Statistical Science
2003, Vol. 18, No. 3, 346–356

A Conversation with John W. Tukey

Luisa Turrin Fernholz and Stephan Morgenthaler

This conversation with John W. Tukey took place on June 20, 1995, at Princeton University's Jadwin Hall. The questions were asked by Luisa T. Fernholz, Stephan Morgenthaler and others among the public present. The conversation was taped and what follows is a typescripted and slightly edited version of these tapes. The conversation was previously published in *The Practice of Data Analysis* (1997).

JOHN W. TUKEY'S HIGH SCHOOL AND COLLEGE DAYS

Q: I am going to start with a somewhat personal question. We heard yesterday that you did not have a formal education, but were educated at home. Could you tell us a little bit about that?

A: Okay, well, by the time I was five, my parents had settled in New Bedford. My father was head of the Latin Department in the high school. In those unregenerate days a married woman couldn't be a teacher in Massachusetts. So, my mother wasn't a teacher, but she was a substitute. And I have heard it claimed, that between the two of them, they ended up teaching everything in this high school, except bookkeeping and physical education. I think you have to add chemistry to that. And rumor says that my mother decided that it would be bad for me to go to school because, either I would get very lazy, or I'd be a problem, or something. And, so, there wasn't too much formal education. But I spent a lot of time in the public library. New Bedford had a wonderful public library in those days. Not only did it have the *Journal of the American Chemical Society*, but it had the *Transactions of the American Mathematical Society*. And I think the reason that I went to Brown as a chemist was because I could read the *JACS*, but I couldn't read the *Transactions*.

Luisa Turrin Fernholz is Professor of Statistics, Department of Statistics, Temple University, Philadelphia, Pennsylvania 19122-6083 (e-mail: fernholz@temple.edu). Stephan Morgenthaler is Professor, Institute of Mathematics, Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland (e-mail: stephan. morgenthaler@epfl.ch).

Q: When you went to college did you regret being brought up in this isolated environment?

A: It wasn't that isolated, in the sense that I am not sure that the environment was more isolated than if I had gone to the high school. I actually went to the high school for one term in French and some mechanical drawing. I am not sure if I am the last person to enter Brown with credit for mechanical drawing or not.

Q: How did you enter Brown?

A: College board exams. And so I went in and ended up with advanced credits in mathematics and, I guess, German. So, I went to junior differential equations as my freshman math course. We had a cousin who was the head of the mathematics department in the high school. But, again, there wasn't too much formality. But I worked lots of problems in a calculus book, and that seemed to produce the necessary effects.

Q: You did your Ph.D. at Princeton in mathematics. Tell us about that.

A: I came to Princeton in 1937 as a graduate student in chemistry, and ended up being a lab assistant in one of the freshman inorganic courses. In Princeton you had to be a Ph.D. to be a lab assistant in physical chemistry, which worried me a little because I had been a lab assistant in physical chemistry for a year and a half at Brown. But, anyway, I fell over the fence the summer before I came to Princeton. I came in chemistry, but I spent a lot more time in old Fine Hall than in Frick; and I took prelims at the end of that year.

Q: Harking back to Brown once more, do you remember a particular professor or course you liked?

A: I don't think there was one that was an obvious dominant influence or anything of that sort. I knew most of the professors in the mathematics department and most of the professors in the chemistry department. I was there four years and, at one point, I was going to take a master's degree and an Sc.B. at the end of the four years. The department didn't like giving two degrees at the same time and so they decided to give me an Sc.B. after three years. But W. A. Noyes, Jr., who was later the editor at *JACS*, used to claim that he did the glass blowing for my thesis experiments in my fourth year. I was well enough tuned in on the scuttlebutt in the math department. There was a lady graduate student who went to Illinois and the word came back that she was going to marry Pierce Ketchum who was a professor there. And so, I know who it was who proposed (in the math department) to send a telegram reading “Congratulations, you Ketchum.” So, I didn’t feel isolated! But, probably among the mathematicians, it was Tamarkin from whom I took a graduate course in the second, or third year; he was sort of the senior research man in math at that point. Among the chemists—the course work—I don’t think I have anything to say. I spent time with the physicists (and geologists), too. Bruce Lindsay, who was then chairman of the physics department, had come from New Bedford originally. I don’t think I have a better answer for you.

Q: When you went to Princeton as a graduate student, did you take courses, or did you immediately start on your research?

A: No, no! I went to lots of courses and seminars in Fine and to the chemistry courses that a graduate student (first year) might reasonably be supposed to go to. Now, if I hadn’t already fallen over the fence, I don’t know. Henry Eyring was here in those days. Henry was a theoretical chemist who was the salt of the salt of the earth. And if I hadn’t been so far over the fence—and if Henry hadn’t been away for a semester—I might have stayed in chemistry a little longer. But in those days the Institute for Advanced Study and the math department were all mixed together in old Fine and people sort of didn’t segregate, either in lectures or in turning up for tea. Along in the spring of that first year when I was still a chemist, Marston Morse asked me whether I was at the University or at the Institute. Since I was still a chemist, I thought that was interesting. But what had been going on is that Norman Steenrod and I had been sitting in the two back corner seats in old Fine 113 and doing our best to keep Marston honest in the seminar he was giving.

Q: In general topology I learned about Tukey’s lemma and I know that you have made other important contributions to mathematics. I was wondering, how your transition from mathematics into statistics happened?

A: Practice! May 1941 I went to work for Merrill Flood in Fire Control Research—boom–boom fire control, not hose fire control. Charlie Winsor was there, and except for the year and a half that Charlie wasted in Washington in the Mine Warfare Research group, Charlie and I spent the war pretty much either in the same office or across the hall. Charlie knew an awful lot of statistics that wasn’t in the books then, and I am sure a certain amount that isn’t in the books yet. So, I learned by talking to Charlie and by doing things and by reading.

Q: Going back to your days as a graduate student in mathematics, who were the professors among the mathematicians at Princeton whom you remember most?

A: Again, I am not sure that one can pick individuals. I quote you a verse from the faculty song. “Here’s to Lefschetz, Solomon, L. unpredictable as hell, when, laid at last beneath the sod, then he’ll begin to heckle God.” That was the Fine Hall verse that senior graduate students, probably years before, produced. That was a private verse for the faculty song. So, Lefschetz is one, Bohnenblust is another because he was a clear lecturer, Steenrod was on the faculty at that point, but we interacted lively, Tucker. Although there were a lot more contacts in Tucker’s case during the war than before. I don’t know just what to say.

Q: What about life in the graduate college? Do you remember any particular incidents?

A: The situation in the graduate college was that fairly soon I joined a group of people who ate together. It was a continuation of a group that had eaten together the year before. This was before Hitler’s invasions so it is perhaps not surprising that it was known as the Fuehrocracy, and Lyman Spitzer, who just recently retired from across the street here, in astrophysics, was officially the Fuehrer. He sat at the end of the table and if spare ice creams needed to be divided it was his responsibility to divide them fairly. But there was a physicist or two, a couple of astrophysicists, one theoretical chemist, several mathematicians and one romance-linguist, who was a courtesy member of the group, but who had been authorized to put anybody into a Klein bottle that he wanted to. But, Ralph Boas, who was a national research postdoc, and Frank Smithies, who was over for a year from Cambridge, and I ended up hanging out together a fair amount. So, we were just three people who sat through Aurel Wintner’s lectures on convolutions and so when the Princeton notes came out we were the note writers, the only other person was Cyrus McDuffie—and so, the Library of Congress card to this reads, notes by Ralph Boas, Frank Smithies, John Tukey with sympathetic encouragement from Cyrus C. McDuffie. And when the seminar came to the end at the end of spring, McDuffie packed us all into his car and the whole group went up to North Jersey for the day. Again, there was no shortage of interaction.

I am not sure which year it was, it must have been the next year. Arthur Stone bought some paper for his ring notebook at Woolworth's and since he had a British notebook, he had to cut an inch off it. Since he had all these strips of paper, he had to do something interesting with them. So, he was folding the regular polygons and he was smart enough to recognize the first known hexaflexagon when it was made. So that Arthur and Brian Tuckerman and Dick Feynman and I spent a fair amount of time on flexagon theory. The asymptotic formula is known for the number of different hexaflexagons with n sides, for example. It turns out to be equivalent to the problem of triangulating a regular polygon, etc.

Q: Who was your advisor?

A: I think Lefschetz. You know, Peisakoff played the version in the opposite direction. I think I was supposed to be his advisor, but I was not at all convinced.

Q: I was going to ask you whether you advised any math Ph.D. students after starting to work as an instructor.

A: Not that I can specifically think of. But I was really only two years there. In May 1941 I pulled out across Nassau Street.

THE WAR YEARS

Q: Can you tell us a little bit about the work in the fire control group during the war?

A: It started out as a project to study the training of height and range finder operators. Do you know what a stereoscopic range finder is? It is a situation where you see the field differently in the two eyes and there's a reference mark and you can try to make the reference mark appear at the same distance as the target. Details are not that important for the present purposes. And it was mainly stereoscopic height finders, i.e., they had automatic conversion of range to height for antiaircraft fire and then all the big naval guns used stereoscopic range finders, because most of the fire against naval targets was against targets that you could see at least the mast. Particularly if you put the range finder high in the other ship's mast. And Brock McMillan, who was I guess a year or so ahead of me as a postdoc, ended up with other people at Fortress Monroe running a field laboratory and the group in Princeton was a combination of true target position and analysis. The point is, if you're going to test height finders and height finder observers, you have to have some way to know how high the target really is. And so there were recording photo-theodolites which somebody tried to keep pointing at the target. Where the image showed in the frame you could correct and get a good idea of what the angle of the target was. Then we had one of the first IBM multiplying punches and so we actually got IBM calculations of what the true heights actually were. But there was a lot of physiology and fairly soon there was more to do with the hardware. Why there were temperature errors? We pioneered filling the height finder with helium instead of air, the point being the thermal conductivity of helium is about seven times that of air. So the temperature gradients inside the instrument were much smaller if you filled it with helium. And this broadened to get into armored vehicle fire control. We had various interactions with Frankford Arsenal and eventually we were, for a while, funded out of there. We had some civil service personnel here at 20 Nassau and they couldn't be told what to do by anybody but a government employee. So, a couple of us became part-time technical experts so we could supervise the civil service people. Then Colonel Trichel moved to Washington and we ended up getting involved in testing rocket powder (because he took over the corresponding section of the Office, Chief of Ordnance). And then later on we came back to NDRC because there was this project (AC-92) which was trying to do all the fixes on the B-29 as an operational device that they could. And we ended up being the coordinating group and, as I say, that's when I learned to ride airplanes. The Mt. Wilson Observatory people were hanging up little models of aircrafts with light shining out of them to see what the defensive fire coverage really was for different formations, and there were people in two or three places in Texas, and also something was going on at Smoky Hill Army Air Field in Kansas. Those who've never seen a loaded B-29 take off in Kansas probably missed a sight that will never recur again. This was the really flat Kansas and loaded B-29's would go down the runway—I think they had somewhere between a 10,000-foot and 12,000-foot runway—and then they disappeared under the curve of the earth and you wouldn't see them till they were maybe 10 or 20 miles out. With that engine, by the time you got up to somewhere between 100 feet and 200 feet, the temperature on the engine was over red line and you had to flatten out (to ease the load on the engines and let them cool a little) before you could fly up the rest of the way. This is why the airstrips on Tinian and Saipan, and so on, always went to the water's edge to get maximum clearance, except for wave heights.

After that job wound up I went to work for Murray Hill (Bell Labs) and was involved in the first paper-and-pencil study for what was called AAGM1 and later called Nike and later called Nike–Ajax after it started having big brothers. This was an antiaircraft guided missile. Bernie Holbrook (who was, we finally decided later, something like a ninth-and-a-half cousin of mine), a switching engineer, and I, we did trajectory, aerodynamics and war head for the paper-and-pencil study, both of us being "experts" in all these fields. But on the other hand the state of supersonic aerodynamics was poor even if we got all of the best information. We went to Langley and talked to the people around the wind tunnels and so on. Then, when Don Ling came to work for the Labs somewhat later, he produced a little pink paper—meaning unofficial draft memo—ascribing to me the theorem that if a semicontinuous function had its values known at three points it was well determined, but if it was a continuous function you only had to know its value at one point. That was about the state of knowledge of supersonic aerodynamics. People were still taking seriously the linearized theory that in particular said that the control surfaces would start to work oppositely when you went through Mach equals root three. Of course, nothing of that sort ever happened. So, I stayed with this. Afterwards I spent a fair amount of my Murray Hill time in connection with Nike for quite a long time. I got to go on impact parties out at the White Sands Proving Grounds, which meant going around through the boondocks and seeing if you can find the pieces that came down.

STATISTICS IN THE 1940s AND 1950s

Q: Coming back to your statistics education. You were turned into a statistician through practice.

A: And eating an average of 1.9 meals a day with Charlie Winsor over probably the equivalent of three years.

Q: So, he was a major influence on you?

A: Yes.

Q: Did you ever read a statistics book?

A: Oh, yes. When I was at Brown I read a lot of miscellaneous books in the math library, including some statistics books. Even back at that stage. I used to have a little tin container with 3 by 5 cards in it that had interesting looking tables of critical values for statistics. That didn't mean I had any feel for them. But on the other hand I am one of the—I suspect not inconsiderable number—who taught a graduate course in statistics before he ever sat in a course of statistics.

Q: You took no course from Sam Wilks?

A: No course of this sort, no. But I saw the books and looked through them. Sam and I did a little joint work.

Q: Around 1945 through 1950 could you describe the statistical community in the U.S.?

A: Well, I'll just have to try to isolate that period by guess. ASA had been meeting yearly since lord only knows when, maybe since the foundation. The Institute I think came into existence in 1938. In those days to join the Institute you had to be proposed by two members. They were worried about keeping the nonmathematical statisticians out for a while. That they gradually recovered from. The Biometrics section of ASA had been in existence and ENAR came into existence probably in late 1945. This is the Eastern North American Region of the Biometric Society. I think it would have been late 1945 and not 1946, one day there was a meeting in Woods Hole to set up the Biometric Society followed by the first meeting of ISI after the war in Washington. And Linder and Fisher and I and someone else shared a sort of a four-room suite at the hotel at this (Washington) meeting. But it was somewhere about this time there was some discussion about the vigor of comments in biometrics meetings. We had one of the very good biologists give a talk. And people had asked questions as vigorously as usual. So, there was some question at lunch whether the speakers had been unfairly treated. The outcome of this was that it was agreed that at a biometric meeting one was entitled to ask any question that one felt like. So, the biometric thing was an area of activity. I think more so than mathematical statistics. But mathematical statistics more so than ASA in general. Who were the key figures? After Sam Wilks, Hotelling and Wald, really I think you have to count Gertrude Cox and George Snedecor. Although Cox and Snedecor were not research contributing types, they still played a large role. From an older generation A. T. Craig in Iowa City and on the coast of course Neyman and the people that he'd drawn together. And fairly soon Bowker at Stanford because of his building-up powers. This isn't an exclusive list. Now, if you took a biometric flavor, you would get a different set of people. You probably have to get Cochran in both sets. People who'd be respected on the biometric side certainly would include our friend at the Mayo Clinic, Joe Berkson, who was a red-headed Irishman. Red was pretty much faded in the hair but not in the spirit. Since he had both an M.D. and a Ph.D. you couldn't put him down in any obvious manner if you got into an argument. Now, maybe it's relevant that I think I made some comments at about that time that the person that I would be most careful with if they were in the audience and I was giving a paper was Milton Friedman. Because Milton had worked with Allan Wallis at a different Columbia research group during the war. This is the one that produced the sequential analysis things. One day they took Jack Wolfowitz out to lunch and worked him over thoroughly about the importance of doing something with this. A month later when nothing came of that, they took Abraham Wald out to lunch and he came in the next morning with the fundamental identity of sequential analysis. Milton was well acquainted with the statistical side and very sharp. Probably easier to cut yourself on him than anybody else in the Biometric Society or the Institute. Of course Jimmy Savage was well started on his way up by that time. As of 1946 if either Jimmy or Fred Mosteller was in a room out of sight talking, no one could tell which one it was. They'd spent so much time working together during the war that they ended up equivalent in accent.

JOHN W. TUKEY'S WORK FOR THE FEDERAL GOVERNMENT

Q: I would like to go on to your work with the government. You were a member of the President's Science Advisory Committee. Do you recall incidents from those meetings?

A: Jaa. I'll give you an anecdote. But let me remark, I don't think that over those years I did any statistics (for PSAC). But one of the earlier environmental reports was being discussed and there were people in from some of the government agencies. And it became clear to them that PSAC was going to have, if anything, kind words to say for Rachel Carson in *Silent Spring*. And the people from agriculture practically wept in their beer. They didn't think that she should receive any mention or notice whatsoever.

Q: The FFT got started at a meeting of PSAC?

A: That's not quite the story. The FFT's realization was partly influenced by PSAC meetings. I used to sit next to Dick Garwin down at the far end of the table. This was in the room in old State which had once been the secretary's office. This was the room in which the Secretary of State saw the Japanese envoys just before Pearl Harbor. I was sitting there scratching and Dick wanted to know what and I told him what it was about generally. He went back to Watson Lab and he had some things going on that required some large Fourier transforms and so he tried to get Jim Cooley to program this. And after a while Jim did regard it solely as a programming exercise—all the theory was done. So I wouldn't have thought all the theory was done, and this is why there is the Cooley–Tukey paper. It really didn't start there and the initial reference is probably the *Princeton Notes* on a graduate course in time series, which I think either matched or predated this. And what really happened in due course is that Jim Cooley produced one algorithm and Gordon Sande produced the transposed algorithm. And eventually IBM kicked Jim into publication because they decided they didn't want to try to patent it and they didn't want anybody else to. And somebody fell out to be a co-author and I sort of floated along and didn't take adequate action to see that Gordon got his stuff out, which I always felt bad about.

Q: Your work on the PAQAB seems at least to have in part led to the award of the Presidential Medal of Science. Could you elaborate on the work you did there?

A: I doubt if there was very much connection. That sounds to me like something that got mentioned in a list because it was handy. PAQAB (President's Air Quality Advisory Board) was moderately effective. This was in the first Ruckelshaus era as EPA administrator and Ruckelshaus was still optimistic, feeling that if you talked nicely and informatively to the polluters you could get them to stop polluting. He learned quite rapidly. It was a respectable advisory group but nothing in particular.

Presumably, that got cited along with "Restoring the Quality of Our Environment," a report that was originally a Great Society task force for (President) Johnson. There were ten such, nine of them reported through one channel in the White House, we reported through another one. Nine of them leaked; one of them didn't. And after things were over, it was decided that this ought to be converted into a report for the general public and this was done under PSAC auspices. So that this is where the "Restoring of the Quality of Our Environment" came from. That was the first all-types-of-pollution moderately comprehensive report. We beat the Academy group under Athelston Spilhaus by about 6 months. So, PAQAB sounds interesting but is not worth particular attention, by comparison.

STATISTICAL EDUCATION: THEORY, CONSULTING, EDA, ETC.

Q: What are your views on education? How should we train the future generations of statisticians?

A: Now, let's just get the ground rules clear! Are we talking about education of statisticians, education in statistics or education in general.

Q: Let's start with education of graduate students in statistics. How much math, how little math? Does it help, does it hurt?

A: I think the answer to that one is: First, there shouldn't be a single answer. Secondly, it's not the whole story. Mathematics didn't hurt me; as I understand it, when one graduate student came to Princeton, the word among the grad students in the graduate college was—here is someone who is never going to get a Ph.D. from Princeton in math. He met the formal requirements all right. But, if he had been pushed too hard in math, something would have had to break. And statistics couldn't have done without having him as a statistician. So the answer is there are places where it can hurt badly. You can't live these days with no math. But neither, I suspect, can the average graduate student live with all the math required to adequately cope with what's in the journals. They talk about practicing defensive medicine. Statisticians may have to practice defensive mathematics. I well remember a remark Charlie Winsor made walking down just in front of old Fine. Charlie said: "Sam Wilks trains good mathematical statisticians and it's surprising how soon they become good statisticians." Now, what I worry about most about the math is well, (firstly) the loss of some people who can think well and who can do good things and (secondly) also the deflection of people away from thinking about what they can do much more effectively. I think everybody appreciated Paul Velleman's speech yesterday. Paul came to Princeton from essentially a mathematical sociology program and he managed to survive the necessary mathematics. But I think it was survival rather than anything else. So, if, given a choice between turning out proto-statistician B with no mathematics and proto-statistician F with no feeling for analyzing data, I think I'd almost rather turn out proto-statistician B although I wouldn't feel it would be fair to him or her, because it would not leave them in the position of practicing defensive mathematics after they came out. That would be protecting them and protecting their position among the purists in the field rather than what was needed to do the job. They would feel maybe they need some more mathematics. Now, this is an uncomfortable side of things—it is much easier to teach mathematics, that is to those who will be taught, than it is to teach some of the other things.

There are places where it can hurt badly. I may have caricatured my position a little, but not very much.

Q: How about consulting? Should that be an important part in the training of Ph.D. statisticians?

A: If you can do it right. If you have a department where none of the professors have ever done any consulting, then it's—to say the least—dangerous. You can do some of that by interaction. The old Applied Statistics seminar here that I ran for many years which pulled in people from outside to talk about their problems and who were told there were only two axioms for the seminar, (1) they need not know the answers to the problems they talked about and (2) the audience could ask almost any question any time. I think that had many of the virtues of a formal consulting situation as far as the graduate students were concerned.

You heard Karen yesterday say that what she was talking about and what she's worked on mainly since she got her degree, came out of a graduate student project. There were a number of graduate students who had a feeling they just didn't have any feel for data and so we got together and we decided to try to do something with some of the cancer atlas data. The better consultants you got on the faculty, the more importance you can afford to give to various sorts of consulting arrangements. Because you'll be able to teach the students not only the feel for data and thinking about data, but also about interacting with people in the consultant's role. And that's almost equally important.

Q: What is your present view on EDA. How should it be taught, to whom should it be taught?

A: Well, as some of you know—most of you don't—in principle, work is going forward for a second edition. And the main thing that will happen is some of the more unnecessarily complicated things in EDA will go out and be replaced by simpler ones. There'll also be a few things that have come along since then and need to be added.

How should it be taught? I guess my only answer to that is "Whatever way a really interested teacher wants to teach it!"

Who should it be taught to? In one extreme, David Hoaglin taught it in a graduate course at Harvard for a while. He was thinking of teaching it for a quarter and the students wanted more. Charlie Smith's mother tried some teaching in high school with it.

To whom should it be taught? I think anybody who is willing to stand for it. There is a famous example that I can't report in complete detail because I've forgotten some of them—of somebody at Chicago who really couldn't stand it and ended up by starting statistics courses in three different divisions of the university at three different levels from undergraduate to graduate and kept dropping them because they started to teach EDA. There are people like that—not very many, we hope—it doesn't pay to teach it to them. But, if you're going to teach it to people who have a statistical background that's more difficult than teaching it to people without. But they are entitled to get more supplementary material and some indication of how things lock together or do not lock together.

Q: In your case, a second important part of your education was chemistry.

A: First! I have as many chemistry degrees as pure mathematics and none in statistics.

Q: Do you think it is a good idea to have an undergraduate degree in statistics?

A: It seems to me the places who do this and take it seriously do fairly well. Have you ever seen the paper "The Education of a Scientific Generalist"? This represents an optimistic view of the kind of diverse education that it might pay to give some people who wanted to head for scientific generalist. Those who haven't heard of the paper should know that the four authors are Hendrik Bode, who was at Bell Labs and was the man who had a lot to do with feedback technology, and Mosteller, Charlie Winsor and myself. We didn't have any great difficulty agreeing on something to put down as a proposal. But I think the answer to this one depends on what you've got in the way of an academic organization to fit into. Doing this in a mathematics department, at any place, Yale, Princeton, Harvard, Brown, can only—I think—be described as totally unfeasible. If you have an interested statistics department and people in other departments who are willing to at least be useful contact points and so on, the situation is very different.

If there was no reason for diversity the University of North Carolina would not require three statistics departments which is sort of historically what they had. I don't know the institution as an establishment at all, but I guess with much less than three you couldn't have covered the waterfront as well as you did—you probably didn't cover it far enough anyway. It's the practicality of fitting into the establishment that controls that one, not otherwise.

Q: How would you organize a statistics department?

A: Do the best you can fitting into the establishment. Probably straining the establishment a little, but not too much. A very delicate operation. As I understand things, Don Rubin at Harvard has actually made enough contact over enough time with the economists that really at that particular institution the economics–statistics gap is maybe almost gone. That's a thing that many other institutions—I think—would like to copy, if they knew how.

QUALIFYING EXAM QUESTIONS

Q: If X is a Poisson random variable with expectation 2, what is its median?

A: I have difficulty inventing a third answer. There are at least two respectable answers. The strictly formal answer that says which jump in the cumulative includes $p = \frac{1}{2}$. That's formally correct, but not very helpful. Now, let's see. I have to do a small amount of mental calculation. If you like to use halves, the cumulative at $1\frac{1}{2}$ is clearly less than $\frac{1}{2}$ and the cumulative at $2\frac{1}{2}$ is bigger. And so, I would plot the cumulative at $1\frac{1}{2}$ and the cumulative at $2\frac{1}{2}$ and draw a straight line connecting those and say where that line cuts cumulative equal $\frac{1}{2}$ is a respectable definition of a median for this situation.

Q: The answer is 2. What's the answer when the expectation is 10?

A: Well, one would have to compute, wouldn't one? Whichever definition we use, the question is: Do I have to compute one value of the cumulative of the Poisson or two values of the cumulative of the Poisson? (You would have to compute two, yes.) To be sure, probably you compute two and if you compute two, then you can do the interpolation. So, the computational load for the two definitions is approximately the same.

I am not that well acquainted with what it is you really want to look at here—the percentage points of a chi-squared and use Wilson–Hilferty? Can I have a show of hands of how many people know Wilson–Hilferty as such? 30%.

Q: You don't need to use Wilson–Hilferty.

A: You mean $\mu - 0.667$ is close enough?

Q: No, the answer is 10.

A: You mean the first answer is 10. The second answer is probably 10 and a bit, but I don't know. Let me indoctrinate the audience on Wilson–Hilferty. Wilson–Hilferty says—this is the other E. B. Wilson, Edwin Bidwell Wilson, and a lady named Hilferty, about 1920—that to get a respectable percentage point of chi-squared you take ν times the cube of the expression "$1 - 2/(9\nu)$ plus the corresponding normal deviate times the square root of $2/(9\nu)$." This is remarkably good for ν from roughly, say, 2 up. You don't need it for 1, because you can get the percentage points for chi-squared on 1 out of the Gaussian table.

But this is one that every statistician ought to have in their back pocket.

Q: When testing for significance for contrasts under conditions of multiplicity, we now can control the "false discovery proportion" instead of the family-wise error rate, thereby increasing power and rendering findings relatively invariant over changes in family size. With large families, the advantages are considerable. Two questions. How should we most constructively think about how to extend such advantages to the establishment of confidence intervals or what impediments stand in the way of such extensions?

A: How many people—I am going to continue this polling process—how many people know about the Benjamini and Hochberg stuff? Running 20% maybe, okay! Those who look at big multiple comparison situations, I think, predominantly feel that 5% simultaneous is being too stiff and that 5% individual is being too darn loose. Now, long ago I suggested using the average of the two guides as an intermediate significant difference, I never pushed it very much and I doubt if I could have sold it real well. Last year or two, Yoav Benjamini and Yosef Hochberg, both in Israel, have been working on the idea of controlling the fraction of the positive statements that you make that are wrong. Now, what this means, you see, is if you have things being compared and there are going to be very few significances, no matter how assessed, the FDR is going to be very much like the individual rate because you are going to make maybe one positive statement on average. You are entitled to make, by this formal thing we have to correct in a moment, a twentieth of 1. On the other hand, if you have a point, a point, a point and the standard error is about this much (using hand), there is only one difference left that you haven't cornered. Then this thing (the FDR) is going to tell you, you are going to behave very much like the individual thing, and that's fair enough, it's really simultaneous on 1. It does seem to have good properties. Some of us, Lyle (Jones) included, believe that the first place you come to on this is a matter of direction only; that is, the positive statements that you might make are that they differ in this direction or they differ in that direction. But the statement that they nearly differ is silly, because they all differ in some decimal place anyway. And if you talk about directions, then when you are getting very few definite ones, half of them have to be wrong because clearly these are things that have come about from small differences. That's oversimplified heuristics, but roughly right. So, you have to fix up the game somehow. The two Y's like to fix it up by saying that if you get 0 over 0 you count that as one full case of 0. I like to fix it up by saying if I'm doing things at 5%, I'm entitled to make $2\frac{1}{2}$% as many false positives as 1 plus the number of positives. You have to seed things somehow at the beginning and there's a choice on how you do this. If Stephan will read the first part of this question again, I will tell you why I think it is not adequately formulated.

Q: How should we most constructively think about how to extend such advantages to the establishment of confidence intervals?

A: The point is that confidence intervals are not confined—the positive confidence intervals are not that different from negative confidence intervals. And, my guess is—and I haven't thought this through long enough—but my guess is you end up using simultaneous things for the confidence intervals anyway. I doubt that a very hybridaceous thing that says, well we aren't going to make a positive confidence interval statement erroneously more than $2\frac{1}{2}$% of the time we make positive confidence interval statements and negative confidence interval statements we don't want to be wrong more than $2\frac{1}{2}$ experiments out of a hundred. That sort of hybrid thing that says: "once the statement ceases to be positive, you change the bases on which you evaluate suddenly"—I don't see how to make that one fly anymore. I think I have a suspicion it's never going to fly. So, I don't see anything wrong in principle with using the false discovery proportion for directionality and the simultaneous calculation for the confidence intervals. The fact that I will get some upward directional statements where the confidence interval includes values less than 0 doesn't bother me very much. I don't have to have a seamless connection between the two. I offered an extension that is viable and the impediment that stands in its way is it's not as logically seamless as you might like. But I think it's better than any seamless one I see. Directional statements versus "I'm uncertain about direction" is a very great difference. A confidence interval that just doesn't quite cover 0 and one that just does cover 0 are very much nearly the same thing. They ought not to be connected in a seamless way.

ARGUING THE FIDUCIAL ARGUMENT WITH R. A. FISHER

Q: The correspondence you had with R. A. Fisher about the fiducial argument ends suddenly. You offered some counterexamples, and he said you were foolish. Then you offered a different counterexample, and he said something else that was sort of rude, and then you said you were going to England to visit him and that was the last was in that correspondence. I was wondering if you could finish the story.

A: Well, I was talking with Sir Ronald in his office. And I think roughly what I said was that I didn't see the logical strength of the fiducial argument but I gave a lot of weight to the fact that he thought it was a good idea. At which point he did his best to show me the door. Since Elizabeth was out in the garden in the other direction, talking to one of the daughters, I didn't get shown the door. So he grabbed his hat and his cane and went toddling out the door himself. I would say that was the end of the correspondence.

Q: So, we never got to the bottom of the fiducial argument then?

A: No, we got to a place where it was—I think—mutually felt that further debate between these two parties would not get us any deeper. Now, whether we got to the bottom at that point or two different bottoms or whatnot, I leave that for other people to judge.

STATISTICS 411

Q: We heard a lot in this session and yesterday about a course Stats 411 for undergraduates. Can you explain what this course was, what sort of topics were involved and what happened to the notes?

A: Well, I'll answer in reverse order in part. The notes take up about a filing case upstairs. That's what happened to the notes. I wouldn't be surprised if you could get a better answer from some of the people who took the course than you can from me. As to what was in it and so on. Who wants to volunteer?

Q: I've got part of an answer. I taught 411 two years ago and prior to doing that I was suitably humbled and I asked John if he would let me have a copy of his notes and he did. He sent me—not all of it—but a subset which I'll be happy to share if it's all right with John—and I read them and they're fascinating. But there wasn't a chance in the world that the students that I knew would be able to understand it—at least with me as a teacher. So, we did other topics. This was a course for seniors in engineering. And it tends to be the last course they take.

A: This is *not* an adequate description of the situation when the course was being given by me. This was a course—as somebody had said around here—for seniors majoring in the department, graduate students, etc., etc. and strays from all sorts of places. Henry Braun said he took the course for six years in a row while he was a faculty member. The one thing that I always used to do was to get the seniors to sit up front at the table and make it clear I was going to answer *their* questions before I answered anybody else's. Otherwise I think the moral pressures on the students would have been bad. What it was, was an attempt to start at the beginning and be serious about it. Not serious mathematically, but serious statistically. That might mean a couple of weeks talking about one-sample questions, taking the view that the standard assumptions are almost guaranteed never to be the truth. So, you want to understand what happens when they're not true. Does this accord with your readings, Howard?

Q: I think I tried to do that.

A: I think Stephan certainly sat through this course at least once.

Q: Twice, I think. The course was like this. John came in with a Ziploc, the lectures were nicely parcelled out, the course was in topics, numbered, and the numbering varied from year to year. So there were several numbers for each topic. A topic could be 14N but at the same time also topic 9 for the course given in 1979. Clearly the course content evolved quite a bit during the years. It was well structured. It started out with single-sample—single-batch—questions. How do you estimate location? How do you estimate scale? What do you do if you have several of them? It went on I think—in the years I took it—up to ANOVA.

A: It included (ψ, w)-technology, (g, h)-technology, orstats, gaps, simultaneous confidence, lots of things.

THE FUTURE OF DATA ANALYSIS, INCLUDING STATISTICS

Q: What do you see as the future of industrial statisticians?

A: I never was on the industrial firing line in any real sense anyway. The more difficult question—I think—is how much better will people making greater use of statisticians fare in the industrial competition than those who make lesser use of statisticians? Again, this is an establishment question. Some years ago, there was a period of half a dozen years when IBM didn't talk to BTL because the lawyers were feuding over the patent agreement. And after that relaxed, there was a delegation of brass from the IBM research laboratory who came to Murray Hill, and a detachment (from BTL) that went up there. I think the two largest surprises—because, you see, at this point you had people in the range of administrative levels that would get involved, who wouldn't have been involved seven or eight years ago and they really didn't know anything in detail about the other organizations. The thing that struck them the most, I think, was how many statisticians there were at Murray Hill and how few there were up in suburban New York. And this was at a time when some people would have said these were the two best industrial laboratories, or maybe, two of some very small number of the best. So, even in research, things were not uniform—now, I don't think that IBM's difficulties came about from not having more statisticians—it would be nice to think that, but there were other reasons. They were generally a stick-in-the-mud outfit as far as their computers went. And they were shocked when they first realized they were spending more money on software development than on hardware development. If industrial statisticians are going to flourish, they are going to have to do different things than they used to. Industrial arithmeticians are not going to get hired as such. Arithmetic is a well-stabilized field and to the extent that it's needed, people pick it up. But statistics and data analysis isn't at the moment a well-stabilized field, and hasn't been, and there can be a particular need for having people who are reasonably up-to-date and are growing forward and in various directions. Whether this means TQM or not, I don't know. George isn't here, we cannot put the bite on him and see what he had to say. I think there's a strong future if they adapt well enough to the changing needs—and I just wonder as a whole, how many of them have been changing at all. The industrial statisticians, in double red quotes, that we had here these two days include Colin who is talking about the kind of experimental design that would drive a classical experimental designer up the wall and John Chambers who was talking about the origins of S and where it might go and whether it might accommodate EDA. Times are changing and if the industrial statisticians don't change, other people will do the change—maybe called something else. Remember the early days of cathode ray tubes—TV display tubes—the net result of the situation was to invent a new profession, called shrinkage analysts. Because at that time about 3% of the things that started out to be TV display tubes came out at the other end of the process as satisfactory finished ones. So, what you had to analyze was how things shrank as you went along. Now, the fact that they got to be shrinkage analysts says that the industrial statisticians either weren't there or that the industrial statisticians at the time didn't see the need to think differently.

I see Stu in the background there. And about the only thing we are willing to fight about publicly is Taguchi, I think. I regard much of what Taguchi said as overexpressed, unwarranted language. But it is not clear to me that in the early stages of fixing processes you need the degree of security that comes with classical experimental design. And that if you operate à la Taguchi you may get most of the gold at the grass roots and after that, if you don't go and do something better, it's just too darn bad. But that doesn't leave me feeling bad. Now, I don't have to deal with these people who learn Taguchi as a watch word and don't know anything about him. Those who do, have my deep sympathy.

Q: I tell people, John, that they should listen to what the gentleman has to say philosophically and avoid his technology.

A: What I'm saying is maybe the opposite. Don't believe what he says about the security of his results but in early days the technology may be a good one. Maybe—I don't know. I have not been on the firing line in this direction.

Q: It's like pulling weeds.

A: Well, I pull weeds.

Q: My question relates to a question I asked you in 1961.

A: Well, I hope I won't give the same answer.

Q: The question at that time was—I was starting out in statistics—I asked you what I should read. You told me, the early *Journal of the Royal Statistical Society*, *Supplement* and Discussion. I found this a very important learning experience and have also been using it in teaching. I was wondering if a young person came to you today, what would you tell him to start reading?

A: I guess you have to start at the same place, because to start anywhere else, you assume that they are a lot further along than you are when you start. And that one refers also to some of this thing about consulting, because the nearest thing to a surrogate for consulting that I know is to go and read the supplement to *JRSS*. This is something that only requires a library and it's not going to penetrate nearly as much as experience would, but it's going to penetrate in some of the same directions. There was a reasonably savvy group of people who were in the Industrial Application section of the RSS in those days and the agricultural departments, and a lot of this does come through, as you and I both know.

Q: How many Ph.D. students did you have over the years?

A: Stephan, you have been looking at lists, maybe you know?

Q: I don't know.

Q: Eileen said last night that you had over 40.

Q: What have been some of your greatest satisfactions and regrets over the years?

A: I've avoided the classification.

Q: What do you see as the future of statistics in Princeton?

A: (Laughter.) I have various sized crystal balls, but none of them big enough for this. We're in a time of academic retrenchment. We're presumably going to lose some good statistical departments in some places. If there were sufficient numbers of analogs of me, so that teaching things could be tried out and enough books written, then.... There's a basic difficulty which Dick Link would present as saying that a good statistician must be a schizophrenic, because he has to deal with uncertainty and the measurement of uncertainty—that's his main task—and to do this using the most certain tool we have which is mathematics. He has to bridge across the gap here. I don't see any reason to believe that statistics in a mathematics department is going to be other than a hard and dangerous life. On the other hand, statistics as a separate entity is going to have a different set of reasons for being a hard and dangerous life. Paul Velleman was arguing about statistics being a science. I would tend to think it would have been more accurate to say science-and-technology. We face the facts, but the academic technologists have by and large ceased to teach engineering. And the way academic society is organized it's not clear how a pure technology acquires intellectual stature, except through individuals. But for my money, statistics, along the lines of the Mosteller and Tukey review in the *Handbook of Social Psychology*, "Data Analysis, Including Statistics," is a pure technology. In physics, theoretical physics draws the attention of most of the undergraduate students or most of the graduate students. Now, that's maybe a good thing in a backhanded sort of way. Somebody was telling me yesterday that they're making 1400 physics Ph.D.'s a year in this country and there aren't going to be that many who do theoretical physics. On the other hand, theoretical physicists, it's been well established, can be converted to almost anything. Experimentalists probably can't. So maybe it's good that theoretical physics attracts the crowd. The point, is though, that physics has had an old well-established intellectual reputation. And I don't think we'd get away with training 1400 Ph.D.'s a year in statistics and have most of them leave and go into all sorts of other things. I don't think that would be as acceptable as it is for theoretical physicists. There are some internal contradictions at very high levels. If you wanted to ask me where I think I failed the profession most, it would be in the direction of (not) doing something about this.

REFERENCE

BRILLINGER, D. R., FERNHOLZ, L. T. and MORGENTHALER, S., eds. (1997). *The Practice of Data Analysis*. Princeton Univ. Press.