research thought this would tie me to Leeds for a few years, which in a sense it did.

Reid: Was your "wool" paper [1] your dissertation for your Ph.D?

Cox: That was a bit of it, that was a chapter basically. There was also a chapter on long-range dependence, and various other things.

Reid: Your book on statistics and textiles [2] went to a fifth edition, in 1960.

Cox: Oh, but that was just an account of very elementary methods for quality control.

Reid: Yes, but it's very elegant. There was recently a suggestion that we should have a list of the good cookbooks, cookery books you call them over here, and I was thinking it should be on the list.

FIG. 1. *At Storrs University, Connecticut, February 1994, to be videotaped for the American Statistical Association's distinguished lecturer series.*

What about your time at Cambridge as a student: did you learn any statistics?

Cox: The short answer is no. Harold Jeffreys gave a short course which was intriguing, but almost totally incomprehensible. I sat in on a rather longer course that J. O. Irwin gave, where he was essentially reading Wishart's lecture notes because Wishart was away on war service. This calculated the moments about the origin, about the mean, the factorial moments, the cumulants and the factorial cumulants, of a considerable number of known distributions. It wasn't terribly inspiring. Although in a certain way it was a very systematic course, and there was more to it than I just described.

During the war people came round to the universities and decided whether you went into the army or air force or went to a research establishment. There was an enormous shortage of statisticians, and the notion was that anyone doing a mathematics degree and doing reasonably well knew something about statistics or could learn it very quickly—a totally false assumption. I was sent as a statistician a' though I didn't know any statistics.

Reid: So you were sent to the Royal Aircraft Es tablishment as part of your draft work or war commission. How is it that you have an MA after just two years as an undergraduate?

Cox: Under the war regulations, you see, you could get a degree in two years. Then Oxford and Cambridge have this strange system: once you've got a Bachelor's degree you can get a Master's degree automatically a certain number of years after your BA.

Reid: What kind of problems did you work on in the Royal Aircraft Establishment?

Cox: I was in a department of structural and mechanical engineering: it was mostly strengths of materials. They did testing components of aircraft and to a certain extent testing whole pieces of aircraft: they used to break aircraft wings. But there were also miscellaneous other things like where German rockets landed, and a big thing on aircraft accident rates.

Reid: Were there other statisticians working there?

Cox: No.

Reid: The opportunities that you had to be so involved with applications in these early years seem to me to have influenced all your later work.

Cox: I was involved in all these applications but it wasn't what I was interested in. I was interested in mathematics. I wanted to be either an analyst or possibly a mathematical physicist—I didn't want to be a statistician. I mean I don't regret being one, but that was what I was interested in, for quite a while.

Reid: Where did that come from, that interest?

Cox: This is part of the Cambridge tradition; and I went to lectures by several great mathematicians. On the whole not very good lecturers, but inspiring.

Reid: Did you always have a notion that you would get this practical work out of the way and get back to mathematical physics?

Cox: Oh, as I got older, I got more and more interested in applications. But as late as 1955, certainly as late as 1950, I would have still seriously considered giving up statistics. Partly because the career opportunities in statistics at that time seemed terrible. There was this burst of activity during the war and immediately following, but at one point there were only two or three professorships of statistics in the whole of the country and the possibility of ever becoming a full professor in the university system seemed very remote indeed.

Reid: Was that your goal at that time?

Cox: [Laughs.] No, no, not at all. The goal was to survive, get enough to eat, somewhere to live; survive. No, no, no: that wasn't a goal at all. I never really thought much about goals. Except one thing I would really like to do is to make a contribution to probability and the foundations of quantum mechanics. But I don't think I'll ever do that. Partly because it's too difficult. Primarily because it's too difficult.

Reid: You left the Wool Industries Research Association to go to Cambridge as a lecturer?

Cox: Assistant lecturer.

Reid: What was Cambridge like then, going back as a lecturer?

Cox: Scientifically, fantastic. Personally, terrible.

Reid: Please elaborate.

Cox: Well, first of all, scientifically, I mean, marvellous, absolutely first class students; and excellent colleagues: Wishart, who was not really very active scientifically, was a very good director of the laboratory. Daniels, Anscombe, Lindley, and then lots of visitors. From that point it was absolutely marvellous. Personally, well, one was very badly paid. My wife and I had two young children and certainly couldn't afford a car, couldn't really heat the house properly and so on. I had a post that was limited to five years and that was unsettling and very discouraging.

Reid: How did you come to have such a terminal post?

Cox: That's all there was available.

Reid: Even for the best and the brightest?

Cox: Basically any post that became available in mathematics that I might have been appointed to

FIG. 2. *This portrait in bronze by Martin Jennings was commissioned by the fellows of Nuffield College, Oxford, and is displayed in the Senior Common Room of the College.*

would have been filled by a pure mathematician. I didn't take it personally, but it was just unsettling. At the end of the five years it seemed I had really no option but to go to the United States, which was an extremely fortunate thing to do.

As I say, scientifically it was fantastic and yet going to the United States was a revelation. Not just the physical standard of living, but the enthusiasm of the people was very encouraging. There were even people that had read papers that I wrote. For example, I gave a lecture at Princeton on conditioning; if I'd given that lecture in London, oh, I don't mean people wouldn't have been interested, pleasant and encouraging, but there wouldn't have been any sense of vigour.

Reid: Yes, I think I know what you mean.

Cox: You know what I mean I'm sure; and it was more noticeable in those days. Without going into the sociology of it, I think the attitude in the United States then was that everything is possible. Now a sort of British cynicism is to some extent apparent and people in the United States see the difficulties rather more than they did then.

Reid: It was economically a tremendously rich time for the United States.

Cox: Yes, that's right; and people were enormously helpful and pleasant and encouraging.

Reid: You mentioned once that you might have stayed.

Cox: Yes, very easily.

Reid: How close did you come?

Cox: Oh, very close.

Reid: And how did you come to come back?

Cox: Well, suddenly there started to be all sorts of jobs appearing. I had accepted a job in the U.S. and then was offered the Birkbeck job, which was in many ways very nice indeed. For the next few years, I had a succession of very tempting offers in the U.S. to which Joyce and I could have very easily succumbed—which would have been marvellous, I'm sure. I don't in any way whatever regret staying in the U.K., but I'm sure I'd have been very happy in many places in the United States.

CONDITIONAL INFERENCE

Reid: Your 1958 paper on conditioning [6] was presented when you were in North Carolina?

Cox: Well, it was actually presented in Princeton. I was visiting North Carolina at the time.

Reid: Was it considered controversial by the audience, do you remember?

Cox: Yes. It's the only occasion I've ever given a lecture where people came up and were still talking to me about two and a half hours later—Allan Birnbaum in particular.

Reid: Talking about?

Cox: The implications of conditioning, although there's a lot more in the paper than the conditioning, things about the difference between p-values and tests of hypotheses, for instance.

Reid: Your weighing machine example introduced in that paper is possibly the only thing about conditional inference that everybody at least thinks they understand. Why is conditional inference so difficult?

Cox: Well, there are two aspects, aren't there: is it conceptually difficult, and is it mathematically difficult. It seems to me that conceptually it's not difficult, it's a very clear consequence of wanting to make long-run probability calculations relevant to the interpretation of sets of data. Of using physical probability epistemologically. Of calculating confidence coefficients and significance levels whose interpretation is based on long-run frequency but which you want to be relevant to a particular set of data. How does the long run become relevant to a particular set of data? Well, by being suitably conditioned. The arguments for this seem to me absolutely overwhelming; but to convert that idea into definitions, formulae, algorithms and so forth, then it gets much more difficult. I think that's the point at which people find it hard going. I find it hard going. [Pause.] Yet it's strange, isn't it, that an enormous number of people must learn about statistics and perhaps even do Ph.D.'s in statistics and not think about this at all.

Reid: Well, I guess that might partly be the American training.

Cox: Oh, I don't think it's as nationalistic as that.

Reid: Was your paper written to make a statement against decision theoretic formulations?

Cox: I think partly, yes. I don't exactly recall how it came about but the background, more or less, is that in the 1950's, particularly in Cambridge, there was intense interest in these philosophical issues. Fisher gave three famous public lectures (well, one was a vote of thanks to somebody else and the two others were lectures) in which he put forward the ideas that to some extent appeared in his last book. There was an enormous amount of discussion of this. Of course Don Fraser was active in these things at that time as well. Then I got to the United States and I was invited to give I think it was a special IMS lecture, or one of these things, and I hadn't the remotest idea of what to talk about. I thought, well, perhaps this would be a good topic. With the weighing machine example, you see, I was trying to reduce the argument to its absolutely simplest case: it was a pedagogical example to try and put the argument in its very simplest form.

Reid: And when you get to the examples that are somehow not so clear-cut, getting back to your comment that you want the long-run frequency to be relevant to the data at hand: it seems terribly difficult to mathematize that.

Cox: Yes.

Reid: Is that just the state of affairs or are we all missing something?

Cox: Oh, I expect we're all missing something, but I don't know what it is. [Laughs.]

Another aspect is to minimize differences from Bayesian arguments, particularly with standardized priors, which is how I first learned about statistics in Jeffreys' lectures. I feel the differences between the various schools of inference are emphasized too much and the similarities not enough. Bayesians will achieve this conditioning automatically, so as compared with unconditional Neyman–Pearson, say, conditional inference is going some step towards Bayesian conditioning without having to bring in priors.

Reid: And is that the right direction, towards Bayesian inference?

Cox: Well, yes; but in talking about Bayes, I feel one has to distinguish very sharply personalistic priors from standardized priors. Personalistic priors have a role, if you are strongly interested in the aspects of how personal judgment enters into analysis, but most of the problems I look at are not like that. To me Bayes with some sort of reference prior seems quite appealing, although I don't regard it as the absolutely ultimate criterion. All these things are essentially measuring devices for measuring how much information there is in data, and you test a measuring device by seeing how it works when you use it. From that point of view the absolutely ultimate criterion must be some sort of notion of probability of correctness.

Reid: In a long-run frequency sense.

Cox: Well, in some sense—yes, in a frequency sense: that if hypothetically we were to use this procedure again and again, then its properties would be reasonable. That's much weaker than saying, you know, 95% coverage is all that matters. But it is saying if you had a procedure that in hypothetical repetitions did badly, it can't be a good procedure. If there is an ultimate test, it is that.

Reid: Could I come back to the talk that you gave, that the paper is based on—you said it generated a lot of discussion. Was it well-received and was the discussion friendly but puzzled or was it hostile?

Cox: It was in no sense hostile; and it wasn't discussion, it was six people gathered around a blackboard after the lecture. It wasn't discussion in anything like the normal sense of a scientific meeting. For one thing, for some strange reason it was held in the evening, something like seven o'clock in the evening.

Reid: Do you remember who the six people were?

Cox: Not very clearly. I know that Arthur Dempster was there, I can't remember whether he took part in this discussion. Allan Birnbaum was the main person. [Laughs.] I'm probably offending somebody by leaving them out—they were all people who then and now are quite well known. It's a long time ago. But in no sense was it a hostile discussion.

Reid: How long were you in the U.S. at that time?

Cox: Fifteen months—Princeton, North Carolina and Berkeley.

Reid: It must have been a busy 15 months.

Cox: Yes, and very exciting.

Reid: Who were the people you interacted with the most?

Cox: Well, John Tukey very particularly, John perhaps more than anybody and very intensively. Martin Wilk arrived in Princeton the same day that Joyce and I did. He had just finished at Iowa State, and I had many discussions with Martin. Bernard Greenberg, who was the head of the Biostatistics group at Chapel Hill, I couldn't say I had many technical discussions with him, but he was enormously helpful to me. At Berkeley, I don't remember scientific discussions, with particular people there, although I took part in all sorts of seminars. I suppose, if anybody, it would have been Mr. Neyman.

Reid: Your whole approach is so different than the American school of the fifties.

Cox: Yes. [Laughs.] I can remember Mr. Neyman telling me off, in a very nice way. I gave a talk in his seminar about some problem in stochastic processes and I used Dirac delta functions and he came up to me afterwards and said, "Yes, that really was quite interesting, but we don't do that sort of mathematics here." But the irony was that a few years later Laurent Schwartz visited Berkeley and, of course, after that it was perfectly respectable. But Mr. Neyman was very nice about it.

[Schwartz's books (Schwartz, 1950, 1951) established a mathematical foundation for the study of Dirac delta functions.]

Reid: Was your 1958 paper your first paper on pure theoretical statistics, so to speak?

Cox: I'm not quite sure, but I think it must have been. Yes, and forced on me by having nothing better to talk about.

STOCHASTIC PROCESSES

Reid: Could we turn to your 1955 paper that you read to the RSS [3]? That was the work you'd done at WIRA and Cambridge?

Cox: Well, some of it was done in the period from 1946 to 1950, and then it continued on at a lower level of intensity between 1950 and 1954. The trou-

FIG. 3. *With Nancy Reid and David Hinkley at CIMAT (Centro de Investigaćion en Matemáticas), Guanajuato, Mexico, March* 1993.

ble with the paper is that there is far too much in it. Doubly stochastic Poisson processes, all sorts of tests to do with empirical series of point events, a certain amount about unbalanced variance components, a certain amount about overdispersion in Poisson models and quite a bit about some of the peculiar sampling problems that come up in those sorts of studies. That makes it all a bit of a mish-mash. There was even more in the original version. In those days you didn't submit something to the RSS for reading until you really felt you had spent a long time on it. That was the idea I picked up anyway; I don't know how true it really was.

Reid: Almost all your early work that wasn't in design was in stochastic processes; you have several books [9, 10, 12, 13] following on from the 1955 paper [3].

Cox: Well, in some ways preceding the 1955 paper. When I first went to work with Henry Daniels he said the up-and-coming subject of the next umpteen years is stochastic processes, and he even arranged for me to go across from Leeds to Manchester every other week to listen to Maurice Bartlett's lectures. Leeds to Manchester is about 70 kilometers but in those days that was a major journey, like going to the North Pole, particularly going from Yorkshire to Lancashire. Then he gave me various suggested readings: Chandrasekhar's famous paper [Chandrasekhar, 1943] and S. O. Rice's papers [Rice, 1944, 1945] and Bartlett. There wasn't much else, you see. It was the idea then that if you're a statistician, you'd jolly well better be interested in stochastic processes. So I had never thought of stochastic processes as separate from statistics. Just as I don't think of time series as separate from stochastic processes or from statistics.

Now, there were various problems in my thesis about time series and stochastic processes. There was a textile problem of a queuing kind that led to my interest in queuing theory. So I was quite interested in stochastic processes by the time I went to Cambridge in 1950, and that interest has continued. In those days, while it was a difficult subject to work in, it wasn't highly technical. I mean, you needed to know some matrix algebra and some differential equations and preferably some partial differential equations and Laplace transforms—the standard elementary techniques of mathematical physics—and I did know that, that was what my training was. Of course, if you were like Henry and were a wizard at saddlepoints and so forth, all the better. All the same the subject wasn't highly technical. You looked at a particular scientific problem, you saw if you could formulate it somehow or other as a Markov process and you set up your differential equations and you had a go at solving them. You had better solve them analytically because unless they were very simple, solving them numerically would have been a bit of a pain. It's the sort of mathematics I like doing and the concepts are fairly straightforward. Some of the more theoretical work on stochastic processes that I did in a couple of papers in the *Proceedings of the Cambridge Philosophical Society* [4, 5] were basically about how to take non-Markov processes and build them into Markov processes.

It seems to me that if you look now at even relatively elementary books on stochastic processes, they require very heavy apparatus. Things like Itô calculus and so forth that seem to demand a lot of what I would regard as heavy mathematical background. I'm not convinced they really do need all that paraphernalia—that one couldn't study stochastic processes more in the spirit of Bartlett's great masterpiece [Bartlett, 1955], which is difficult reading but not because of an overelaborate mathematical formalism. Anyway, I've always been interested in stochastic processes from that point of view, and at that mathematical level. Nowadays I feel out of touch. If I went to one of the meetings of the Stochastic Processes Group of the Bernoulli Society, I suspect I wouldn't understand anything. It's all so dressed up in this great apparatus. Which has a beauty of its own, I'm not debunking it in any sense, it's great stuff, but is it really necessary for looking at scientific problems, if that's what one wants to do. I don't know.

The main area that I have continued with is work that I've done primarily with Valerie Isham. We addressed a long-standing problem connected with what is called covariance counting in physics [18].

It's a method of counting particles which depends upon looking at correlation between numbers of particles observed in overlapping periods. It's one of these sort of blocking problems. There had been various attempts to solve it which hadn't worked.

And then more recently we've worked with hydrologists on marked point process models for rainfall [21–23]. There again we used the traditional kind of mathematics and my book with Valerie on point processes [19], while we range over a pretty wide variety of models and there are various new things in there, I think, it's not done by bringing a vast mathematical armoury to bear.

And there is a close connection between all that and the proportional hazards model. Particularly in point processes one naturally thinks of saying, "Here we are now, and we've got a certain history; what's the probability of an event in the next short element of time?" Now the mathematicians make a great hoo-ha about setting up such a function, but it's physically absolutely obvious that such a thing uniquely defines a point process and it takes half a sentence to say so and Bartlett said so many, many years ago. That's the kind of thing that rather worries me; the notion that you can't just set up a complete intensity function or whatever you like to call it without a great mathematical paraphernalia of filtrations, and this and that. If you want a very general mathematical theory, it's clear that's the way to do it; but all I'm saying is that it should be legitimate to treat the subject at this more informal level as well. I'm not saying that it isn't valuable to have the very general theory; I think it is, but not everybody has to go that way. And if it means that you get a situation where lots of statisticians don't know anything about stochastic processes because they haven't the time to master all the elaborate apparatus, then that's bad news.

Reid: Could I go back in time and ask things I thought of when you were talking earlier? For example, you went to Bartlett's lectures; were they enlightening?

Cox: That's a very difficult question. I suspect they were highly enlightening. They were not particularly easy. I mean, my admiration for his work is enormous. I think his book's a great masterpiece and I'm shocked at how few people nowadays who call themselves experts in stochastic processes have read it. I'd have thought it was one of the most important books in our field in the last 50 years: the number of ideas per page is incredible.

Reid: And how did you get interested in the rainfall work, in particular?

Cox: Very soon after I went to Imperial College I had some contact with hydrologists and that led to a Ph.D. student, Gideon Weiss, who did some nice hydrological work on runoff modeling. Beyond this sort of vague interest I didn't do anything very much until I went once to Caracas, really to meet someone else. I met there a famous hydrologist, Ignacio Rodriguez-Iturbe, who had worked on point process models in rainfall. We developed a collaboration following from this. It really was a kind of accident arising out of visiting Caracas.

Reid: A happy accident.

Cox: Yes, indeed extremely.

Reid: There's an enormous literature on rainfall in the physical literature. How did you avoid being swamped by that and just getting on with something new?

Cox: Well, really by drawing on Ignacio's deep knowledge of that. I mean I have talked to meteorologists and hydrologists a bit as well, but it's largely Ignacio.

Reid: Are you still working on it?

Cox: Yes, but it goes very slowly. I'm working with Valerie on it; in fact we gave a paper at a conference three or four weeks ago on extending the sort of models we had in the earlier papers to several sites.

PLANNING OF EXPERIMENTS

Reid: Design of experiments was another part of your work at the Wool Industries Research Association.

Cox: Yes, there was a fair amount of use of balanced designs in the textile industry. Henry Daniels had persuaded a lot of people that this was more or less inevitable. So I was interested while I was in Leeds, and when I went to Cambridge I was also involved with agriculturists and others. I also taught a course in design for two or three years. I had an interest in both the practical and theoretical side.

Reid: Had you ever had to design an experiment?

Cox: Oh yes many. Not recently, unhappily. It's the most interesting side of statistics in some ways.

Reid: Is *Planning of Experiments* [7] based on your lecture notes?

Cox: No, not at all. The lecture notes are mathematical, about things like Galois fields and combinatorics and the derivations of various standard designs, fractional replication and so on.

Reid: I guess you would have had to be quite mathematical in the mathematics department.

Cox: Not necessarily, although the students would have expected a fairly mathematical treatment. It was something like 24 hours of lectures, and the pace of lecturing in Cambridge in those days was very fast. Some people could do it by writing on a blackboard with their right hand and rubbing off with their left hand while they were talking about a third thing. I could never master that technique, but I lectured pretty quickly.

Reid: What was the context for writing the book?

Cox: I'm not absolutely sure, but I think I had the sort of vague idea of writing up something like the lecture notes as a theoretical book on design of experiments. Then I decided for some reason that it would be much more difficult and much more useful to write something that was aimed at scientists with a minimum of technical statistical analysis, and so I tried to do that. In particular there are very few formulae, and the handful of formulae are written in words. You could call it a gimmicky thing, if you like, but it was meant as a way of saying no working scientist should be put off by the fear that they don't know enough mathematics. And yet at the same time it's an attempt to explain things, not to just say dogmatically, "Randomize," but to explain why you should randomize.

Reid: It's a lovely book, and quite unusual now: was it pretty unusual at the time as well?

Cox: Yes. It isn't a textbook, you see. In fact in a certain way none of the books I've written are textbooks. None of the books I've written are other than extremely indirectly based on lectures, they're not really meant to teach courses from. They are attempts to write down a subject that I've thought about for a while, as it seems to me, or whoever I've been working with, and I've been very very fortunate in the people I've worked with. Maybe it would have been better if they had been textbooks, but they aren't. The publishers may pretend they are, but they aren't.

Reid: Even *Theoretical Statistics* [16]?

Cox: Well, that's a slight exception. I used to give 12 hours of lectures on that material and cover quite a lot of it. [Laughs.] So you could say it grew out of some lectures, but only in the most tenuous sense; and it's not an easy book to teach from, I would imagine.

Reid: I can attest it's not an easy book to teach from.

Cox: I've never attempted it; I wouldn't dream of doing so. The publishers ask for comments on books and David and I did get the comment that the book should be totally rewritten in the form of theorems and proofs. I'm afraid I sent a rather sour answer to that, which I hope reached the stupid person who made that remark. I mean, I understand why the remark was made, but all the same...

Reid: Just to get back to *Planning of Experiments*: do you remember how it was received at the time?

Cox: I don't really remember how it was received. Most of the reviews were reasonably favourable. I do remember one by a mathematician that said something like, "This book is on the whole a competent account of principles of experimental design for non-mathematicians. What a pity that biologists and others don't learn more mathematics so that such

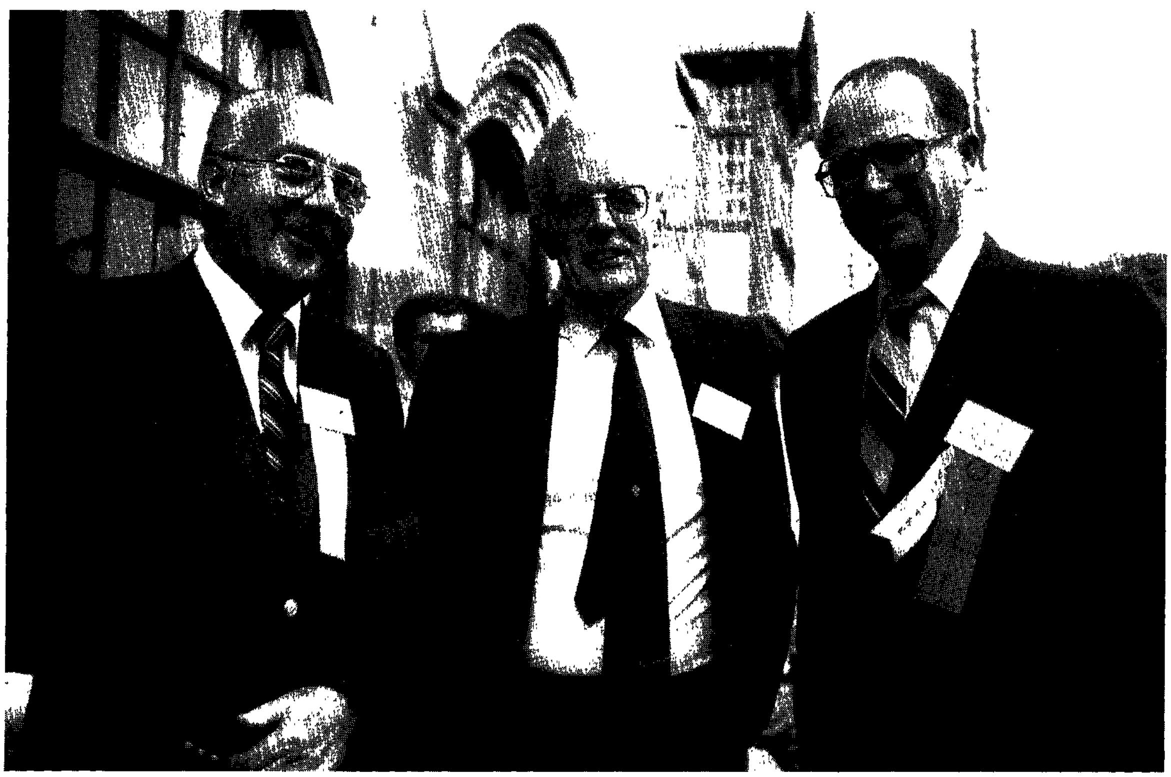

FIG. 4. *With George Box and Stu Hunter in front of Meadow Brook Hall, Oakland University, Rochester, Michigan, in June* 1986.

books would be unnecessary." I thought that missed the point rather dramatically. This is an underlying point about all sorts of statistical issues: to what extent these matters are really mathematical, and to what extent the important things are concepts which then have to be translated into mathematics. The other thing about the book is the pattern of sales was very curious because it sold extremely well, by my perception (Wiley might think differently), for almost 30 years and then suddenly the sales dropped off. So that would suggest that the reception was reasonable.

Reid: Perhaps what the reviewer meant was that there are occasions when it is nicer to see a formula than to see a formula written in words.

Cox: On that particular bit yes of course, but I don't think that's what he meant. No, I think he meant that we have a book on the mathematical theory of design of experiments and that's it, and everything else is deduction from this.

Reid: You wanted to talk about split plots.

Cox: Partly in connection with *Planning of Experiments*. One of the advantages of writing in a non-mathematical style is that it forces one to explain, in qualitative terms, absolutely crucial, vital issues which everybody has to understand about split plot experiments. Why certain contrasts have different errors from other contrasts. This isn't just a question of deducing it from some model that one's plucked out of the air, but in seeing why physically this is so. And then once you've mastered the split plot experiment of course you can go on to the split split plot, or the split split-split-split-split plot, with no further problems other than a certain messiness.

But the importance of it goes beyond design of experiments because it does bring home in its simplest form the point about hierarchical error structures. There is a tendency if you see random variation on different individuals, to model it by iid random variables. The split plot is a warning against this: there may well be correlational structures or hierarchical structures in the error which mean that some comparisons have quite different precision than others. That's a point of importance far beyond just the classical design of experiments. Although one could have a working life and never come across a split plot design, there are other areas where it's a wholly natural design and there are other areas where it's being used implicitly, without people quite understanding that this is what is being done, and getting incorrect estimates of precision. So I think the moral is that these very important, absolutely central issues, like the split plot, have to be understood, and it's more important to understand them at a qualitative level than it is to plough through a lot of algebra.

Reid: And some areas where it comes up all the time?

Cox: In all sorts of contexts where you can have repeat measurements, whether it's an experiment or observational study: agriculture, of course, where it originated; any multistage industrial process or physical process where let's say, simplest case, you have two stages of processing some material. You are interested in a factorial experiment where you make some changes in stage A, and some changes in stage B and it may be entirely natural to have a design for your stage A material, a stage A treatment that's processed and produces some material that is then divided into parts and processed in a different way in stage B. It may be almost necessary that way.

Then there's the possibility that you have factorial experiments where certain kinds of adjustment to a process or a system are extremely difficult or expensive to make and other adjustments are very easy to make. Then it would be natural to hold the expensive factor fixed for a certain length of time, have a number of runs in which the inexpensive factor is varied and then switch your expensive factor and do some more. Certain areas of experimental psychology as well. Mostly when it does arise, it arises inevitably from the nature of the constraints on the way the experiment has to be set up. I don't think it's very often imposed. It's more often that the physical constraints of the experiment dictate that a split plot structure is used.

Reid: Perhaps it needs a better name.

Cox: I used to think it didn't matter what things were called, but that's a very mistaken view; there is a certain skill in naming things and methods in a way that will make them attractive. Well, I've had my say about split plots and even split split-split plots.

Reid: I was wondering what your thoughts are on sequential design and analysis?

Cox: Perhaps I could ramble on a bit about my interest in sequential analysis. I first came across this when I was at the Royal Aircraft Establishment and Wald's secret report was circulated to government research scientists. And in the department of structural and mechanical engineering, people were very interested, because the claim was that you could halve the amount of testing. One of the really expensive things they did was to take whole aircraft wings and load them until they broke, measuring all sorts of things, strain gauges all over the place. Being very good engineers they'd always break two aircraft wings, which was a very very expensive business, and I can remember one of the heads of a group saying to me, "Perhaps this means we need only break one aircraft wing." He took a bit of persuading that this wasn't quite what the implication of Wald's work was.

Also the first Royal Statistical Society meeting I ever went to was during the war when George Barnard spoke. He developed something very similar to Wald's work, but put very differently, which was quite widely used in some of munitions factories. So when I went to WIRA, I had this interest in sequential sampling, I did consciously look around for applications there, and the people there were very receptive to new ideas, which is not always how people are. There were one or two things where we tried some simple sequential tests, but it was pretty clear it wasn't actually at all effective. What seems to have happened very recently is that medical statisticians have become extremely interested in the subject, partly, I think, because of ethical questions about not continuing trials with major life-threatening events involved any longer than necessary. So the subject is having a very interesting renewal.

Why was the previous work pretty ineffective? Too slanted to simple decisions like accept/reject (which is after all what it was set up for) rather than measuring things, estimating things. And when you come to sequential estimation, what the sequential method does, and it can be very important, is adjust the sample size to the requirement of the precision that you want. It's not a question exactly of saving observations, but more a question of getting the precision that you need.

Reid: When you said that recently there's been quite a bit of development in medical applications, do you mean in the last 15 to 20 years, or in the last 5 with the AIDS crisis?

Cox: No, AIDS is just one aspect of it. Work on group sequential sampling seems to be being used, and very interesting; and adaptive randomization.

But even in a major clinical trial, where sequential stopping may be a very relevant and contentious issue, the most you can look for is reasonable advice as to when one should think seriously about stopping. I think a decision about stopping is always going to be taken on grounds that are statistical and ethical, involving prior knowledge and perhaps even political considerations, quite properly. And for a statistician to come up with a procedure to simply say you stop now or you must continue, is a misformulation.

BINARY DATA

Reid: I'd like to ask you about binary data. I think that your first binary data paper was 1958 [8] and the conditioning paper was 1958, so you did have something else to talk about at the IMS meeting.

Cox: Well, the IMS meeting was in 1956, you see. The *Annals* didn't actually leap at the thought of publishing this paper.

Reid: Really, you went several rounds with the editors, did you?

Cox: Well, if you look at other papers in the *Annals* at that time, this is not in the conventional *Annals* style at all. I don't mean they were fractious or unreasonable, but it took a while.

I think the binary data work was partly motivated by things in Chapel Hill, but more particularly by going to Birkbeck College, London. There was a very strong psychology department, mainly experimental psychology, and they were very concerned about analysing binary data.

Reid: And what was available before your 1958 paper?

Cox: I'm very unclear about the history. Analysis of binary data I think was often thought of as chi-squared basically. You find observed frequencies and you find fitted frequencies and you compute observed minus fitted squared over fitted; you forced it into that mold. Now that isn't totally true, because contrasted very much with that was Finney's work on probit analysis. And I think historically, I don't know who first did logistic regression, but quite possibly it was Jerry Cornfield on the Framingham study a few years before 1958. Possibly Nathan Mantel would have been involved. I certainly met Cornfield, so whether I was influenced by him is entirely possible, but I don't remember. What I was trying to do was set out a systematic framework for analysing binary data that is somewhat analogous to least squares regression.

Reid: Which you did, beautifully. I was a little bit surprised when I read the paper to see you derive in detail conditioning on a sufficient statistic for the nuisance parameter in a canonical exponential family; would that have been relatively new at the time?

Cox: Well, I thought it was new, and the motivation was new but of course as probably somebody pointed out in the discussion, it was regions of Neyman structure; from the Neyman–Pearson point of view it would have been orthodox. But the justification I gave for it was not a Neyman–Pearson justification. It was Fisherian in the sense of saying this part of the sufficient statistic doesn't tell us about the parameter of interest and so we'll condition on it. It's a nice question, I'm very unclear about it, as to whether that isn't really a better justification for conditioning than achieving exact similarity.

Reid: It's still not clear to me why you're so sure that there's no information about the odds ratio in the distribution of the marginal total.

Cox: Well, I'm not sure. Efforts still continue to define that elusive notion. [Laughs.]

Reid: That was a read paper, to the Royal Statistical Society. Was it your impression that it was well received? Those discussions to me always sound terribly critical: what was your feeling at the time?

Cox: Well, in those days, it's probably changed now, you offered a paper for reading after working on it for some years. It was not something you did lightly, and you expected an onslaught. My impression is that it was much better received than I expected it to be. Well, there are one or two quite inaccurate comments on it, which is inevitable because people had a very short time to read the paper. I think it was reasonably favourably received.

Reid: Your 1970 book on binary data came more or less from that binary data paper, I guess?

Cox: Yes, it should have been written in 1960; that was a mistake. I don't know quite why I didn't write it then.

Reid: Why do you say it was a mistake?

Cox: Well, it's more like a book. Computing was changing extremely rapidly at that time. You see in 1960 computers were there and all sorts of things could be done but only with enormous struggle. By 1970 things were much more standardized. I mean nothing like they are now, but things were much easier. And things like the emphasis on weighted least squares, which was perfectly appropriate in 1960, was a bit old hat by 1970. In any case I had all the material; in fact I had lectured on most of it here and there. It was stupid not to write it down.

Reid: You might have been too busy writing all your other books.

Cox: [Laughs.] Could be.

Reid: I don't remember seeing in the paper analysis of binomial data, but there's a lot of that in the book.

Cox: No, that's in the book. That's almost in the 1955 paper [3], although for the Poisson case rather than the binomial case.

Reid: The binary data paper in many ways is quite similar to your 1972 [15] paper in that it highlights features appearing in a number of applications and then presents a relatively simple systematic way of approaching it. And yet the 1972 paper shot to meteoric fame, and I'm not sure the binary paper did.

Cox: No, and I think it's related to what I was saying about computing. Nobody took much notice of the 1972 paper for a while, until various people started to write software which was widely useful.

Reid: It was pretty fast, I think.

Cox: Yes, it was 6 or 7 years.

Reid: At Oxford they did the software before you gave the talk, I think. Peto mentions it in the discussion.

Cox: Ah but there's a difference: I had the software to do the examples but there is a difference between just writing the program to do an example and having something polished enough to distribute around the world.

PROPORTIONAL HAZARDS

Reid: Could you describe the background for your 1972 paper [15]?

Cox: Quite a few people—I think particularly of Peter Armitage and Ed Gehan and Marvin Zelen, and I think there were others—said they were getting a certain kind of data, censored survival data, with a lot of explanatory variables. Nobody knew quite how to handle this sort of data in a reasonably general way, and there seemed to be dissatisfaction with assuming an underlying exponential distribution or Weibull distribution modified by some factor. It seemed that something slightly more general was called for. Well, in the light of all sorts of things I'd done in stochastic processes it's entirely natural to approach this in terms of hazard. So the specification of some basic function of the underlying time scale, multiplied by a factor, that's sort of immediate and obvious really. I don't know whether it's new to that paper, I think probably it is, but anyway it's sort of immediate.

Then the question was how to actually do the statistical analysis. I wrote down the full likelihood function and was horrified at it because it's got exponentials of integrals of products of all sorts of things, unknown functions and so forth. I was stuck there for quite a long time—I would think the best part of five years or maybe even longer. Then suddenly I thought that the obvious thing to do was to concentrate on the part of the likelihood that actually gave you the information about the regression coefficients that you were interested in. It was absolutely obvious how to do that, and so just write down the answer. It occurred to me while I had a high temperature and was in bed with flu. It suddenly struck me that you could do this, and then when I felt better I tried to recover the argument and couldn't. But I was so convinced that when I was ill I had done this, that I tried again and then I saw what it was that I'd done.

Then, of course, the paper has a lot of other things in it, which came later: particularly the numerical example, and the idea of time-dependent covariates, and multivariate generalizations and so on. But the key thing was to see that it was obvious that you just ignored parts of the likelihood. Now of course that did raise the question of whether when you'd thrown away certain factors of the likelihood, you could still apply standard maximum likelihood results. Somehow or other, I don't exactly remember how, I persuaded myself that it was quite clear that the results did apply. So the paper just says that and doesn't really give any careful justification.

Reid: I think you used the word martingale in the paper.

Cox: Did I? Well, I was well aware it was connected with martingales. I'd been influenced by David Silvey [Silvey (1961)] who I think fifteen years before that had pointed out that there is a very strong connection between maximum likelihood estimation with dependent random variables and martingales. Anyway, I didn't spell it out in any detail: that was done two or three years later in a *Biometrika* paper which sets out the idea a bit more systematically [17]. I'd called the thing in the 1972 paper *a* conditional likelihood and I was taken to task by various people who said that it wasn't *the* conditional likelihood—I thought that was rather odd. It really was *a* conditional; it was a form of conditional likelihood.

So that's the essence of it. It didn't come from one particular application, but it came from perceiving, on the advice of others, that in medical statistics people were getting a certain kind of data that they didn't know how to analyze. And I think, though it's a long time ago and I don't remember too clearly, I could conceive that in industrial reliability and perhaps other fields essentially the same problems were arising.

Reid: Between the time that you suddenly realized how to do the likelihood and you finished the paper—about how long was that?

Cox: I don't remember; it wouldn't have been very long. Not more than a few months.

Reid: At that point did you have a feeling of excitement that one assumes goes with a great discovery?

Cox: Yes, I think so, because I'd had this problem at the back of my mind for a long time, and it was awfully nice to feel that I'd got somewhere with it.

Reid: I think that comes across in the paper; it almost seems to have been written in one go.

Cox: It would have been written extremely quickly, yes.

Reid: How do you feel about the cottage industry that's grown up around it?

Cox: Don't know, really. In the light of some of the further results one knows since, I think I would normally want to tackle problems parametrically, so I would take the underlying hazard to be a Weibull or something. I'm not keen on nonparametric formulations usually.

Reid: So if you had a set of censored survival data today, you might rather fit a parametric model, even though there was a feeling among the medical statisticians that that wasn't quite right.

Cox: That's right, but since then various people have shown that the answers are very insensitive to the parametric formulation of the underlying distribution [see, e.g., [20], Chapter 8.5]. And if you want to do things like predict the outcome for a particular patient, it's much more convenient to do that parametrically.

Reid: The paper has had an enormous impact, as you know, in many different directions. What do you think are the most positive benefits of the work?

Cox: Handling in-study covariates, that is, time-dependent covariates, I think is rather important—and the fact that it's readily adapted to multiple events, what the sociologists call event history analysis, for instance. It's the basis for really lots of further things in a fairly immediate way. Of course, another issue is the physical or substantive basis for the proportional hazards model. I think that's one of its weaknesses, that accelerated life models are in many ways more appealing because of their quite direct physical interpretation, particularly in an engineering context.

BOOKS

Reid: We did talk about your books to some extent when you mentioned earlier that none of your books are textbooks. I've always thought of *Theoretical Statistics* [16] as a textbook, and I've used parts of *Stochastic Processes* [12] as a textbook.

Cox: Yes, *Stochastic Processes* would be the nearest. But although it now looks very elementary, there was quite a bit of new material in it, both of Hilton's and mine.

It might help explain my attitude to these books to point out the traditional British method of university teaching, though I don't say this is good or that I altogether approve of it. You see, although I've taught quite a wide range of courses at one time or another, I've only once ever used a textbook. I would not normally, even for a moment, consider using a textbook in a course of lectures. I would refer the students to several different books, and see the role of books as backup for the teacher of a course, the opportunity to choose what he or she thought was important. Not to set out an exact prescription for somebody to follow. And that's why, perhaps, my co-workers and I have written books in the particular way we have. So, the serious point is the attitude to textbooks which is still I think entirely different in the U.K. than in the U.S. I don't say that in any spirit of thinking the U.K. system is better, just totally different. And I guess the best method is somewhere in between.

I've only once taught a course from a textbook, and that was in Berkeley in 1956, when one of the standard things one did in those days, if one was lucky enough to be invited to Berkeley for the summer, was to teach a summer school course from Mr. Neyman's elementary book on statistics, which I think is no longer in print. At coffee many mornings, Mr. Neyman would say anxiously "Have you done the example on so-and-so?" and he'd mention one of his favorite examples; and I cheated slightly, because I could almost always say, "Yes, I've done that example" but what I didn't tell Mr. Neyman was I often demanded more of the students than he had.

Reid: You're famous for your books being extremely concise.

Cox: Really?

Reid: [Laughs.] Yes. Is it something you need to strive for or is it something that comes automatically to you?

Cox: Well, I'm gradually coming around to the idea, and this is something I've only learnt perhaps in the last year or so, that it's not something to strive for so much as something to fight against. All my inclination, and all my training, is to write with not an unnecessary word.

And I suppose it's something about me personally; I find it much easier to understand something that is clearly put with a minimum number of words, when you know you've got to look at each word and think what it means. The notion that if you then double the number of words you make it any clearer, I think is not right. You know, if you double the amount of information or explain something at more length or give an example or something, that's helpful, but.... So I would in principle want to claim maximum conciseness is also maximum clarity.

There's also the psychological point. I find the notion of trying to explain a certain moderately advanced subject in a couple of hundred pages, the essence of it, much more appealing than the 800-page encyclopedia on something. In certain very particular subjects there's a need for an encyclopedic treatment but I don't think very many. I mean one could write a book a thousand pages long on the linear model, but would it be a good idea?

Reid: You said a minute ago that in the last year or two you've come to the notion that you needn't be quite so concise?

Cox: Hmmmmm.

Reid: Do you have a favorite book? of your own?

Cox: Well, either *Planning of Experiments* [7] or *Point Processes* [19], I suppose.

Reid: Any you really don't like or wish you hadn't written?

Cox: No, actually. Perhaps I should have, but I don't. Of course many of my books, and papers, are collaborative efforts. It's been my enormous good fortune to work with a succession of friends with whom collaboration has been both very enjoyable and from my point of view extremely fruitful. But, they're all a considerable pain to write. It's satisfying when it's done and on the whole, if you've thought about a subject for a considerable time and feel you have something to say about it that isn't in the literature already, it is entirely sensible to write it down as a book of some sort. Whether it is a textbook or not doesn't really matter. But with one exception I think all the books have taken a mighty long time to do, the exception being the little book on renewal theory [10], which took about three months. The slowness comes partly from not concentrating on one thing at a time, but more seriously it comes from a lack of clarity about what you want to say.

Reid: Five years you said once, on average it took you to do a book.

Cox: Did I?

Reid: Yes, which means you sometimes must have been doing two at the same time. Do you enjoy writing just for the sake of writing?

Cox: No, I certainly don't. I'd rather look at the problem which somebody brought in this morning, so to speak.

THEORY AND APPLICATIONS

Reid: Before this interview you suggested talking about the motivation of theoretical research.

Cox: I think what I had in mind was perhaps this: that people say theoretical work in statistics should be motivated by applications because it's a practical subject, and that of course is true. On the other hand, I think theoreticians have to try and stand back from individual applications. Some of the papers I've written have been very strongly tied to, for example, solving a particular problem in experimental design in a very particular context. Okay, that can be worth doing maybe, but the things that are more likely to be widely useful are those where you stand back from one very particular application and say here's a whole family of problems that arise in applications in several fields, and try to address that. That's a better way to go if you can. And in a sense, you see, the work on conditional inference is one step further back still from that, in that it was very strongly motivated by practical experience and yet on the other hand, I couldn't say it arose from one particular special type of problem. It arose in a sense from all the applied work I'd done to that point.

Reid: Are you saying that you saw something missing from the theory?

Cox: No, but that I felt, for instance, that various aspects of the Neyman–Pearson theory—choose alpha, choose a critical region, reject or accept the null hypothesis—give a rigid procedure, that this isn't the way to do science.

Reid: To me it's so obvious that it's not the way to do it, but presumably not at all obvious in the fifties.

Cox: I agree it's obvious but why then do people write books that say this is what you should do? But, Neyman talked a lot about inductive rules of behavior, and it seemed to me he took the view that the only thing that you could ever say is if you follow this procedure again and again, then 95% of the time something will happen; that you couldn't say anything about a particular instance. Now, I don't think that's how he actually used statistical methods when it came to applications; he took a much more flexible way.

But even apart from that, you can say, is this notion of 5% or 95% region—is this just an explanation of what a 95% confidence interval would mean? A sort of hypothetical explanation, if you were to do so and so, such and such would happen? Or is it an instruction on how to do science? It seems to me okay as the first, in fact very good as the first, terrible as the second. I don't think this has always been very clear. Do you think so?

Reid: No, I agree not. It always seemed clear to me, if I thought about it, which wasn't terribly often I suppose.

Cox: Yes, but if you do, if you're involved in doing applied work, you don't necessarily have to think about it but you have to have a broad approach. The theoretician's job is partly to try and capture how you should use these techniques as well as just to form a theoretical basis.

Reid: So it really is motivated by applications, or motivated by science.

Cox: Yes, totally so. But not by saying here is this problem in mineral technology or something to which here is the answer.

Reid: Yes. We seem to hear a lot about this nowadays. Unless you're going to the lab and standing shoulder to shoulder, . . .

Cox: Well, I'm a terrible experimenter, but I have spent a fair amount of time in labs, although not as much as I would have liked to. But of course some of this discussion has a strong antitheoretical tone to it which seems to me destructive and totally unnecessary: "We're practical chaps and we don't need all this theory. All we need to do is plot a few graphs, and be sensible." Now it's important to plot a few graphs and be sensible—it's important and difficult, but if you remove theory then the whole subject becomes nothing. It becomes a collection of fragmentary tricks.

Reid: The theory that you're describing where you see a common thread in a variety of applications, it's really rare to see that kind of theoretical work done. And in that sense one can sympathize with the more practical people in that a lot of the theoretical work does seem very self-motivated.

Cox: There's nothing wrong with that but different people get their motivation in different ways. I don't think you can lay down any law. And of course, as the subject gets more and more specialized, it's getting more and more difficult for particular individuals to know enough about more than at most one or two fields of application. Forty years ago it was perhaps a bit easier to be wide ranging in applied interests.

Reid: Let me ask just one more question on the theory topic. I'm not sure if it has a sensible answer, but what's your favorite part of theoretical statistics?

Cox: I've never thought of that. Well I'm tempted to say, what I'm working on at the moment, but that's a slightly facetious answer. You're thinking now as a teacher?

Reid: I suppose I mean possibly as a teacher or expositor.

Cox: It's an interesting question. I very much admire subjects like certain areas of pure mathematics, or for that matter certain areas of probability theory, where a seminar can be given in the old style with a very large blackboard, in which the lecturer begins in the upper left-hand corner of the blackboard and develops a theme and ends 50 minutes later down in the bottom right-hand corner of the blackboard. Then you've got an area before you—a beautiful painting, almost. It would be excellent to be able to do that with a topic in statistics. I feel it's very hard, partly because the difficult aspects are often more conceptual than mathematical. But if one could think of it, a piece of theory, a new theory that could be laid out like that, I'd find it extremely appealing.

Reid: Turning from theory to applications, I wondered if you had any specific applications or consulting problems that you especially enjoyed or you thought were particularly well done and useful.

Cox: Well, I collaborated when I was in Leeds very closely with a textile physicist. That was very interesting indeed, and I suppose if I had stayed in Leeds we might have developed that a bit further. It was some mixture of physics, statistics and classical applied mathematics. More recently, I tend to get involved in applied problems almost at second hand via other statisticians or epidemiologists or whatever.

The only clinical trial I've ever been deeply involved with at first hand was a big primary prevention study on hypertension. I was the only nonmedical person on the management committee, and that was extremely interesting. Not perhaps so much from the statistical point of view, although it was reasonably interesting statistically, but to get some insight into how these things go. Otherwise, I've tried to really range as broadly as possible in applications, and I can't immediately think of any particular one that stands out.

Before I came to Oxford I'd had dealings with a lot of people in different areas of work, in the physical and biological sciences, in technology and to a limited extent in some other fields. I hadn't had any very systematic contact with social scientists. One of the appeals of coming to Nuffield College, which is a postgraduate college in the social sciences, was to broaden my applied experience and to get a better idea of what social scientists are up to. I have found that extremely interesting in various ways. It's left me with a great respect for sociologists, in particular, who in the empirical tradition here are very careful about their data collection and their analysis, and cautious in interpretation. And, of course, they are working totally with observational material, whereas I've tended on the whole to work with experimental material. That brings to the forefront issues like to what extent can you draw reliable conclusions from observational data and also this issue of trying to get somewhere vaguely approaching a notion of causality from observational material. It's very interesting and important.

Reid: I'm not sure students get the same exposure to applied problems that you did.

Cox: Well I was exceptionally fortunate in most respects. I didn't learn perhaps a great deal at the Royal Aircraft Establishment, although I expect I learned more than I think I did. But WIRA was fantastic experience. And then Cambridge also, there were a lot of people around who were very good scientists who one could talk to, who wanted to have discussion of their problems. There's a great danger I think that statistical consulting, so-called, in universities can end up rescuing not very good or bad doctoral theses in other subjects, rather than talking to the leading scientists in those subjects. It's the second we should be doing, not the first. The first is a kind of moral duty, up to a point, but not at the expense of the second. Too much of the first leads to all sorts of undesirable things; in particular, a certain arrogance amongst some statisticians towards engineers and scientists, which seems to me absolutely ludicrous and very dangerous.

Reid: It's very difficult though, isn't it, that kind of collaboration that you're describing. You need to find the people or they need to find you.

Cox: Yes, it's very very difficult—I'm tempted to say a matter of luck.

Reid: Do you sit down and fool with data much.

Cox: I sit down, and [laughs] make suggestions to other people as to what do to and sometimes I hover annoyingly over them while they do it.

Reid: I wondered what you'd like to say about computing?

Cox: Well, in the days of the electric calculator and the old hand Brunsvega, I did an enormous amount of numerical work. Then when computers came along I went to a programming course in London. In those days the notion was you had to learn to program in machine code, before you were let loose on Fortran or anything like that. And I did write a program for linear regression in 1957. I think I realized at that point that I had to make a choice either to spend a large proportion of the next 20 years on this, and it was obviously going to be extremely important subject, or to spend nothing and just try and follow what was going on so I had an idea what was feasible, what wasn't feasible.

I decided on the second, and I think for me that was the right decision to make. The whole business is a great miracle and I think people take it for granted. There was a period, perhaps in the late fifties, when if you went to a statistical meeting all you'd hear would be people talking about whether they'd got their regression program to run. I can remember going to one of the leading British computer scientists and saying I was interested in inverting matrices. "Ah, yes," he said. "Can you do a 3-by-3 matrix?" I said. "No, not yet. Tomorrow?"

I did, with Katherine Booth, construct some designs on the computer in a paper in *Technometrics* in the late 1950's [11]. I think that might have been one of the first uses of a computer to construct experimental designs. So I'm sort of interested, but also being an extremely impatient person, I still find the fiddley details irritating. People have been saying for 20 years, "Next year everything will be absolutely painless and foolproof, very, very easy to use." Well now what's happened it seems to me is that, although what can be done has increased by a fantastic factor in that period, and it has got easier, it's not got easier enough for the casual user.

I was also very fortunate to spend six months at Bell Labs in 1965, where the main topic of discussion was in what direction statistical computing was going to go. Martin Wilk, I think, was the driving force behind it. I think the notions of S and S-PLUS were in a way foreseen by Martin. As I recall it he was saying, "We don't want a package for regression and a package for analysis of variance, and a package for binary data, a package for time series, and a package for multivariate analysis; we want something general enough to see the common elements in these things and be able to move from one type of calculation to another." But even with the enormous resources at Bell Labs behind it took a long time for that to get to fruition.

Reid: Well, it really sounds like a prescription for object-oriented programming.

Cox: Yes, almost.

ENCOURAGEMENT

Reid: The only other thing I wanted to ask you about is something that you mentioned to me in a letter a while ago which I think was connected with your knighthood. Your words were roughly that, after feeling when you were younger that you didn't get very much recognition for your work, you now felt you were receiving a "bizarre excess."

Cox: Yes, I think that sums it up adequately. Well, everybody needs encouragement, and of course as you get older you still need encouragement. But the time you most need it is when you're starting. It would be quite wrong to think that people were ever discouraging, they weren't. It was all very low key, typically British understatement sort of thing. You never really knew quite what people thought, despite the relative frankness of RSS discussions. And I'd published a few papers with very little notion of whether anybody had paid any attention to them. Until I first went to the United States, where people would come and say, "Oh, I read that paper, I think we could do so and so." That sort of thing is very encouraging. And then, more recently, I've been absurdly lucky with all these pieces of recognition. Of course the system's a bad one in one sense, that if you get one piece of recognition it's more likely you'll get another. It ought to be the other way around.

Reid: When you were younger and you were not getting initially positive feedback, was that just typical British understatement?

Cox: I think so. Well, I don't even remember thinking about it at the time; this is a retrospective feeling. I wouldn't have expected it, but I perceive now that it would have been good.

Reid: You've been very encouraging to young people; I think that's the first thing anyone I speak to about you mentions.

Cox: Well, I hope I have been; I've certainly tried. I think because of this perception in my own case that it was important.

Reid: Was there a time when you realized that your fortunes had changed?

Cox: Well, two occasions I suppose. First of all I bumped into George Barnard in the street and he said something like, "Have Birkbeck put you in for a professorship yet?" He must have known that they were intending to, which I didn't know, and I was absolutely astounded. Years before when I was at Cambridge I would have said my chances of ever getting a full professorship in a British University were virtually nil. So that was one thing. And then Fellowship of the Royal Society—I was astounded to be nominated, and some people go berserk with anxiety over this; it's probably very similar in Canada and the United States, and in other countries. You're nominated and then some years go by, and if you're lucky you suddenly get a message. I just dismissed it from my mind, I thought it would never go through.

Reid: But those were, as you describe them, surprising, unexpected encouragements. Was there a time when you suddenly felt you were a Very Important Person?

Cox: Well, I hope I've never thought so. [Long pause.] In a sense, the only thing that matters is if you can look back when you reach a vast, vast, vast age and say, "Have I done something reasonably in accord with my capability?" If you can say yes, okay. My feeling is in one sense, I've done that: in the tangible sense of books and papers, I've done more than I would have expected. In another sense I feel very dissatisfied: there are all sorts of problems that I nearly solved, and gave up, or errors of judgment in doing a little something and not taking it far enough. That I nearly did something you see, this is the irritating thing. You know, if you'd no idea at all, well it doesn't matter, it's irrelevant, but if you feel you were within an inch of doing something and didn't quite do it...

Reid: David, your energy and modesty continue to be an inspiration to a great many people, including myself. Thank you very much for the privilege of this interview.

CITED PUBLICATIONS OF D. R. COX

[1] (1949). The theory of drafting wool slivers, I. *Proc. Roy. Soc. London Ser. A* **197** 28–51.

[2] (1949). *An Outline of Statistical Methods for Use in the Textile Industry.* Wool Industries Research Association, Leeds. (1st ed. 1949; 5th ed. 1960; With A. Brearley.)

[3] (1955). Some statistical methods connected with series of events (with discussion). *J. Roy. Statist. Soc. Ser. B* **17** 129–164.

[4] (1955). A use of complex probabilities in the theory of stochastic processes. *Proc. Cambridge Philos. Soc.* **51** 313–319.

[5] (1955). The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proc. Cambridge Philos. Soc.* **51** 433–441.

[6] (1958). Some problems connected with statistical inference. *Ann. Math. Statist.* **29** 357–372.

[7] (1958). *Planning of Experiments.* Wiley, New York.

[8] (1958). The regression analysis of binary sequences (with discussion). *J. Roy. Statist. Soc. Ser. B* **20** 215–242.

[9] (1961). *Queues.* Methuen, London. (With W. L. Smith.)

[10] (1962). *Renewal Theory.* Methuen, London.

[11] (1962). Some systematic supersaturated designs. *Technometrics* **4** 489–493. (With K. H. V. Booth.)

[12] (1965). *The Theory of Stochastic Processes.* Methuen, London. (With H. D. Miller.)

[13] (1966). *Statistical Analysis of Series of Events.* Methuen, London. (With P. A. W. Lewis.)

[14] (1970). *The Analysis of Binary Data.* Methuen, London. (2nd ed. 1989, with E. J. Snell.)

[15] (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

[16] (1974). *Theoretical Statistics.* Chapman and Hall, London. (With D. V. Hinkley.)

[17] (1975). Partial likelihood. *Biometrika* **62** 269–276.

[18] (1977). A bivariate point process connected with electronic counters. *Proc. Roy. Soc. London Ser. A* **356** 149–160 (With V. Isham.)

[19] (1979). *Point Processes.* Chapman and Hall, London. (With V. Isham.)
[20] (1980). *Analysis of Survival Data.* Chapman and Hall, London. (With D. Oakes.)
[21] (1987). Some models for rainfall based on stochastic point processes. *Proc. Roy. Soc. London Ser. A* **410** 269–288. (With I. Rodriguez-Iturbe and V. Isham.)
[22] (1988). A simple spatial temporal model of rainfall. *Proc. Roy. Soc. London Ser. A* **415** 317–328. (With V. Isham.)
[23] 1988. A point process model for rainfall: further developments. *Proc. Roy. Soc. London Ser. A* **417** 283–298. (With I. Rodriguez-Iturbe and V. Isham.)

REFERENCES

BARTLETT, M. (1955). *In Introduction to Stochastic Processes.* Cambridge Univ. Press.

CHANDRASEKHAR, S. (1943). Stochastic problems in physics and astronomy. *Rev. Modern Phys.* **15** 1–89. [Reprinted in *Selected Papers on Noise and Stochastic Processes* (N. Wax, ed.). Dover, New York, 1954.]

HINKLEY, D. V., REID, N. and SNELL, E. J., eds. (1991). *Statistical Theory and Modelling: In Honour of Sir David Cox, FRS.* Chapman and Hall, London.

RICE, S. O. (1944). Mathematical analysis of random noise, I. *Bell System Technical Journal* **23** 282–332. [Reprinted in *Selected Papers on Noise and Stochastic Processes* (N. Wax, ed.). Dover, New York, 1954.]

RICE, S. O. (1945). Mathematical analysis of random noise, II. *Bell System Technical Journal* **24** 46–156. [Reprinted in *Selected Papers on Noise and Stochastic Processes* (N. Wax, ed.). Dover, New York, 1954.]

SCHWARTZ, L. (1950). *Théorie des Distributions* **1**. Hermann, Paris.

SCHWARTZ, L. (1951). *Théorie des Distributions* **2**. Hermann, Paris.

SILVEY, S. D. (1961). A note on maximum-likelihood in the case of dependent random variables. *J. Roy. Statist. Soc. Ser. B* **23** 444–452.

WHITTLE, P. (1993). A conversation with Henry Daniels. *Statist. Sci.* **8** 342–353.