Statistical Science
1994, Vol. 9, No. 2, 285–303

A Conversation with William Kruskal

Sandy Zabell

Abstract. William Henry Kruskal was born in New York City on 10 October 1919. His basic education was primarily in the public schools of New Rochelle, New York, a suburb of New York City. He attended Antioch College for two years and then transferred to Harvard College, from which he received the S.B. degree in 1940 and an M.S. in mathematics in 1941. Then he went to the U.S. Naval Proving Ground in Dahlgren, Virginia, first as a civilian and later with a USN commission.

After World War II, he worked toward the Ph.D. in mathematical statistics at Columbia University, but joined the faculty of the newly formed statistics group at the University of Chicago before he completed and received his degree in 1955. He has remained at Chicago except for a summer at Harvard University and visits to the University of California at Berkeley and the Center for Advanced Study in the Behavioral Sciences at Stanford.

Kruskal's research and teaching are closely linked. Among his primary research areas have been linear structures, nonparametric procedures, the taking of censuses, government statistics in general, the history of statistics, clarification of such concepts as representative sampling and normality, miracles and statistics, and the relative importances of cause-like variables.

From 1958 to 1961, Kruskal was the Editor of *The Annals of Mathematical Statistics*. He was a member of the 1970–71 President's Commission on Federal Statistics, out of which grew the present Committee on National Statistics in the National Academy of Sciences–National Research Council. Kruskal headed this Committee during its first six years. A different kind of activity was his editorship of the statistical part of the *International Encyclopedia of the Social Sciences* and his co-editorship (with Judith M. Tanur) of the *International Encyclopedia of Statistics*. He has for years been a trustee of the National Opinion Research Center (NORC).

He held a Senior Postdoctoral NSF fellowship and was a Fellow of the John Simon Guggenheim Memorial Foundation. He has held a number of offices in professional organizations, including the presidencies of both the Institute of Mathematical Statistics and the American Statistical Association.

At the University of Chicago, Kruskal chaired the Department of Statistics for six years and later was Dean of the University's Division of the Social Sciences for nine years. He also served in 1988–89 as Dean Pro Tempore of what is now the Irving B. Harris Graduate School of Public Policy Studies. Since 1973, he has been an Ernest DeWitt Burton Distinguished Service Professor (now Emeritus).

In 1942 he and Norma Jane Evans, alas no longer alive, were married. There are three children: Vincent Joseph, Thomas Evan and Jonas David.

Sandy Zabell is Professor, Departments of Mathematics and Statistics, Northwestern University, Evanston, Illinois 60208.

FIG. 1. *Taken in approximately 1935—left to right, front, is Joseph (father), Joseph Jr. and Lillian (mother). Rear is Molly, William, Martin David and Rosaly.*

The following conversation took place in Chicago, Illinois, during the fall of 1991.

Zabell: Bill, perhaps we might start with your family background and early education.

Kruskal: Fine. My family comes from central and eastern Europe. My father arrived in America when he was about six; my mother was born here. My father was an energetic and successful businessman, not in any towering way, but in a comfortable way. When I arrived, the family lived in New York City, but I hardly recall that. We moved to New Rochelle, a suburb of New York, when I was small and I went to public schools there through the eleventh grade. For grade twelve I attended Lincoln School in the city.

I had two sisters and two brothers. My sister Molly is dead; my sister Rosaly lives in Israel. My two brothers are Martin David and Joseph. They both are in the course of vigorous mathematical careers, Joe's substantially in statistics.

ANTIOCH AND HARVARD

Zabell: How did you, a New Yorker, come to attend Antioch College in Ohio?

Kruskal: I went to Antioch College in 1936 for my first two college years mainly because I was attracted by the co-op plan whereby you had a job and attended college in alternate periods. I had two mathematics teachers at Antioch, Irving Burr and Max Astrachan, both of whom later went into statistics. Irving Burr taught statistics at Purdue, and Max became active in industrial quality control.

Zabell: When did you shift to Harvard?

Kruskal: In 1938 I discovered W. V. Quine at Harvard and became enthusiastic about the opportunities. So I transferred to Harvard for my last two undergraduate years. . .wonderful years. For example, I met Saunders McLane, then a young faculty member at Harvard, and of course Quine himself. My first teacher of statistics was E. V. Huntington at Harvard. He did not give a really fine course, mainly working in boring detail on the analytic geometry of the bivariate normal distributions, the family of ellipses and so on, but he was a fine person.

Zabell: Which department was Huntington in?

Kruskal: Mathematics. He is well known for his work on the mathematics of splitting political representation when it has to be in unit terms because you can't split people, and almost any arrangement can lead to paradoxical results. As you know, there's been a recent flurry of interest in this, and in fact the question comes up in connection with Census accuracy, where I have a current concern.

Zabell: Was interaction with Quine the first experience that really awakened your interest in mathematics and statistics?

Kruskal: Well, yes. While at Antioch, I went through a frequently found syndrome of worrying about how you know things are true and asking—in backward sequence—why, why, why, and ultimately finding there is no end to this. Those questions brought me to the relevant literature in philosophy and logic, and that led me to Quine, who was a great beacon. Then I realized that he was in Cambridge, at Harvard, and that one could actually get to talk with him. I always was in awe of his work even when he says puzzling things, perhaps especially because he says puzzling things. The resources of Antioch were tenuous compared to those at Harvard, so I managed to transfer.

Zabell: Did you end up taking a typical mathematics major or leaning toward logic?

Kruskal: I sat down and tried to decide what courses I wanted and what teachers I wanted. The result was a joint program in mathematics and philosophy. Then I stayed on for another year and got a masters in mathematics. In 1940 the war was looming. It seemed to me that it might be sensible to point toward possible military activity. In fact, that's what came about.

DAHLGREN

I went to the U.S. Naval Proving Ground in Dahlgren, Virginia, part of the tidewater country. I met various people there, including Herman Chernoff. We were in different parts of the Proving Ground. My little area was aviation ordnance, and mostly what I did, at the beginning anyway, was connected with the Norden bombsight, which had some reputation at the time. For example, people said you could drop a bomb in a pickle barrel from an altitude of 5,000 feet. I wasn't sure what a pickle barrel was.

Zabell: Who brought you to the Naval Proving Ground?

Kruskal: The chap who brought me to Dahlgren was L. T. E. Thompson, a physicist there. I think he was seeking young people, and he offered me a job. So I went there as a civilian. To jump ahead a little, later on I was commissioned in the navy partly because civilians had a difficult time at Dahlgren; for example a civilian might have special trouble getting classified documents.

It was at Dahlgren that I got seriously interested in statistical issues as epistemological problems.

Zabell: So you learned about statistics on your own at Dahlgren?

Kruskal: Well, I read on my own, and I met people. Another person at Dahlgren was Francis Dresch—Fran Dresch—who had recently gotten his Ph.D. under Neyman at California. I learned a lot from Fran. Also, Brock (Brockway) McMillan was at Dahlgren. There were other interesting people. Then various statisticians came to visit. Allen Craig, of the Hogg and Craig textbook, came. Neyman came perhaps more than once, and Hotelling came on one or two visits. The Hotelling visits were arranged by the statistical research group at Columbia. Allen Wallis' 1980 memoir in *The Journal of the American Statistical Association*, with my comments, lays out the details. The statistical research group at Columbia, as you know, was very active during the war. It included Wald, Wolfowitz, Wallis, Bowker, L. J. Savage, Mosteller and others. (There was also a statistical research group at Princeton and I think that Wilks, Tukey, Scheffé and others made their contributions via that Princeton group.)

FIG. 2. *William Kruskal in* 1974.

[*Note*: A history of Dahlgren is available by Kenneth G. McCollum (ed.), *Dahlgren*, Naval Surface Weapons Center [new name], Dahlgren, Virginia; June 1977.

Statisticians interested in coincidence might be amused to know of another book by Kenneth G. McCollum: *Nelson Algren (A Checklist)*, Bruccoli-Clark, Gale Research, Detroit, 1973.]

Zabell: So what view did you have of statistics, coming from the outside?

Kruskal: Well, many came from the outside—for example, Tukey and L. J. Savage, but not Allen Wallis, who already was immersed in statistics as an economist.

Zabell: But you came without preconceptions. What sense did you have of the field?

Kruskal: It was a welcoming sense; here was something important that other people had been thinking about. For example, these bombsights were shipped down from New York City, where they were made, to Dahlgren. They were given an extensive check in our shop. There was a wonderful civilian there named Middlebrook who was in charge of that side of things. Then they were flown, or some fraction of them were flown. They were put into a plane, an expert operator (the bombardier) operated them, and there were actual bombs dropped. . .not real bombs, but sheet metal simulacra filled with heavy iron ore that had laboriously been shipped from Labrador, or somewhere else in Canada. This process had been worked out well before I got there, mainly by the physicist I mentioned, L. T. E. Thompson. Navy seamen on the shores of the Potomac, three of them, each with a theodolite, were the observers. The bomb would be dropped from the airplane at a target, a platform built in the Potomac. There would be a splash of water, quite a substantial splash, and the sailors were supposed to rotate their theodolites till the crosshairs were on the splash, and then write down the scale reading. You'd get three rather tattered pieces of paper coming in with lists of run numbers, degrees and minutes. Then I, or an assistant, would take a prepared, map-like sheet and with a straightedge draw lines of sight. In principle, the three lines of sight would intersect at a nice point, the point of impact. Then you would have the geographical heading of the plane—and you'd lay a plastic grid over the maps and read off the errors in range and in deflection, perpendicular to range. So each splash gave you a two-dimensional vector of components of error. Well, of course, the people doing the measurement were sometimes inattentive, or made mistakes. The lines would not cross at a point; sometimes they were way off, so you would get quite a substantial triangle. My instructions, when I first came, were to take the middle of the triangle by eye; I did that, but I stewed about it. Later, I learned there was literature on this topic and I tried to develop an approach without success, but it was a genuine statistical problem. Much later, I learned that Henry Daniels had worked on a related problem. It was more complex, with spherical geometry and great circles on the Earth.

Zabell: What were some other statistical problems encountered at Dahlgren?

Kruskal: I managed to get my Naval Aviation colleagues to do replications with the same bombsight. The usual procedure after a bombsight was flown was to take it back to the shop, readjust it and take it up into the air again. But there didn't seem to be any clear relationship between the two results (before and after). On the other hand, it was a sample of bombsights that had done poorly at first. So we did an experiment, a real experiment as I would now call it, where bombsights were not chosen in this way, but maybe in order as they came off the assembly line. You would send one up in the air and get the results, bring it back and send it through the regular shop procedure, and then put it up in the air again. So we used little blocks, twofold blocks, and there didn't seem to be any connection. The variability was apparently all due to weather conditions, day-to-day changes in the bombardier's ability and so on. Anyway, Hotelling came from the Statistical Research Group, and I prepared data that he used to exemplify his study. His analysis took a dreadfully long time. By the time he sent us his paper, this whole program had been discontinued for reasons that I never really knew. This is all described in my 1980 discussion of the Wallis paper. The work was originally classified for security reasons, but later declassified.

Zabell: It's interesting to see that your first real experiences with statistics were so closely tied to applications of immediate importance.

Kruskal: Yes, these were genuine statistical problems. Here's another. At one point we were comparing two different bombsights, the Norden and some other that had been proposed, and my commanding officer said: "Well, why don't we put in a third bombsight? We have the setup all there, the bombardiers, the planes, the theodolite operators etc. If there's a third bombsight, it would be interesting to include it." So we did, and examined the results. If you only had looked at the two primary bombsights, the difference between the two would be of statistical interest in the usual significance sense. If, however, you looked at all three, the third one would fall between the other two. Grinding out a standard analysis of variance would lead to no statistical significance in the usual sense, and that seemed to me a big puzzlement. It still seems to me a big puzzlement. Since then, of course, we've had multiple comparisons and other approaches, but I still think it's an intrinsically difficult and interesting problem that comes up all the time.

Another example: there was interest in simulation; if we could only simulate this bombing business, think of the money we would save—all that gasoline, people's time etc. So we had built a sort of cart with a raised platform on it, and it rolled over the big hangar floor. It had the bombardier in it pretending to be on a plane. It was simulating the bombing procedure, so then you'd have a simulated point of impact. Some colleagues decided to analyze these by looking at averages. I claimed that you should do it by looking at variability. Average errors you can always correct, I thought then, by just a change in some knob. That was an introduction to the question of what criteria do you want to look at in a quantitative study. I think of that as a bellwether theme, because it has held my interest ever since. What aspects of your measurements do you really want to examine and what interpretations do you want to give them? It comes up everywhere, for example, in the work that Leo Goodman and I did on measures of association for cross-classifications. That stems directly from the theme, because most of the existing measures of association for cross-

FIG. 3. *October 26, 1973—on the occasion of an honorary degree being awarded to Fred Mosteller. Left to right is Michael Perlman, Leo A. Goodman, Paul Meier, Fred Mosteller, Michael Wichura, David Wallace, William Kruskal, Patrick Billingsly, Shelby Haberman and Joseph Whitaker.*

classifications were not interpreted in any serious way by the people who used them. Those joint papers argued that in statistical work you should think in terms of sensible interpretations. We used the term "operational interpretations," but maybe that was too specific.

Zabell: These episodes illustrate how many people in American statistics got into it because of wartime experiences. Like you, they were in some subject perhaps related to, but still different from, statistics: in your case logic, and in Tukey's case topology and chemistry. Yet as a result of the work during the war, many emerged afterward as people committed to statistics.

Kruskal: I think what you say is absolutely right. Decisions had to be made all over the place, and people came to them by different channels. For example, Joe Hodges and Erich Lehmann, later at Berkeley, got into it in part through operations research in the Pacific, analyses that quickly got one's hands dirty.

Zabell: I understand that you left Dahlgren and later came back. Can you tell us a little about that?

Kruskal: I went to Dahlgren in 1941, and about that time I had the truly good fortune to marry Norma Jane Evans. On the other hand, the living conditions at the Naval Proving Ground were poor, so Norma and I struggled with miserable temporary housing (which is still there). We got restless and decided I'd be better off wearing a uniform. So I applied for a commission in the Navy. I had some problems: first height, I was half an inch or so under the minimum, and second, I'm color deficient. So I couldn't possibly stand on a bridge on a foggy night, look at a light and decide whether the light on another ship was coming towards me or going away from me. That was negative from the Navy's point of view. But one way or another I got the commission.

Then I went to work with Alston Householder, a mathematician, who later was at Oak Ridge for many years and who has a fine book on linear computation. Alston was with a psychological group at the D.C. Naval Gun Factory on fire control, mainly antiaircraft fire control, so it was not unrelated to what I'd been doing at Dahlgren. The group was headed by a statistically inclined psychologist named Bill Kappauf, who was at Urbana and went back to Urbana after the war. I haven't kept up with him. So I went to the Gun Factory for perhaps half a year. Then the folks at Dahlgren managed to get me back. You weren't supposed to go back as an officer to a place you'd been as a civilian, but they created an exception. The trouble was that when I got back, there really wasn't much for me to do.

FIG. 4. *Navy officer William Kruskal, Washington, DC, 1943.*

COLUMBIA UNIVERSITY

Zabell: When did you leave the Navy?

Kruskal: I left active service in '46, went into the family business in New York for a couple of years, and started taking courses at Columbia University, which at that time had extensive night courses. I had to get special permission for one course because I didn't have the prerequisites. It was a course in multivariate analysis taught by Ted Anderson. In fact, it was the first course he taught after getting his Ph.D. at Princeton. I'll never forget my interview with Abraham Wald, who was chairman of the department and whose permission I needed to take the multivariate course. I hadn't had the basic introductory course or anything intermediate. So he said, "Well, you try it." He was realistic and sweet. It was amazing. I had a difficult time, but it was stimulating and informative. Charles Stein was in the class, just head and shoulders above everybody else in ability and insight. Charles took the notes for Ted. I think that Ingram Olkin was in the class, and I can't remember others. But I ground through it and it served as quite an introduction. Then later on I took the basic course taught by Wolfowitz.

Zabell: What was Wolfowitz like as a mentor?

Kruskal: He was violently negative about Bayesian approaches and he didn't mind saying that and excoriating people who were positive about Bayes. So that was part of my initiation also. On the other hand, I really learned from him. Here was a man who was primarily a theoretician with very strong views. I got good lessons from him about the importance of looking at measurement methods, the data and so on.

Zabell: Whom else did you work with at Columbia?

Kruskal: I worked with various people: Henry Scheffé and Howard Levene, in particular. Wald gave great courses.

Zabell: When did it start becoming clear to you that statistics was going to be a long-term professional interest?

Kruskal: The decision was made around this time. There was the question of staying in the family business or going into statistics. Should I stay at Columbia or should I go to some other place? At that time the most immediate other place to come to mind was Chapel Hill. For example, Ingram Olkin went to Chapel Hill after having been a student at Columbia. In principle I should have thought about Berkeley, but I didn't really know much about Berkeley; it had not yet achieved the fame that it did soon after. This was in the summertime and I recall going to see Jack Wolfowitz at a summer cottage he had on the Hudson, an hour's drive from the city. Norma and I went there and talked to him. His children played in the sand; one of them is now a significant figure in our Department of Defense. And Jack was shocked that anybody should spend more than two moments deciding whether to go to Columbia or North Carolina. "How can you dream of doing anything else, but coming to Columbia?" So I did and stayed there in residence for two years.

Then it was 1950 and I was anxious to get a regular position; I had a family and I was older because of the war. Of course, there were many others in that position. I had an invitation from Bill Madow at Urbana; he was more or less in charge of the statistics operation which was part of mathematics, and that was attractive. We had been living in Manhattan on 22nd Street, just a terrible place to raise kids. Norma said, "Let's go to a semibucolic place," and Urbana sounded right. Then someone told Allen Wallis and Jimmie Savage about me. I had not met Jimmie Savage, but he had known my brothers. Another connection was my admiration for and friendship with I. Richard Savage, brother of Jimmie and also a graduate student at Columbia. Allen invited me to come out and meet people in the newly formed statistics group called the Committee on Statistics. The half-joke is that it was called that because Hutchins was against the over-departmentalization of knowledge. I don't really know how much truth there is in that; perhaps a good deal. So I had a very exciting trip to Chicago. I met Raj Bahadur, Leo Goodman, Murray Rosenblatt, Tjallings Koopman and of course Jimmie Savage and Allen Wallis.

Zabell: This was in 1950?

Kruskal: It must have been in the spring of 1950. I had in my pocket an Urbana offer that I was considering sympathetically. Chicago made a competing offer and I was embarrassed to present this to my wife, because she very much wanted to move out of big cities. What on earth would be the sense of going from Manhattan to Chicago, another big city? Of course, the University of Chicago area is nothing like Manhattan. So I put Norma on a plane to go out and visit there herself. She came back and the first thing she said was: "Why didn't you tell me that the two salient things about Allen Wallis are (i) he's tall and (ii) he's got red hair?" I said "Well, I certainly should have told you about the height, but the red hair I just don't see very well." Anyway, she agreed, and we came to Chicago and I think that it was exactly the right thing to do.

EARLY WORK ON ROBUSTNESS

Zabell: I think other statisticians would be interested that you only got your Ph.D. in 1955.

Kruskal: Well, it took a long time. Things were expanding very rapidly, of course, and I had thought I would finish up my dissertation within a year or so, which is not unheard of, but then I didn't. I was very busy getting started, and I met with research frustrations. I was interested in what we would now call robustness against nonnormality for procedures that assumed normality, a central problem. My approach was via neighborhoods of a normal distribution. How much could you deform a statistical process while staying within a small neighborhood of the assumed distribution?

Of course, you had to get a distance that made sense and I explored metrics that other people had used. This is again a case of concern with meaningful characteristics. I finally decided that the L_1 distance was especially meaningful because it had a simple relationship to the biggest difference in probabilities over measurable sets. I could do something in the simplest case of just a sample average, but anything beyond that defeated me and it was extremely frustrating.

Soon after I got to Chicago, Allen Wallis proposed what led to our joint paper on analysis of variance for ranks, the H test. He had been thinking about it, but hadn't quite got the mathematics worked out. So I turned to that problem, which in a sense was complementary to my interest in robustness, for the use of ranks lessened dependence on parametric assumptions like normality. My dissertation finally was in two parts: one was what I had managed to do on the robustness front, and the second was the work with Allen Wallis. The latter resulted in two substantial 1952 papers, one of my own in *The Annals of Mathematical Statistics*, and the other with Allen in *The Journal of the American Statistical Association*. The work on robustness really never got published except for a 1951 abstract in *Econometrica*. Since then others have done an enormous amount of work on robustness.

Zabell: Was there someone at Columbia who encouraged you?

Kruskal: Yes, Henry Scheffé. But when I came to Chicago, Henry went to Berkeley, and Howard Levene became my Columbia thesis advisor. I'm most grateful to him.

Perhaps this was a case where I should have been warned away from a difficult topic, but the topic was attractive to me because it was, and remains, a basic problem throughout statistics. How sensitive are your results to errors in the assumptions?

Zabell: But it had a fruitful outcome in that it led you into the whole area of ranks and nonparametrics.

Kruskal: Well, that's right. Things are tied together in all sorts of ways.

Zabell: It's curious that your two papers on analysis of variance by ranks both appeared in 1952 and yet somehow it took Columbia three years to getting around to giving you a degree for work that was pretty famous.

Kruskal: Oh, I don't know. I guess I didn't push as hard as I should have. I was brought here to Chicago as an instructor, maybe the last instructor we ever had, with the understanding that I would be an assistant professor when I got the degree. Probably there were assurances from Columbia that the degree was coming, so I became an assistant professor after the second year here.

Allen Wallis himself only had a bachelor's degree and was almost proud of that. At least he made a point of telling people, for whatever reasons. Jimmie Savage was totally a free spirit in this and other matters, so it never really bothered me. I guess it took a long time to get the opus written up in dissertation-like form. The delay was largely my own fault; I was captured by other activities. After a while it became laughable and I decided I just had to push it through.

CHICAGO

Zabell: Describe what Chicago was like in the 1950's.

Kruskal: Paul Halmos was here at the time working with Jimmie Savage on their Radon–Nikodym paper and other things, and was a marvelous colleague. Marshall Stone was here, and Allen managed to get him on our side, which I think was quite a coup. Oh, lots of other people! Walter Bartky was dean of the division, an astronomer interested in statistics. Don Fiske and L. L. Thurstone in psychology, George Platzman in meteorology and Phil Hauser in sociology all interacted with us. In economics, a lot of people: Milton Friedman was actually a member of the Committee on Statistics for a few years. There were visitors in the early years, all kinds of wonderful people: Joe Hodges, Charles Stein, Dennis Lindley, Ingram Olkin, Don Darling.

In the department, I've already mentioned Leo Goodman and Raj Bahadur. Also, Harry Roberts, John Gurland, now at Madison, the late K. A. Brownlee. And then others started coming: Paul Meier, David Wallace, Pat Billingsley, Hans Zeisel, Esther Seiden, your now Northwestern colleague Shelby Haberman, Mike Perlman, Joe Eaton and yourself, Sandy Zabell. Also Kit Bingham, Gus Haggstrom and others.

Zabell: Now you say "the department," so the name must have changed.

Kruskal: The Committee on Statistics was established in 1949. The committee turned into a department in the mid-50's, although there was some opposition to it. Opposition, I think, for two reasons. First, because of generalized hostility to Statistics on the part of some chemists and physicists. Second, more immediately, a worry about budgets and the distribution of University resources among departments. A department is somehow more permanent than a committee. But anyway the change was made, and it reduced confusion since the group was from the start a department in all functional ways: its own faculty, its own students, its own budget, its own degree recommendations. Allen was rightfully proud of the change, which was really more than nomenclature.

Zabell: What role did Allen Wallis play in setting up the department?

Kruskal: Oh, Allen and Jimmie were really the people who set it up. Allen had been brought here from Stanford to get statistics rolling. He was marvelous at the interpersonal and administrative areas. He dealt with Hutchins and then Kimpton in masterful ways, at the same time remaining active on the scientific front. Jimmie was so stimulating. He was just starting on his subjective probability work and had not yet become inflexible in his views.

Zabell: What did you think when Savage first became interested in Bayesian methods?

Kruskal: Well, I thought: "Here's something interesting. I'm delighted to be a colleague of a person like Jimmie who is so intelligent and so serious about foundations." My own point of view then was, and still is, an eclectic one. It seems to me there are at least two kinds of probability, probably seventeen or so, and why knock oneself out to incorporate them into a single kind of probability?

Zabell: You certainly paint an exciting picture of the department then.

Kruskal: Oh, it was exciting. I remember in Jimmie Savage's kitchen composing the statements for the departmental announcements and other materials, most of which have been changed by now, but maybe not in essentials.

One other thing. Allen and Jimmie had brought me to Chicago in part for a specific functional need, to operate what Allen liked to call a statistics clinic. Other people call it a consulting program, a term I detested from the beginning, much preferring "joint research." I spent a lot of time in the early years here on that kind of activity. To remind myself about the various projects, some of which died early, and some of which stayed on for many years (for example, long-term work with meteorologists on weather modification), I kept a notebook. Later, I stopped using the notebook and this particular role of mine tended to decrease as I got into other things, but it was a fine thing to do. Fun and frustrating at the same time. Others took over the function as time went on.

Zabell: It must have been time-consuming.

Kruskal: Time-consuming indeed. And it reinforced a theme already outlined, notably striving for interpretability of quantities whether in population form (that is, functionals of the unknown distribution) or in sample analog form.

MEASURES OF ASSOCIATION

Zabell: At this time, you also started your research with Leo Goodman?

Kruskal: Well, the work with Leo on measures of association started with a conversation at a New Year's Eve Quadrangle Club party. Leo had had, as I remember, some discussions with L. L. Thurstone that in effect were about measures of association. I had had some separate discussions with the late Barney Berelson, a senior faculty member in the Graduate Library School. Leo and I were rather scornful of what you found in the textbooks about measures of association for cross-classifications, so we tried to see what we could construct that was interpretable. (One of the measures, it turned out, had already been proposed by Louis Guttman. We didn't learn that until somewhat later, fortunately in time to get it into the published paper.) We tried to lay out some of the considerations that might affect choice of such a measure, including the view that you don't have to have a single one. It's a little like what I call the "who's your best friend problem." Children—maybe adults, too—often get worried by the feeling that they have to have a best friend, and when I was a child it hit me that there was no reason why I should have a single best friend. You have good friends. But that was not part of the childhood culture then and I suppose not now. So, there's no reason why you should pick a single factotum measure of association, no reason why you should only have one kind of screwdriver in your home repair tool kit etc. But then that point kept recurring. People kept asking, "Well, what is the best one? What do you really think?" I suppose there's a limit; you can't have thousands of them, or you'll go crazy trying to make choices.

Zabell: How did you and Leo interact?

Kruskal: Oh, we exchanged draft statements, we talked on the phone and in person. We got after this epistemological issue of interpretability. It was well hashed out between us. Then we got into relevant approximate sampling theory and tried to write it up in an accessible way. I remember that while we were doing this, Leo spent a year in England at Cambridge and I and my family went to England for the summer to visit my sister Rosaly. We rented a house with my sister and her family on the Sussex seacoast thinking to go swimming, but the Sussex seacoast that summer was nothing like the seacoasts I remembered in the New York area. I think we went swimming once. Anyway, we visited Leo, moving from one draft to the next there in England. That was great.

Of course Leo later moved on to linear, log-linear methods and I didn't go with him. He created a whole new field of statistical analysis.

Zabell: What did you turn to?

Kruskal: I moved in a slightly different direction. I'm proud of the work I did on ordinal association. That came out of a course given at Berkeley when I spent the year there.

BERKELEY

Zabell: How did you come to visit Berkeley?

Kruskal: Well, I had met Neyman during the war, we saw each other at meetings and so on. He invited me. It was a regular teaching visit; very busy, very stimulating, a different kind of department from Chicago's. For example, the folks at Berkeley wanted to introduce a new elementary statistics course, and it was taking them something like 18 months or 2 years to get approval. It took me a while to understand what was going on because back in Chicago, if we wanted to introduce a new course, we'd just do it. The only red tape was being sure that the registrar knew in time to make up the requisite forms. So that was a totally different administrative atmosphere, but of course Berkeley was, and is, an exciting place to be.

I remember getting involved—again in the joint research sense—with a chap in the Department of Soil Sciences, a field I had never heard of, by that name anyway. That was good fun. I think his name was Bodman.

One of the courses I gave there was on applied nonparametric analysis, and in talking about rank correlation I realized that there really was a need for clarifying interpretations. I used Kendall's book a lot, of course. It is a fine book in many ways, but it didn't go into this interpretation question enough, so that's what led to the ordinal association paper [*Journal of the American Statistical Association* **53** (1958), 814–866].

Zabell: And didn't you also teach a course on chi-square?

Kruskal: Yes, and that's an example of something I regret never having followed up. It was fun trying to get chi-square-like things down systematically. (Of course the term "chi-square" is shorthand; it's much broader than that.) Fortunately, a few years later Shelby Haberman picked up that theme. The great thing there for me was to see two interests put together—namely, chi-square methods and coordinate-free linear analysis methods, whose intimacy was plain but not truly written up by anybody to my knowledge until Shelby came along and did it. His dissertation was marvelous.

FIG. 5. *William Kruskal with Jerzy Neyman, Berkeley,* 1971.

COORDINATE-FREE METHODS

Zabell: How did you get interested in coordinate-free methods?

Kruskal: Via Jimmie Savage, very clearly, who had in turn gotten interested I believe from von Neumann at the Institute for Advanced Study in Princeton. Once Jimmie said a few magic words, it all became plain, but it needed writing up. So both of these topics, again, are examples of things I regret not having done or not having finished. As you probably know, I worked for years on a book essentially called "Coordinate-free approach to linear this and that." It never was completed and that's a matter of deep regret. It's probably too late now. The mathematics of it goes way back, and there have been a number of books on it, including books that go further, mainly to truly multivariate methods from this point of view. My book, my potential book, my unborn book was the basis for a sequence of courses I gave here for quite a few years, and in it I put a lot of emphasis on interpretation of population functionals. For example, interaction, which is too often just treated in a purely formal sense. Some of my papers are based on this point of view.

Zabell: I remember your 1968 paper in *The Annals of Mathematical Statistics*, "When are Gauss–Markov and least squares estimators identical...?"

Kruskal: That started out as an exercise, an exam exercise in the course I was giving, and then Geoff Watson came along with much the same material; he encouraged me to try for publication. I note another paper, the 1986 one in *The Journal of the Royal Statistical Society, Series B*, "The geometry of generalized inverses." I call that one my string quartet paper.

Zabell: Why is that?

Kruskal: Because in it I use little balloon-like sketches to indicate the hyperflats and linear manifolds, and I have a footnote about how this is an aid to the mind like tapping your foot while playing in a string quartet so that it can't be seen by the audience, but nonetheless is helpful.

OTHER UNCOMPLETED WORK

Zabell: It's obvious you're not shy of publishing, so why were you reluctant to publish your book?

Kruskal: I was an overperfectionist. That was the problem. I opened up negotiations. There were several publishers who came to me. They heard about it. There are other topics that I would have liked to have gone on with, but I doubt that I ever will. Let me mention one or two.

First, I did some work with Don Campbell, social-psychologist (once at Chicago, then at Northwestern, now at Lehigh) on seating aggregation. He was interested in social measurements that were not intrusive. In the seating study, teachers of classes kept track of whether seats were occupied by white or black students, with interest in seating aggregation and segregation. Or you could look at boys or girls. All the teacher did was make a few checks on a form. Don, for whatever reason, looked at the number of pairs of adjacent seats which had a boy and a girl (for example) in them. The question came up: "What do you do about empty seats?" Another question was: "How do you compare classes of different size that have different proportions of boys and girls or blacks and whites?" If you take a really null–null hypothesis approach (which is all we could handle), the model is to put idealized black and white balls in an idealized urn; you mix and draw, then put the balls in occupied seats. That is, if these two seats are occupied and not the others, you restrict yourself to hypothetical repetitions in which these two seats are occupied. You look at how many are black/white or white/black, or same color. That's a combinatorial problem that I later found had been worked on heavily, but it leads to an expectation and variance under the null–null hypothesis and that gives you one way of getting a yardstick if you want to be able to compare seating arrangements. You can standardize both of them by null–null hypothesis expectation and a null–null hypothesis standard deviation. That was my primary contribution. But it was worrisome because obviously people do not seat themselves at random, and I tried to think of some more realistic approach, some model with parameters that you could estimate and that are meaningful social-psychologically. I come back to that problem once or twice a year, but I never got anywhere with it. Maybe it would attract you. So that's another totally different example of something I'm sorry I never did. The paper (by Campbell, William P. Wallace and me) appeared in *Sociometry*, 1966, plus a correction in 1967.

Second example. I've always been interested in history, and there were a lot of history projects that I would have liked to do, but didn't. (I talk as if there was nothing to do for the future, which I don't feel inwardly, but obviously the future's limited.) For example, I've been quite interested in the Dreyfus affair and its statistical aspects. There's a wonderful quotation, or alleged quotation, from Painlevé, one of the great mathematical minds of the period and also a statesman; he was twice prime minister of France. Anyway, I tried to track the quotation truly down, but never did. The issue was this: suppose you find some deviation in Dreyfus' writings, or anybody's writings, from a population average in terms of frequency of using letters or words. The actual one that was of special interest was doubled letters, words with doubled letters. You could then indict any writer—Molière, Racine etc.—because every writer has special deviations from the grand average. This seemed to me a fine example, if I could only get the documented quotation, but I never did. David Hackett Fischer, in his splendid book *Historical Fallacies*, gives a tertiary citation on page 118, but it leads to dead ends.

HISTORICAL INTERESTS

Zabell: This might be a reasonable place to discuss other of your historical interests. The earliest paper on the history of statistics of yours that I know is "Historical notes on the Wilcoxon unpaired two-sample test" (*The Journal of the American Statistical Association*, 1959).

Kruskal: In fact, my very first paper (*American Mathematical Monthly*, 1946) was on the Helmert distribution; you might think of it as historical. Back at the Naval Proving Ground during the war I was educating myself in statistics, and one of the first things I got to was the Helmert distribution and various derivations. I was surprised that I couldn't find one that was inductive; it seemed to me it ought to be easy and it was easy, and that led to my first paper. There was some history in the joint paper with Allen Wallis about analysis of variance by ranks. As you know, once you get caught up by history it's pervasive. The historical paper on the Wilcoxon test is closely related to the H-test paper. The measures of association work with Leo Goodman includes a good chunk of history. We kept running into the following phenomenon. Somehow we'd get a lead to an article on measures of association, let us say in the geological literature. That would open up a whole new bubble of titles. I'm sure there are big literatures we totally missed.

Zabell: Have you found as time goes on that you have increasing contacts with professional historians of science?

Kruskal: Not nearly as much as I would have liked. Some of them came about relatively recently via my paper on miracles, my Presidential paper for the American Statistical Association (*The Journal of the American Statistical Association*, 1988). The miracles paper deals with stochastic dependence and the dangers of blithely assuming independence in contexts such as the witnessing of unusual events.

Zabell: An example of a partly historical paper is your long review (*The Journal of the American Statistical Association*, 1980) of Joan Fisher Box's biography of her father.

Kruskal: Yes, that was fun to do. Steve Stigler invited me to review the biography of R. A. Fisher by his daughter Joan, and I found it a fascinating activity. It turned out to be quite a long paper that raised questions such as: "Do we have a right to be interested in the biographical details of a great scientist's life, or should we restrict ourselves to the work?" That, of course, is a big continuing issue among historians of science. There are at this university people like Noel Swerdlow who feel strongly that you should just look at the work, and others like Allen Debus who feel quite the opposite: that it's important to look at the setting. I tip in that direction, although I see the danger of getting too far away from the work.

R. A. FISHER

Zabell: Your review seems to sum up a long period of interest in Fisher.

Kruskal: Certainly, teaching linear hypotheses materials inevitably brought me back time and time again to Fisher. He had a wonderful mind, but one can get so angry with him.

Zabell: Would you like to say something about how Fisher was regarded back then in the 50's when you came to Chicago and first became connected with academic statistics? That was certainly a period of considerable controversy.

Kruskal: Well, another great man, Jimmie Savage, expounded often and at length on what a great man Fisher was and why. Fisher paid several visits to this campus; we gave him an honorary degree on one of these visits. As I say in the review of his daughter's book, on one of these visits Jimmie and I plotted a little to get R. A. F. off into a corner and really pound him politely about attending to alternative distributions in hypothesis-like testing situations. We weren't going to use the word "power" because that would be a red flag. "Power" had been introduced by his antagonist, Jerzy Neyman. We got Fisher off in a corner and kept saying: "How can you choose the test statistic without having some idea about alternative distributions?" In the end he had to say: "Well of course, how stupid do you think I can be?" But I don't think that made any difference in his other public utterances.

Fisher was stimulating to talk to, but just as irritating to talk to as to read, in my humble view. And of course *he* got irritated. He spent a year, or what I think was supposed to be a year, at Michigan State, and they kept asking him in his lectures to give definitions: "Would you please define 'sufficiency' Professor Fisher?" His mode of thinking was just the opposite of what was so mathematically fashionable; he liked to do things in terms of examples. He finally just left Michigan State early; there was so much tension.

Zabell: What did people in the profession make of these disputes, say between Fisher and Neyman, in the 50's?

Kruskal: I don't think there have been major secular changes. There are now more attempts to try to tease out exactly what the issues were and weren't, and we've got access to papers and correspondence that we didn't have back then. On the other hand, back then you had the people. The whole Fisher–Neyman controversy, especially in its later years, became ridiculous and repetitive. Most unfortunate, but they were human beings.

BEING A DEAN

Zabell: When I first came here to Chicago in 1974, you had just become a dean.

Kruskal: Of the Division of the Social Sciences, which is not the division in which the Department of Statistics is. That's the Division of the Physical Sciences. Isn't that a little anomalous? I was asked to be the dean of Social Sciences by Edward Levi, at that time president of the university, and one reason he asked me was that he liked to have deans of divisions whose home bases were in other divisions. He'd say it was a little paradoxical and he enjoyed that. Secondly, and more pragmatically, a common problem is that a dean comes out of the Department of X, so there's a possible conflict of interest. Is the dean going to lean over forwards and favor his home department, or lean over backwards and disfavor it? Whereas, in my case there was less of that because my own department was in a different division of the university. Of course there were intellectual patterns. For example, I wasn't very close to anthropology and I was quite close to sociology, but some-such is inevitable.

I found the deaning extremely interesting for statistical reasons. It seems to me that the most important part of the job is making recommendations about appointments, promotions and so on, because that's where the future of the institution lies, and it's inherently very difficult, even in your own field, to come to a reasoned evaluation of a scholar. To try to do it for one person where you knew the area would really take weeks, easily months. To try to do it when you don't know the area is essentially impossible. Now, this is an old problem that, for example, Walter Lippmann talks about in his book, *The Phantom Public*, from the viewpoint of citizens who want to be responsible in voting. Yet how can they do it? They've got distracting families and jobs. Lippmann, himself, as a full-time public commentator could pay serious attention only to four or five topics at a time. How much more difficult for a steel worker in Gary to consider any national topics. Lippmann had some suggestions for the problem, but I don't think they really were solutions. Still it seemed to me a challenge and fascinating to study department views and letters of so-called evaluation, and to try to come up with reasonable conclusions. Then and now the dean's recommendations are typically, not always, but typically, adopted. If they weren't, the motivation to do the job carefully would decrease sharply. Part of the problem is that many letters of evaluation from the outside, which are an important part of the process, are not truly informative. They say things like "brilliant thinker," without giving examples of brilliance.

When I was chairman of the Department of Statistics, earlier, I was in effect writing recommendations to my dean and I tried to explain people's contributions in understandable but not condescending terms. I couldn't do it every time, but I did it a fair fraction of the time. And in the social sciences, there were traditions in some departments of what seemed to me rather sloppy compilations and evaluations. You might have fifteen letters and you knew no more after reading them than that people were willing to put vague words down on paper. I think the norms and cultures here differ widely among universities, and among departments of a single university. I tried to push in what seemed to me stringent directions; I hope with some success. It may be parochial to say it, but the fundamental questions are statistical. You're dealing with a sample of opinions and often ill-expressed opinions. You need to put it all together in less than a hundred years.

Zabell: Do you find yourself looking back now upon various decisions and having a clearer sense of how the process worked?

Kruskal: Of course it's a selective situation. The people that didn't stay here, or that you lost track of, it's hard to know about. So in that sense, it's highly selective. A few people hate me as a result of evaluations that didn't go their way. I had never really experienced it before, but as a result of negative actions taken, I was cut, given the cold shoulder, not greeted when I ran across a person. Some of these dislikes go on for a long time, ten or fifteen years. It doesn't really bother me, but it's a phenomenon that deserves study.

Zabell: You apparently didn't find such issues arising when you were department chairman.

Kruskal: Well, I think we were careful in the department to try to avoid personal antagonisms, and actually that's one of the great things about the Department of Statistics here. I hope you agree. By and large there have been pleasant constructive relationships among its members. The only major exception was when Jimmie Savage left in a complex huff.

STATISTICS: ITS ORGANIZATIONS AND IMAGE

Zabell: Besides being a chairman and a dean, you have also been president of the Institute of Mathematical Statistics (IMS) and the American Statistical Association (ASA). Do you want to say a little bit about either of those organizations?

Kruskal: Well, they are quite different, of course. The ASA has a big staff, and bureaucracy, and a lot of momentum. IMS was, and is, much smaller, less formal and it's international. I thought that one responsibility of the president of IMS was to preserve its international character because it was easy to neglect that since most of its members were American. I think that's been meliorated since my presidency. At ASA, of course, the chief executive day-to-day person really controls a lot of things. That is Barbara Bailar now; it was Fred Leone during my presidency. ASA is avowedly a nonelite organization and I think that's important. It should be a nonelite organization and it is. But it sometimes does strange things; for example, I think that ASA is oversectionalized.

Zabell: Were there any special projects that you had in mind when you became president of the ASA?

Kruskal: I might mention the possibility of getting a building of our own. I was much in favor of that and tried to move it along. It didn't actually come into being until after my term of office. What I had been hoping for was a building right there on Pennsylvania Avenue with flags flying, because statistics is usually such a misunderstood and kicked-upon field.

I must've told you the story about running into a former student at O'Hare. He flooded me with flattery about the wonderful statistics course he remembered, and the thing he remembered most clearly was something utterly wrong. What does one say?

Zabell: Do you think public attitudes towards statistics have changed since the 50's?

Kruskal: No, I don't think so. There are certainly more people in both absolute and relative terms who understand something about statistics, but ubiquity of misunderstanding continues. In some ways it's our own fault. Some of the problems arise because of total misunderstandings about hypothesis testing. The use of the term "statistically significant," for example, is often just pyrotechnic show-off. Is it worth fighting lifetime battles over these things? Especially when perfectly good words like "disinterested" get misused in the best family newspapers? I don't know.

Zabell: I know you had an interest in usage of statistics in the journal *Science*, and the quality of its statistical refereeing.

Kruskal: Indeed, yes. We had a good thing going with a person on the staff, a senior editor, but then she retired. I think Section U, which is the statistics section of AAAS, has tried to be influential. Joe Cameron worked part-time as statistics editor. I don't know what's happened to that arrangement. But it's a good example. The use/misuse of hypothesis testing in the journal *Science* is, I think, notorious. One misuse is to do a significance procedure, get a *P*-value, and then call the result the probability of the null hypothesis being true with a semi-demi-Bayesian–flavor interpretation. This is an old story; there's a wonderful paper by Fred Mosteller pointing out the difficulty, and it's in my own *Encyclopedia* paper on significance testing.

Zabell: Statistics does seem to have an image problem.

Kruskal: My Vancouver talk on statistics as a profession discusses statistics as it tries to present a better view of itself to the world. Other professions face the same difficulty, but statistics has practically no songs and stories in which the central characters are heroic statisticians. The one good example I have is Ford Madox Ford's novel *Some Do Not....* One important plot and character development turns in large part on a statistical problem: What to do if your boss tells you to distort the data?

Zabell: Yes, a very topical concern.

Kruskal: It takes place in WW I times. Of the two statisticians described, one of them refused to distort and the other distorted. He got knighted, or whatever, plus political points and maybe money, too. I know a few other such literary examples, but Ford's is the most clear-cut because it deals with a real statistical problem in a worldly sense. I'm sure it happens all the time: you might call it "statistical harassment."

Zabell: That problem also arose during the Viet Nam war.

Kruskal: Oh, that's another example, a beautiful and horrible example. But it's not in a novel. Who did it? Which human beings succumbed to pressure and distorted data? I don't know. It had several aspects including body counts, "pacified" hamlets, counts of our own troops.

THE CENSUS

Zabell: This might be a good place to discuss your interests in the census and related matters.

Kruskal: Fine. Let's talk about the decennial census of housing and population. My first formal connection had to do with the 1970 census. A National Academy of Sciences group was established to look at undercount problems. Richard Savage, Norman Bradburn and I were members, along with other colleagues. The problems of understanding counting error within the census, or any kind of error in the census, are extremely difficult and extremely important. We made only a start. Then in 1980 and 1990, the whole thing came up again. The world was waiting for somebody with a smashing new idea to solve the problem. I doubt that such a hero will appear. So many minds have been thinking about it, and, as far as I know, not coming up with anything wonderful. One thinkable radical change might be to make census counting your ticket of entry into a grand national lottery, so that a few people get millions of dollars and more people get little prizes. But you have to be counted by the census. I'm afraid that would not seem tasteful to a lot of citizens, but I don't know.

Zabell: In 1970 though, you're saying that the accuracy of the count really was not as contentious an issue as it is today?

Kruskal: That's right, but maybe that's just the temporal discounting of memory. The topics have stayed the same really, except that it's only in 1980 and then again in a bigger way in 1990 that specific methods of estimating the population for small areas have been extensively and seriously proposed—namely, bias and variance of counting via the capture/recapture approach which has been widely used for wildlife populations. There are problems there, too, of course, but in some cases you can truly tag birds or fish caught in the first netting with physical tags. You can't very well do that with people, thank goodness. For census purposes, capture means counted by the census; recapture means counted by a postcensus resampling.

Zabell: So in 1970, it was mostly an in-house type of affair where people discussed the problem, but no one really advanced a particular solution at length.

Kruskal: The census was advised to work harder, and to pay more attention to basic social science. For example, one approach that I think may lead somewhere if only given a chance is called the ethnographic or participant observer approach. The observer, a trained anthropologist or ethnographer, goes to live in an area or spend a lot of time in an area and gets to know a few households very well, so you really understand who's in there, and then compare those results with the census. The whole trouble with the census problem is how do you calibrate? What is an accurate way of checking it out, especially when you're talking about segments of the population who may be hostile to government activities, segments like criminals, undocumented immigrants etc.? In addition, most statisticians at the census were mathematically oriented, middle-class good boys and girls whose approach to American culture might be limited (like my own). Should there not be some way of broadening that point of view and certainly getting further into basic social science? Well, the ethnography approach was first tried to my knowledge around the 1970 census. Charles and Betty Lou Valentine, ethnographers, went and lived in a low-income, primarily black, part of Brooklyn. But then they became so upset about what they regarded as official municipal mistreatment of these low-income folks that they became political activists. They said: "We obviously can't continue this scientific study and at the same time be activists, so we resign." But they published fascinating materials. Some of this continued in 1980 and 1990, but the samples were still small. In 1990 I think there were some 30 or 40. Further, these were not chosen by any structured sampling method, and we still (October 1991) don't have the results.

So, it's too late to use them, even if we wanted to, in making some kind of change in the 1990 small-area estimates. Part of the problem is that you need millions and millions of estimates and that's not a typical statistical problem at all. Some people, like our friend Bruce Spencer, think that it's enough to use a loss function to combine areas, yet others of us think that any choice of the loss function is arbitrary and the results are bound to be strongly dependent on the choice. We worry about that.

One of the arguments for adjustment is that in most cases adjustment, by most methods, increases the estimated counts. A few of them decrease, but most of them increase. But in my opinion, that's really irrelevant because what what you should be looking at, if you could do it, are not the absolute increases, but the relative increases. Looking at the increases of two geographical units, both might go up, but the relative order between them might switch around very easily; in my report I gave a specific example of that. Of course, what a lot of people say is that, in effect, a rising tide lifts all ships; but that's not true if some of the ships have holes in their hulls.

My opinion is that we were far from having evidence that the undercount would affect the process in the kind of radical way that was proposed by the plaintiffs. There was a lot of material in the papers about how the big cities were going to lose not only money but votes because of the net undercount. I think a great deal of that was exaggerated because had the adjustment method been used, while the plaintiff big cities would come up with higher numbers in general, so would other big cities, and the question is of the net increase. Most uses of census data for distribution are in terms of a fixed pot of something to be distributed. It's not at all clear to me that it's like saying you lose \$150 for every Chicago citizen not counted. There are a few exceptions. For example, in some distribution legislation there's a threshold and the city comes under the rubric of the distribution if it has at least a million people, or some such. So in those cases the absolute numbers do count, but most distribution cases are relative.

Zabell: I know that there is census-related litigation. Have you been involved in litigation or testimony before congressional committees?

Kruskal: Well, in 1980 I gave a deposition for the court case that was then coming up. New York then, as now, is the lead litigant among the plaintiffs. This time, Judge McLaughlin of Brooklyn ordered the litigants to hold everything, to recess the trial. A panel of eight would be formed, of whom four would be tilted toward a so-called adjustment procedure. (I don't like the word "adjustment" because to use it almost implies that there is an accurate way of adjusting, and I don't see any reason to think that.) Then four would tilt in the other direction. The way this was done was that the City of New York named six people who resonated to the New York theme, and from these the Secretary of Commerce picked four. Then the Secretary of Commerce picked four other people, including me, presumably to take the other point of view. These eight panelists all were to make individual recommendations.

In the quartet that included me, the person I talked to most was Ken Wachter (Berkeley). The two others had backgrounds in public opinion measurement and political consulting, namely, Lance Tarrance, who runs a public opinion company centered in Texas, and Mike McGehee. The other quartet was chaired by Gene Ericksen, a very vigorous person. It included John Tukey, whom I've known and admired for many years. The other two were Kirk Wolter, who had recently left the Bureau of the Census to go to

A. C. Nielsen, and Leo Estrada, an academic demographer from California. The ground rules pointed to individual reports, but the four people in the other quartet issued a joint statement and then separate individual ones. This was against a time limit of mid-June 1991. The Secretary of Commerce, Mr. Mosbacher, had to make *the* official decision as of the middle of July. The eight individual recommendations, plus the fourfold joint one, are public documents. One of the things we're seeking now is publication together in some more or less conventional form for future reference. My report was brief and nontechnical. Ken Wachter wrote a wonderful report. Tukey and Ericksen wrote very interesting ones. I don't fully understand them nor do I agree with them.

This formal two-quartet structure for presenting professional opinion was a novelty, certainly new to me, and by itself of considerable interest. Maybe it's a good way of handling statistical issues, but it raises all sorts of questions. For example, in both 1980 and 1990 some people at the bureau were peeved because they felt that the decision about whether to have modifications should properly be left to the statisticians at the bureau. "It's a professional judgment," they claim. That leads to other general issues: To what extent should statistical policy matters be decided by statisticians; to what extent should matters of war and peace be decided by generals? (The French statesman Clemenceau said that war is too important to be left to the generals.) Coming back to the census, statisticians and other scientists disagree about whether there should be some kind of modification and, if so, what. I easily drew up a list of twelve or fifteen prominent academic statisticians who felt that modifications were impossible to do properly, and shouldn't be attempted. Nonetheless, colleagues like Steve Fienberg publish papers saying that the vast majority of statisticians agree that there should be an adjustment. I don't think that is true, yet it's not easy to investigate because presumably you only want views from those who are well informed.

Anyway, this census experience is still continuing. I've been subpoenaed by the major plaintiff, that is, the City of New York, and I'm being helped by the Department of Justice, in its capacity as counsel to the Secretary of Commerce. So I've been on the phone often in the last few days trying to make physical arrangements. One of the problems is that I have about six boxes full of papers that came to me as a result of this activity, many of which I haven't read. The panel members have been flooded with papers, and the plaintiff wants to come, look at them all and then question me—as they say, to depose me. At first, I was asked to bring all the papers downtown. Quite a job. The current plan is that the lawyer for New York City will come here on Monday. We will provide an office so she can look at these papers for perhaps a day; then she will question me and I will have as a shield and advisor a lawyer from Justice.

Zabell: Have you been involved in many depositions before?

Kruskal: Just one about ten years ago, and also a census matter. At that time there was no problem about looking at papers, nor was there a panel.

Zabell: What types of questions were raised at that deposition?

Kruskal: I don't remember. I suppose they asked why I wasn't delighted at the compassionate chance to help these poor people, disadvantaged people; that's one standard attack theme. I have a statement on what I call the argument from compassion which gives my views about why that's a poor argument in this context.

Zabell: Ten years later, we are having another census adjustment debate, so it seems that nothing has been resolved.

Kruskal: It's a terribly difficult problem, and the more I think about it, the more I worry about adjustment in general. It arises everywhere. For example, here's another recent case. The government through the HCFA, the Health Care Financing Administration, prints a multivolume publication every year or two about hospitals. It gives mortality rates for various procedures by hospital. Much of the space in these multivolumes is filled with letters from hospitals saying, "Publishing these figures just by themselves is misleading. People will think that we do lousy surgery here. But the reason our mortality percentages are high is that we take the hard cases." So that's an adjustment request, and there's a lot of interest in methods of adjustment. Some think that it's automatic...that there's no problem, that it's a kindergarten statistical problem; you just follow standard methods. I think that's wrong and that really nobody knows how to choose among possible adjustment methods, and even how to think about it. People have indeed thought about it in particular contexts, so there's a whole field there to work on more seriously than in the past.

SEMANTICS AND STATISTICS

A term like "adjustment" is hard enough in its connotations, but some people insert an adjective like "correct," which has even stronger connotations. Of course, you can also use words like "manipulate," which has strong negative connotations. Indeed, semantic issues arise throughout statistics. R. A. Fisher was a master at choosing strong words for the concepts that he introduced or explored: "sufficient," "efficient," and so on.

Zabell: "Consistent."

Kruskal: "Consistent," yes; let's add it. Well, how much should we worry about such semantics? I have one paper on this theme, and Steve Stigler and I are currently worried about another aspect of it, the word "normal," as in "normal distribution" or "normal equations." "Normal" seems to us a tremendously powerful word. It has two broad senses: (i) something that's a goal and (ii) something that's commonly found. These are different meanings, but both of them are used, sometimes in confused ways. Thus, you have "normal" used in medicine: "normal patient," whatever that means, and "normal blood pressure." In statistics you have "normal distribution," a name apparently given to the Gaussian distribution by several people, including Galton, at about the same time. Then there are the "normal equations," which were christened that by Gauss. What isn't clear is why Gauss called them "normal equations." The obvious guess is that Gauss knew very well the geometrical interpretation of solving normal equations. One way of saying it is that you drop a normal from a vector to a hyperplane, so therefore "normal equations." But Gauss doesn't actually say that, although in the case of other terminology he tries to give some explanation. So that's the big gap in our study so far. We also consider normal schools, which started out as a model for children's schooling. But almost at the very beginning (in Vienna at the time of Maria Theresa), teachers—teachers in training—came to these schools, and gradually the term "normal school" came to mean schools for teachers. There's a good example of grown-up powerful teachers taking over a good word from kiddies. So we see examples of semantics in statistics. Looking back, that has been for me a continuing thread of interest.

GOVERNMENT STATISTICS

Zabell: In addition to your interest in the census undercount, you've had a long-time interest and involvement in government statistics. Maybe you could say something about how you initially got interested in that subject.

Kruskal: Well, I've always been interested in it, but it became more focused during the Nixon era when Allen Wallis was asked by President Nixon to form and lead a commission on Federal Statistics. The commission was to study accuracy, organization, timeliness. Fred Mosteller was the vice-chairman, and John Tukey was a member; Stanley Lebergott was also; I was and others. We had a year's life, which was precious little because we had to gear up, get a staff, and then there had to be time at the end for publication. We published a two-volume report that on the whole was reasonably good. One of its recommendations was that, just because the life of this commission was so short compared to what you needed to study issues properly, there should be a more nearly permanent group formed. That turned out to be the Committee on National Statistics, formed as part of the National Academy of Sciences–National Research Council. It was satisfying to play a role in getting this committee organized and off the ground. We fortunately recruited Margaret E. Martin, who had been with the statistics group at Office of Management and Budget to be our chief administrator, help with money raising, get projects going. So I chaired the committee for six years and then others took over. Steve Fienberg did it for quite some time. Burt Singer is now the chairman. The executive director is Miron Straf. (This activity has been emulated in another part of the National Academy, in the mathematics section.) The Committee on National Statistics has turned out many reports on interesting and important topics. For example, here's a topic: responsibility for data. What, for instance, should a scientist do when there's just too much data to publish? What is responsibility for documentation of data, and so on? Another big project was on so-called missing data, a pervasive difficulty.

Zabell: Is it like missing persons in census counting?

Kruskal: Well, yes, in a sense. The census problem is so incredibly huge. It's not just the count, it's all the characteristics, and fundamental ambiguities about concepts: what's a household, what does it mean to be a member of a household? Suppose you had the most cooperative respondents in the world. Are there not operative differences in point of view here by ethnic group, economic group and so on? Suppose you've got kids, adolescent kids, who have affiliations with two or three households, uncles and aunts, grannies and grandpas, who sleep on various nights at various of these households, leave clothes at all of them, eat at all of them. I'm thinking of extended family arrangements of different kinds. You come to a member of one of these households and ask how many people are in it and who are they? If you came to an Americanized adolescent in an Albanian household, you might get totally different results than if you had come to that person's parent. So, that's an example of a statistical question in a sense, but it's just as much political, social and so on, and there ought to be more relevant work.

Zabell: Have you started any new projects recently in connection with government statistics?

Kruskal: You might find interesting a visiting activity I had last year. I accepted an invitation to establish a relationship with GAO, the General Accounting Office. I was a visiting scholar, which meant going to Washington a couple of days a month for a year, talking to their staff and finding out what they're up to. One of their problems was a traffic issue: are small cars really more dangerous than big cars? This is much in the news now. The evidence at GAO—difficult to interpret—was that it's just not so, or at least not so monotonically. Damage goes both up and down as a function of weight. Terribly hard to know how to set it up because in actual accidents you've got weights of both cars. Then there are accidents that might've happened and didn't. Fascinating, but very difficult, almost impossible I think, other than via laboratory-like testing. But then in laboratory-like testing you don't have all the variables—the real variables.

Eleanor Chelimsky at GAO is a wonderful person. She's in charge of the program evaluation and methodology division, which she has built from nothing to an organized large and able staff. She takes statistical problems seriously. I think that GAO is an important place that will become more important.

STANFORD

Zabell: Talking about visits reminds me to ask about your year at the Center for Advanced Study in the Behavioral Sciences.

Kruskal: Yes, 1970–71 was a marvelous year. I suppose everybody enjoys being at the Center. Richard Savage was there the same time I was and that was great. I had a chance to interact with Charles Gillespie, the historian from Princeton, and that was fun. He was working on Laplace. I didn't get as much done as I might have because I was that year President of IMS . Also, I was finishing off a census project—the first census project I was on—and I was starting on the miracles work with both Berkeley and Stanford libraries. Oh that was good fun! I don't know how other people decide what they're going to do on a given day, month, year. I must say I don't do it with much ratiocination. How do you do it?

Zabell: I think it's most profitable to work on what seems most interesting at the moment, but suddenly you need to be much more disciplined.

Kruskal: Yet that affects what's going to be interesting next week somehow. Sometimes I get started on something and I just can't continue, so I give up. Let me give you an example of a simple problem on which I gave up years ago. Consider bivariate distributions, with customary smoothness conditions. So there's a regression curve, expected value of y given x, and another, the expected value of x given y. Plot these. Question: Do they intersect? I worked and worked on that problem and couldn't get anywhere. The problem arose when I was interested in something else and wanted to translate the origin so that it lay at the intersection of these two lines, for convenience. Then I said to myself: "How do I know that they intersect?" I think the answer is "Yes, they do intersect, but you have to allow intersections at infinity. Will they asymptotically come together?" But I don't really know that, and it's such a simple problem to state.

I think simple problems to state have a special charm of their own. One of my favorites comes up in any elementary course. You're talking about histograms and you have some data that are open-ended. Incomes by thousand-dollar intervals, and then an open-ended interval—everything beyond, say, a hundred thousand—at the right. You could also be open-ended at both ends. How do you present this in histogrammatic form sensibly? Now, there's a truly elementary problem. It's not mathematical statistics, although you could run into some mathematics while fitting, which is I guess the usual way if anybody wants to do it carefully. Mostly people just omit the open-ended part; see, for example, in David Freedman's textbook. That might be the most important part. Those comments fall under "teaching and general attitude" if I may introduce that term.

TEACHING

Zabell: There are a lot of problems that come up in an introductory statistics course that are difficult to answer in a way that doesn't confuse students.

Kruskal: Yes; consider teaching about population characteristics. I've always been concerned with that subject from the first time I learned about mean and variance—first and second moments—as measures of something mysterious called location, and something equally mysterious called dispersion. I said to myself, "Why should I pick these particular fuctionals? They're simple, mathematically simple; that probably is the reason for it, but maybe I'm not interested in the sum of squared deviations, for instance, and there are alternative indicators of both location and dispersion. Do I have to choose among them?" Rather, I have to try to understand what they mean.

This is a point that comes up when you first teach mean and variance in an elementary statistics course. Should you worry about the difficulty in interpreting them? I've had long discussions, for example, with Fred Mosteller who says: "No, don't try to interpret them in the first course; you'll just confuse the people. Give it to them straight from the shoulder," as he put it. That means define it and show how you calculate it for a sample, and don't worry too much about the interpretation. And I said: "Well, but is that honest?" and Fred would say: "Well, I don't know about honesty, but if you tried doing it, they will not follow you and they'll get bored." So, it's a pedagogical problem.

I guess it's been a thread in my publications to try to understand what population characteristics you really want to measure; how to measure them, and not to worry that there may be several things measuring the same vague idea like dispersion or association. One of my earliest papers (1953, *Biometrics*) was on this line, called "On the uniqueness of the line of organic correlation." It really is the following: In sample terms, suppose you have a scatter diagram of points looking like a cigar or a long thin football, and you want to fit a straight line to this swell of points in some sense. What's the best way of doing it? And again that has a population analog, and the line of organic correlation is simply the line that goes through the mean and has slopes corresponding in absolute value to the ratios of standard deviations. You're usually interested in more than a two-dimensional version of it, of course, and that has nice properties and leads to puzzles about sampling.

ENCYCLOPEDIAS AND OTHER COMPENDIA

Zabell: Did you want to touch on the *International Encyclopedia of Statistics*?

Kruskal: That's related to teaching. It's another activity I got into via Allen Wallis. Well, there was an *Encyclopaedia of the Social Sciences* published in the 30's, and various people felt that there should be an updating of that. One of those was the publisher Jeremiah Kaplan, who created and ran the Free Press, originally out of a mailbox on the North Shore here, but then with Macmillan. A group was formed to look into the possibility of putting out a new *Encyclopedia of the Social Sciences*. Mosteller introduced the dirt index as a measure of how much people had used the original edition in libraries. You looked at the edge and saw where there were strata of dirt, and those were the articles that had been used more. A decision was made to get a completely new encyclopedia. The chief editor became David Sills, a sociologist connected with Columbia, and there were editors for various fields. Allen Wallis was Chairman of the Editorial Advisory Committee. I became the editor for statistics and I was enormously helped by Judith Tanur, who was associate editor and who lived near New York where the offices were. It was quite a project. After the encyclopedia came out in 1968, it was suggested that maybe the statistics part could be republished by itself with additions, corrections and so on. Judy and I did that. We generated quite a few new articles or completely rewritten articles and made corrections. I'm proud of the 1978 result.

Zabell: The encyclopedia was obviously a several year project.

Kruskal: Oh yes. Well, the actual *International Encyclopedia of the Social Sciences* came out in '68, and then the statistics encyclopedia came out in '78, ten years later.

There were, of course, all kinds of questions about the functions of encyclopedias: What does it mean? How can you organize a field into articles, topics and so on? Most were hard to answer.

Zabell: You were also one of the editors for *SAGTU* (*Statistics: A Guide to the Unknown*).

Kruskal: That's right. There was a lot of concern then, as there is now, about the teaching of statistics, in particular at the pre-college level. So a joint committee was formed between the American Statistical Association and the National Council of Teachers of Mathematics. Fred Mosteller was the initial chairman. The committee continues, and does great things. Most recently it got a substantial grant from the National Science Foundation, which our earlier committee had tried to do and failed. We did, however, obtain some welcome funds from the Sloan Foundation. Our committee talked long and hard about activities and we finally decided there should be two publications. One became known as *SAGTU*, *Statistics: A Guide to the Unknown*. Our thinking was that this was meant for parents, Board of Education people and so on; it was to publish essays about successful applications of statistics—accessible essays. Judy Tanur became editor of that, we all worked on it, and it's now in its third edition. I believe that the primary audience in the end turned out to be mainly students in undergraduate college courses for which people wanted motivating essays.

The other publication, *Statistics by Example*, consisted of short examples, intended to be cribbed by high school statistics textbook writers. So, far from wanting to copyright, cribbing was encouraged. Fred and his friends and students prepared examples. I did several of them. Two of mine were on Tom Paine and the rights of man, and I came to look at the age distribution in England in the late eighteenth century. Tom Paine made an egregious statistical mistake that was fun to untangle. Charles Land, a former student here, did a great piece on calibration of automobile speedometers. So that book came out. I don't think it's had a revised edition, despite some flurries of activity. Since then, I think most of the original members of the joint committee have gone to other pastures, but the group continues.

So those are some thoughts related to teaching. On the direct teaching front, of course, there are immense difficulties, including real statistical problems of how to analyze educational data—for example, comparing one method of teaching history against another, or international comparisons of mathematics teaching. Such questions have drawn a lot of interest, but how on earth do you make a start since the material that gets presented is different in different countries?

Zabell: What are the statistical topics that need to be addressed in the future, both in teaching and in research?

Kruskal: I hope that the future brings more attention first to interpretations of functionals on relevant distributions. For example, what are the really useful interpretations of first and second moments and of the (population) median? Second, I hope for more intense and broader studies of robustness against all manner of deviations from assumptions. Third, I hope that we will not always routinely use conventional concepts. The most egregious example of this is the automatic unquestioning use of simple hypothesis testing.

ACKNOWLEDGMENTS

Professors Zabell and Kruskal wish to thank several colleagues who helped in preparing this article: Leon J. Gleser, Leo A. Goodman, Joseph B. Kruskal, Martin D. Kruskal, Paul Meier and Judith M. Tanur.