Statistical Science
2003, Vol. 18, No. 3, 332–335
© Institute of Mathematical Statistics, 2003

John Tukey at Bell Labs

Colin Mallows

Abstract. John Tukey was a Bell Labs employee for 40 years. In that time, he influenced very many researchers and contributed significantly to the growth and luster of the Bell Labs Statistics Research Department. His counsel was valued by senior management, and his involvement in many applied problems led to important advances, including his work on spectral analysis. I will survey this and other contributions.

Key words and phrases: Appreciation, Princeton, spectra, EDA, FFT.

"MORE MATHEMATICAL"

I will start with a brief survey of the papers in the "More Mathematical" volume of *The Collected Works of John W. Tukey* (Mallows, 1990a). For a more detailed summary, I refer the reader to the introductory material that I prepared for that volume. The volume contains 42 papers that did not fit naturally into any of the other volumes of the series, each of which has a clear theme. I was able to classify some of the papers, so that there are six on mathematical topics (the title of John's 1939 Princeton Ph.D. thesis was "Denumerability in Topology," published in 1940 as "Convergence and Uniformity in Topology" in the *Annals of Mathematics Studies* of the Princeton University Press), four on the fiducial argument, five on transformation (or "re-expression," as he preferred to call it) and 27 papers on a miscellany of statistical topics. These include the original announcement of the jackknife method and three fundamental papers on tolerance regions. All in all, the volume contains what for most people would be a respectable life's work, but of course for John Tukey it is merely the "et cetera" of his output.

The fiducial section includes a 64-page (previously unpublished) examination of "the present state of fiducial probability," which documents John's attempts to clarify the logical basis of the method. He regarded Fisher's development as a "major intellectual accomplishment," but ended his manuscript without resolving the issues. I quote from John's 1990 commentary:

> Today I do not believe there is any chance for a successful single unifying approach to inference. . . . This does not disturb me—growing partial understanding is all that we have seen in such fields as theoretical physics—and the problems of inference are broader and harder to experiment on. . . . I have come to think [of] belief in a unified structure for inference as a dangerous form of hubris.

The papers on transformations span 30 years. In his commentary, John took issue with my comment that a 1979 paper "introduces Tukey's mature approach." I should have remembered that John's thinking was continually evolving; by 1990, he was focused on the use of g- and h-re-expression rather than the "guided re-expression" of 1979.

The three papers on statistical tolerance regions present successive extensions of Wald's basic idea, which John formulated as "Wald's principle":

> If W is a chance quantity, and ϕ a function such that each value of $\phi(W)$ has probability zero, then the conditional distribution of the elements $(w_1, \ldots, w_n)$ of a random sample of size n, given that $\max(\phi(w_i)) = a$, is that of one w with $\phi(w) = a$ and a sample of $n-1$ other w_i from the distribution restricted to $B = \{w|\phi(w) < a\}$.

I will not attempt to survey the remaining 24 papers here. I will take this opportunity to point out that several other documents were available but were not included in this volume; I recall particularly (because I spent a lot of time writing a commentary) a series of three substantial Princeton technical reports that, in

Colin Mallows is Member, Avaya Labs Research, Basking Ridge, New Jersey 07920 (e-mail: colinm@research.avayalabs.com).

the course of developing an approximate result, worked out the "statistical differentials" idea to fourth order. At the end of the third report, after perhaps a hundred pages of detailed computation, John recorded that he had started the work on Friday, and it was now Sunday evening. It would take a normal person at least that long to read the work.

BELL LABS

My main purpose here is to discuss John Tukey's work at Bell Labs. John Tukey joined Bell Labs in 1945 and retired from there 40 years later. His position was always something of an anomaly; one of the strengths of the Bell Labs style of management is that it can adapt to someone as unique as John by creating a position that suits his talents. At one time, John held two positions at Bell Labs; he was an MTS (member of technical staff, the regular career grade), reporting to a department head, who in turn reported to his director, who was John Tukey. Later his position was termed associate executive director, where he had no line responsibility. Many times while some presentation was being made, he would sit at the back of the room working on his "knitting" (the Citation Index), seeming to take no interest in the proceedings, until at the end he would ask some penetrating question that showed he really had been paying attention and, in fact, knew more about the subject than anyone else in the room (often including the speaker).

Bell Labs, as an industrial research laboratory, and Princeton, as a university, are organized differently and differ greatly in atmosphere. At Princeton, John was heavily involved with teaching and directing his research students, many of whom went on to develop his ideas in their own careers. At Bell Labs, his technical interactions were more through joint work and feedback to (and occasionally from) his colleagues. No doubt there were interactions with higher management, affecting our existence, but these were largely invisible to us at the MTS level. Many of us benefited from his comments, though it often took a lot of thinking to absorb them properly. He contributed to many activities at Bell Labs for which his name never appeared as an author. One specific relationship that was valuable both to John and to the profession at large was with Martin Wilk, who in the 1960s was the department head of the research group. Although they published only a couple of joint papers, their interactions were frequent, lively and productive, with Martin trying to bring John's formulations down to earth, making them more accessible.

His first breakthrough at Bell Labs (in 1948) was in the spectral analysis of time series. For decades, communications engineers had understood the relation between a function and its Fourier transform and had had analog devices whose output was an estimate of the spectrum. It was Tukey, however, working with R. B. Blackman and R. W. Hamming, who showed how to compute spectral estimates effectively using digital computers and put the field on a sound mathematical and statistical basis. The 1958 papers (Blackman and Tukey, 1958d, e) in the *Bell System Technical Journal*, reprinted by Dover in 1959, are difficult for an academic statistician, but contain much that would be transparent to a communications engineer. The papers lay out the basic theory of spectral (and cross-spectral) estimation, omitting only the "tapering" concept that John contributed in 1959. In subsequent years, John simplified the derivations and pointed out profound and illuminating analogies between spectral analysis and ANOVA.

In later work, he investigated cepstra—spectra of log spectra—which are useful in determining the presence of echoes. This was stimulated by his work on demonstrating the feasibility of the nuclear test ban treaty. The paper (Bogert, Healy and Tukey, 1963a) is full of exuberant neologisms, which play a useful role in keeping straight the analogies between spectral quantities (time, frequency, phase) and cepstral ones (frequency, quefrency, saphe). All through his life, John delighted in inventing new names for his new ideas; it is well known that he suggested the word "bit" as an abbreviation for binary digit, though Lancelot Hogben said in 1969 (Hogben and Cartwright, 1969) that this "has nothing but irresponsible vulgarity to recommend it." I once compiled a list of over 50 of John's neologisms; not all of these have found acceptance. Maurice Kendall, referring to "polykays," said in *The Advanced Theory of Statistics* (Kendall and Stuart, 1963) (a little more politely than Hogben), "we feel that there are limits to linguistic miscegenation which should not be exceeded." John would disagree. I don't know the reference, but I am sure he said on at least one occasion that it was one's duty to invent a new word for a new idea, since otherwise the novelty will be misunderstood and forgotten. Thus, his use of "batch" instead of "sample," and of "hinge," for which he proposed a precise definition, instead of "quartile," which does not have one (in finite batches).

As far as I know, John did not contribute to the theory of wavelets. He also played no role in the recent exciting developments in the analysis of self-similar

processes, such as arise in Internet traffic, where the classical theory does not apply because variances are not finite. What a great pity this is! How much farther ahead we would be if John were still around to explain to us the basic relationships in these fields!

John was always ready to comment on the work of his colleagues at Bell Labs. I have six pages of comments on two papers of Dave Thomson (Thomson, 1977a, b) who was studying some time series in which (as John remarked) the spectrum had a "horribly large" range. These papers led John to formulate several research questions, enough to keep Dave (and several others) busy for a long time. John commented, "Like R. A. Fisher's books, [reading these papers] once or twice annually might prove helpful." Dave was gratified merely to be mentioned in the same sentence as R. A. Fisher.

I remember the excitement attendant to the development of the fast Fourier transform (Cooley and Tukey, 1965a). John remembers (in a comment in Volume 2 of the *Collected Works*) that this was "a very busy time" for him, but he was not too busy to explain the idea to some of us at Bell Labs and to invite us to investigate how to implement the new algorithm and to explore how it might be used. Morven Gentleman and Gordon Sande (1966) wrote a delightful paper with the title "Fast Fourier Transforms—For Fun and Profit."

Another heady time was during John's development of techniques of exploratory data analysis. The research area of Bell Labs was a test bed for his new ideas; he would suggest new techniques, such as stem-and-leaf plots, or midmeans, and we would try them out and give him feedback. Since so many of these new techniques were obviously useful, the research departments embarked on an extensive in-house education program to spread them through the technical Bell Labs community. The "Bell Labs style" of data analysis, which is strongly data driven, is largely John's creation. Ideas that he emphasized became part of our style, for example, the importance of exploring data flexibly, with much attention to graphics.

SOME TESTIMONIALS

John's contributions to Bell Labs were not merely technical. I have several testimonials from senior colleagues at Bell Labs, written on the occasion of his retirement in 1985. Vic Vyssotsky, executive director of the area that included statistics research, said:

> What impresses me most about John Tukey is that range of issues on which he can resolve problems. I have observed this repeatedly through the years in both technical and management situation—John sees a solution where others see an impasse.

Arno Penzias, Nobel prize-winning vice president of research, said:

> Over the years, Bell Labs has had many outstanding scientists, a smaller number of great scientists, and a few great scientists who are also great people. John is one of these—a great scientist who is also a great man. Few have had a role equal to his in shaping the information age, and it's been great for Bell Laboratories and AT&T that he's been here. It has been a remarkable half-century of association—40 years, actually, but I'll call it half a century and let John deal with the roundoff error.

John's long, close association with William O. Baker, retired chairman of Bell Labs, began on a tennis court in Princeton in 1937. Baker said:

> John Tukey's thinking is so fine and fast that his friends and admirers are forever asking him to do it again. I have been one of those asking, and delighting in the results, for 48 years. For the 40 of those in which we have been closely associated at Bell Labs, John has had an incisive role in each major frontier of telecommunications science and technology: uses of transistors and the solid state; digital coding and computers...; statistical strategy for finding how speech energy is distributed in frequency (an essence of telephony, leading also to important concepts which he named: 'prewhitening', 'aliasing', 'tapering', 'cepstrum'); evolution of software and operations support systems; earth satellite and other microwave techniques; electronic switching; laser-based photonics; topology of integrated circuits; adaptation of behavioral and human factors science to telecommunication.
>
> Collaborating with John in these and a multitude of other missions, I have known his warm and heartening friendship and his unerring assessment of human abilities and temperaments. He has joined in conceiving and organizing most of the initiative in communications principles and science research undertaken at Bell Labs since 1955.

> These skills have enhanced not only his memorable part in the progress of Bell Labs, but have been applied in historic contributions to national security, environmental preservation, and public health. We have watched at least four Presidents of the United States listen to John and heed his counsel.

Another testimonial came from his long-time secretary at Bell Labs, Mary Bittrich.

> He was constantly challenging me to learn more, develop new skills. With his encouragement I've used capabilities I didn't know I had.

At John's urging, she became an expert in UNIX, and with his encouragement taught many people at AT&T and elsewhere how to use that system. In his introduction to the *Collected Works*, Fred Mosteller wrote that

> Tukey has been blessed with the long-time support of Mary Bittrich at AT&T Bell Laboratories.

Mary maintains that the last statement is backward.

Many people have testified to John's unfailing good humor and to the overwhelming effect of even his casual conversation. Perhaps I may contribute some personal remarks. I first met him in London in 1956, when I was a very new assistant professor at University college. We arranged to meet at Liverpool Street Station, which I had passed through every day (for six years) on my way to the college. I was taken aback by the first thing he said, which was to suggest that instead of taking the Underground, we should walk the two miles to UCL. He had less difficulty than I in navigating the tortuous London streets. Evidently, this did not adversely affect his impression of me, because he invited me to spend a year in Princeton with the Statistical Techniques Research Group, and later recruited me for Bell Labs. Clearly, I owe my career to John Tukey. Over the years, I had many interactions with him, but we collaborated on only one paper, a survey of the techniques of data analysis (Mallows and Tukey, 1982j).

I have many files labeled JWT, all full of attempts to understand and develop his suggestions.

We shared a passion for table tennis. I remember two occasions when we played, once in his basement in Princeton, on my first day in this country, and again in Madison where we were attending an IMS meeting. As usual, he had his paddle with him, and he led a small group of us in search of a table, where we spent a happy hour with, as I recall, an untypical absence of statistical discussion. Another interest we had in common was folk dancing. A meeting I remember with affection was in the Lake District in England, where researchers from many different fields met to discuss forecasting. John spent many hours in deep discussion with experts from many fields.

We all miss him. The world is grayer now.

ACKNOWLEDGMENTS

Thanks to present and former colleagues, especially Dave Thomson and John Chambers.

REFERENCES

BLACKMAN, R. B. and TUKEY, J. W. (1958d). The measurement of power spectra from the point of view of communications engineering, Part I. *Bell System Tech. J.* **37** 185–282.

BLACKMAN, R. B. and TUKEY, J. W. (1958e). The measurement of power spectra from the point of view of communications engineering, Part II. *Bell System Tech. J.* **37** 485–569.

BOGERT, B. P., HEALY, M. J. R. and TUKEY, J. W. (1963a). The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe-cracking. In *Proc. Symposium on Time Series Analysis* (M. Rosenblatt, ed.) 209–243. Wiley, New York.

COOLEY, J. W. and TUKEY, J. W. (1965a). An algorithm for the machine calculation of complex Fourier series. *Math. Comp.* **19** 297–301.

GENTLEMAN, W. M. and SANDE, G. (1966). Fast Fourier transforms—for fun and profit. In *AFIPS Proc., 1966 Fall Joint Computer Conference* **29** 563–578. Spartan, Washington.

HOGBEN, L. and CARTWRIGHT, M. (1969). *The Vocabulary of Science*. Stein and Day, New York.

KENDALL, M. G. and STUART, A. (1963). *The Advanced Theory of Statistics*, 2nd ed. **1**. Chapman and Hall, London.

MALLOWS, C. L., ed. (1990a). *The Collected Works of John W. Tukey VI. More Mathematical*: *1938–1984*. Wadsworth, Monterey, CA.

MALLOWS, C. L. and TUKEY, J. W. (1982j). An overview of techniques of data analysis, emphasizing its exploratory aspects. In *Some Recent Advances in Statistics* (J. Tiago de Oliviera and B. Epstein, eds.) 111–172. Academic Press, London.

THOMSON, D. J. (1977a). Spectrum estimation techniques for characterization and development of the WT4 waveguide. I. *Bell System Tech. J.* **56** 1769–1815.

THOMSON, D. J. (1977b). Spectrum estimation techniques for characterization and development of the WT4 waveguide. II. *Bell System Tech. J.* **56** 1983–2005.