Statistical Science
2018, Vol. 33, No. 3, 444–457
https://doi.org/10.1214/18-STS655

A Conversation with Tom Louis

Lance A. Waller

Abstract. Thomas A. Louis received his BA in Mathematics from Dartmouth College in 1966 and his Ph.D. in Mathematical Statistics from Columbia University in 1972. He served as a NIH Postdoctoral Fellow at Imperial College, London, from 1972–1973 and has held faculty positions at Boston University, Harvard School of Public Health, the University of Minnesota and the Bloomberg School of Public Health at Johns Hopkins University. In addition, he served as a Senior Statistical Scientist at the RAND Corporation, and as Associate Director for Research and Methodology and Chief Scientist at the U.S. Census Bureau. Tom has served as President of both the Eastern North American Region of the International Biometric Society and as President of the International Biometric Society. He is a Fellow of the American Statistical Association, the American Association for the Advancement of Science and the Institute of Mathematical Statistics. As of January 2018, Tom is *Emeritus* Professor, Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University. In addition to his many statistical accomplishments, Tom is a strong advocate for professional development and a life-long lover of time on the water.

Key words and phrases: Biography, Applied Statistics, Bayesian Statistics.

[This interview was conducted on October 13, 2014, while Tom was both a Professor in the Department of Biostatistics in the Bloomberg School of Public Health at Johns Hopkins University and detailed as the Associate Director for Research and Methodology at the U.S. Bureau of the Census. The conversation covers Tom's training and early experiences, as well as his thoughts on the "personality" of statistical methods and Bayesian thinking, and advice for junior and senior statisticians. Details have been updated to reflect the passage of time since the initial interview and in response to peer review.]

BACKGROUND AND TRAINING

Lance: It is a pleasure to talk to you today and to reflect on your experiences in and contributions to the fields of statistics and biostatistics. I've always enjoyed your perspective on ways of thinking about and doing statistics.

FIG. 1. *Tom Louis, circa* 2017.

Let's start at the beginning. Can you tell us about your general background?

Tom: I was born in central New York, and went to Dartmouth College interested in both sports and aca-

Lance A. Waller is Rollins Professor and Chair, Department of Biostatistics and Bioinformatics Rollins School of Public Health Emory University 1518 Clifton Road, NE Atlanta, Georgia 30322, USA (e-mail: lwaller@emory.edu).

demics. I had an interest in "hard" science, but not exclusively. I also liked sociology and political science. At Dartmouth, I started out as a Physics major and enjoyed it a great deal, but when we got to relativity I decided I liked math more than physics. I had essentially no exposure to statistics in high school, and it was fortunate that the Dartmouth mathematics department had active teaching and research programs in probability and statistics, both theoretical and applied. The department was also one of the inaugural academic sites for GE's time share computing with the BASIC programming language. So, in 1962 or 1963, I got my first taste of time-shared computing.

I took my first statistics course in the fall of 1964, after having taken most of the core requirements. I learned a lot from Tom Kurtz, John Lamperti and, certainly, John Kemeny. In one of life's coincidences, Don Berry and Kinley Larntz were in the class, and in 1987 we were all at the University of Minnesota (connected by statistics and cold climate).

L: Can you tell me a bit more about the GE time-sharing machines?

T: It was remote access to the mainframe, apparently one of the first, if not the first. Access was via a teletype machine; you would program in BASIC with code and the results would display on a roll of low-quality paper, printed at 32 characters a minute! Code was saved on punch tape, which took a while, but you quickly learned to save frequently, though it in no way was like hitting "save."

L: So you had an opportunity, with access to do cutting edge computing at the time...

T: That's right, at that point I was well ahead of the curve in computing (I'm confident that was the last time I had that status!).

At this time, the Vietnam War was in full swing, and it turns out, luckily or unluckily, I was deferred from the draft because of my extreme myopia and the draft never had to dig so deeply in the pool to pull out a 20/1800 guy. So, I decided that instead of going directly to graduate school, I would work for a year. I moved to New York City to work for Chubb & Son, Inc., an insurance company. They had acquired a GE time-sharing system with BASIC programming to support research and development. Early on, I was the only programmer. I developed and implemented a merit rating system for automobile insurance, which in retrospect can be labeled (happenstance) "empirical Bayes." I had passed a couple of actuarial exams but was then off to graduate school.

In the fall of 1967, I started in the Columbia University department of Mathematical Statistics. Its emphasis was on mathematical theory with signature research in sequential methods. It was said, with almost complete accuracy, that the only numbers in the courses were page numbers in books and subscripts in formulas (at present, it is a far more balanced department). There was some applied content, for example, on how to do an ANOVA. Computing was on an IBM 7094; I took a course in machine language, and also submitted FORTRAN jobs on punch cards. It was hardly my Dartmouth experience!

L: I see, so you were moving somewhat away from the data analysis you had been doing with the insurance company?

T: I would say very far away. On the other hand, that shouldn't be taken as a negative, because I certainly learned a great deal of mathematical statistics and how to do research. It may have been an inefficient way to become a biostatistician, but I eventually balanced the theory with applications, primarily via serving as TA in a course on statistics for the social sciences.

L: Were you aware of the connections between the theory and the applications you had done?

T: Not intimately. I did do a little bit of consulting. The department secretary would occasionally field phone calls and send them my way, if the outside world had a question. For example, a reporter for the *New York Daily News* called (in 1970 or 1971) with a question about the New Jersey and New York State Lotteries. Both lotteries used hard copy tickets. If you bought five New York tickets they were sequentially ordered, but five New Jersey tickets would have "random" numbers. *Daily News* readers wanted to know if there was any benefit of buying New Jersey tickets due to their random numbering. I gave all of the appropriate caveats, stating that if the same numbers of tickets are sold, if the drawing is done at random, if the payment plans are the same; then it doesn't make any difference at all whether you get the random numbers of the sequential ones. My comments were quoted with attribution, and I received interesting letters such as, "Dear Mr. Statistics: You stay up in your ivory tower, we're buying New Jersey tickets." This was my first entry into statistics in the policy arena!

L: That's right, and it whetted your appetite for it.

T: At that point, there were no social media responses–just that I should stay in my ivory tower!

L: What would you say were the most influential components of your training along the way?

T: I find that difficult to answer, just as when people ask me what my favorite movies are, what my favorite foods are, etc., because I have an ensemble of things that I like. The whole graduate experience, including teachers and colleagues, courses and other aspects were highly influential. Herb Robbins, certainly, for brilliance and simplicity. He would give lectures, either in a formal course or a talk, that would start very basic and by the time you were done, it was profound, yet not necessarily decorated with a lot of notation. So, I learned (or at least I saw!) that you could really get to the essentials without having to show off in a notational way. He was perfectly capable of doing that, but he did it only sparingly. Also, Burt Singer, for enthusiasm and virtuosity. I was his TA in a course on statistics for the social sciences and his enthusiasm was as high in that context as it was when he was teaching probability.

I worked during summers and a day a week during the academic year at IBM Watson Lab in Yorktown Heights, NY, hosted by Betty Flehinger. That connection whetted my appetite for applications and also led to my interest in adaptive allocation, which became my thesis.

Robert Wijsman, Serg Lang, John Rolph, Al Baranchik and many other faculty were very influential. Serg Lang in the Math Department, who had an analysis course that I passed (which I see as a triumph!), had phrases like "The difficulty with this theorem is not so much in the theorem, but with those who tried to prove it. I have a way of doing it!" That kind of ego was fun, especially since his was well calibrated.

At Columbia in the late 1960s, John Rolph was the primary faculty member who was embedded in doing applications as well as methodologic work. He was involved in leading the statistical side of a study in the New York City Fire Department about how many trucks to deploy for certain alarms, where to deploy them, where to place substations, and so on. He used this totally alien thing called Bayesian analysis; then he went to RAND and founded the statistics group.

In one of my, I can't really call it enjoyable, but certainly most memorable experiences at Columbia related to my thesis, I was working with Burt Singer on adaptive allocation of patients to treatments and had a nice result on preserving type I and II errors. During my defense, Herb Robbins got excited and said, "If that result is true, there must be a doubly-indexed martingale." Well, I passed my exam, but for the next three weeks, I must have spent 18 hours a day finding, or trying to find, the σ-algebra and martingale that would simplify my proof. Eventually I did, possibly we did, but it ended up being pretty exciting. Exciting and a bit exhausting that after passing my defense, I was still in the fray rather than celebrating and taking a bit of vacation.

L: Tell me about your Imperial post-doc.

T: I had one of these lucky occurrences, very much in the flavor of Dartmouth having time-sharing computing; my post-doc in 1972–1973 at Imperial College with David Cox as department chair. In that year (and certainly something I never could have arranged), in residence were David Hinkley, a faculty member; Brad Efron, Rupert Miller and Shelly Zacks, as long-term visitors, and many others passing through. Rod Little was finishing up his degree. I received at least five years of education in one year, including the clear message of the importance of motivating methodologic work by applications, having them inform and energize each other. We, meaning all of us and probably many others that I haven't listed, read through early versions of the generalized jackknife and other jackknife items, and Brad gave talks on topics that would eventually become the bootstrap. So these within-Imperial activities, coupled with "competitive teas" at Imperial, University College and Birkbeck College, and meeting people like Frank Yates, produced an amazing convergence of people and experiences. In terms of influential components, I don't know whether it should be under "training" or "professional career" because it blurs.

L: That's an impressive convergence of individuals. And after the post-doc?

T: At Harvard, Fred Mosteller was a big influence on me in terms of, well, everything. For example, he stressed the importance of what he called the "zeroth draft," writing you shouldn't glorify by calling it a first draft. Just get things down, and then you and coauthors will have something to point to and build on. Just get it down. Doing so requires that there be a safe environment for having others look at this product. By example, he advised being passionate about what you're doing, being careful about communication and organization. Fred liked nothing better than to noodle around with a draft and get the wording just right. We'd do another draft, then another, in turn. However, there was a point in every project when he'd say something like, "OK, no more bright ideas, you can still come up with a comma vs. a semicolon, but if you have any more bright ideas and you *really* want it in *this* paper, as opposed to the next one, you have to pay five dollars (mid-1980s). We'll either use the idea in this paper or the next; we'll use the money for a little party." The point was don't try for too much in any one paper, get

a nice result, get it submitted, and go on. These were a set of amazingly good lessons on approaching work, whether it be technical or prose; and how to organize a team and get things done.

L: That's good advice on both starting and finishing.

T: That's true. Start, edit like mad, he would endlessly edit, until at some point when he would stop.

Furthermore, he had fun with words: Fred and I were doing a review for *Annual Review of Public Health* on findings for public health in meta-analysis (Louis et al., 1985). We were reporting on one meta-analysis which had to do with the issue of how diet might influence hyperactivity, and how its effectiveness compared to pharmacologic interventions. The last line in our summary of that article was something like, "from the authors' report, it's clear that the dietary influences are "small potatoes" compared to the pharmacologic." The editor was aghast, and we refused five or six times the editor's request to change the wording to: "a much smaller effect" or some such. Eventually, Fred pulled rank and stated that we would pull the article unless the editor let us use the phrase "small potatoes." We/he persevered, and he taught me the lesson of standing behind your wording until you have to acquiesce or withdraw.

L: That's great. Sometimes you have to fight to say what you want to say.

T: That's right. And, it let me generalize in dealing with editors (especially nonstatistician editors in nonstatistical literature) that you need to word carefully so people understand that wording, but you also have to be willing to stand up for what may sound like improper wording to people who have no reason to understand that that the wording is needed for accuracy.

L: Sometimes words in a different order are no longer true.

T: At the same time, you don't want to geek-speak if you are talking to that other audience. You have to be careful.

THE PERSONALITY OF STATISTICAL METHODS: THE INTERPLAY BETWEEN STATISTICAL THEORY AND APPLICATIONS

L: I've always been struck with your intuition linking the theoretical and application worlds as you move between them. My impression is that you like to live between the application and the theory and tie the two together. Can you comment on this bridge between the two?

T: My overarching advice is to have a deep understanding of the statistical approaches that you may apply to an application. I mean by that not just that you know the formula or you know the conditions under which it is supposed to work, but almost the *personality* of the procedure, and little issues that increase understanding, built up by practice. On the applied side, you don't necessarily have to be an expert in the application (of course, the more the better), but you have to know enough to prevent disaster and to communicate with your collaborators so that you can translate statistical issues into their world. On the other hand, it's important for nonstatistical colleagues that we have an expectation that they don't have to know the formula for the t-test or any of those things, but they need to understand enough about the statistical procedures, what they can and cannot do, to co-own decisions on the analysis and reporting. Lincoln Moses and I (Moses and Louis, 1992) wrote this piece on the two-way street to push the idea that, yes, we always hear that statisticians need to know the application and we say, "Yes, and applied colleagues in other fields need to understand enough about statistics." We should balance the books in all cases.

L: You've also said before, to me, that you need to keep your eyes on the prize and answer the real question, is that a component of the two-way street?

T: It's a component of the two-way street and it's also a component of the personality of statistical procedures. I'll give an example that you can consider both technically and somewhat nontechnically. When I teach basic courses and we're considering a confidence interval for the odds ratio or the relative risk, I say "forget about math for a moment. What are some properties you think should apply to the confidence interval, irrespective of the technicalities?" Usually, students will suggest that the point estimate should be inside the interval and I'll say, "Good, that's certainly one thing, how about another?" If no one volunteers, I'll say "What if you have two people computing the odds ratio or relative risk and one of them does the A/B version and the other the B/A version? Do you really want conclusions to depend on whether you chose A/B or B/A?" Usually, they say "no" and we segue to that the interval should be symmetric on the log scale or that the A/B endpoints should be the reciprocal of the B/A endpoints. I stop there and say that there may be a lot of ways to do it but the big point is statistical in the personality sense, something about face validity about no matter whether you're an A/B or a B/A kind of person.

Another example comes to mind, and none of these came at day 0, day 1 or year 25 of my career, they keep accumulating, and they only accumulate by getting involved in applications. Statistical procedures, unless you make them do otherwise, especially if they are likelihood based, will produce optimal estimates based on the assumptions that produced the likelihood. For example, they'll use optimal, variance-minimizing weights; they can't help it! Why would you want them to do anything else? But, if you're in an application where you may have to address, say, a weighted average of hospital-specific length-of-stay where *you* want the procedure to give equal weights to the hospital or you want a procedure where you want to use survey-based weights to infer results for a broader population, you'll have to use a procedure that doesn't do these variance minimizing statistical weights, but addresses the goals *you* want to address. You should always step back and consider the question that you want to ask, and here's how I want to answer it. There may be programs to do it or, you may need to write something new that will address *that* goal as opposed to the default goal.

L: I think a critical component to being a statistical methodologist is understanding what questions a method answers versus what questions we *want* to answer for an application. I really like the idea of a "personality of a procedure."

T: And just to expand on that, the odds ratio is the natural parameter for the problem of comparing two binomials conditioning on marginals, it pops right out. So, what could be more natural than using the natural parameter, right? Well, if you want to evaluate the risk difference, you need to target something that *isn't* the natural parameter. It's an example of where it is good to know the theory, good to know sufficiency and other such things, but it is also good to know what you really want to do in practice and how to accomplish it.

L: And where does design fit into this?

T: Well, it must fit in. Many educational programs have moved away from design, but my attention to it heightened during my time at the Census Bureau. Design, if not everything, certainly is the first thing. If you take it to the extreme, and I'll take a genomics study where you're running an expression analysis and you do something dramatically silly like assay all of the tumors in one day or on one run, and all of the nontumors on another day and another run, you have absolutely no idea whether what you are seeing is a batch or a tumor effect. And it sounds as though I'm stretching, but there have been many examples, even recently, where these high-throughput machines with millions of dollars spent on the machine but not sufficient attention to asking "Well, once we run the machine and generate the data, will we be able to answer the question?" At minimum, to avoid the experience of, "Come to think of it, we can't get there from here…," let's at least have a design, whether it's ideal or not, that will support the primary questions of the study. I'm not comfortable with pushing for *optimal* designs unless you absolutely know that there's only one endpoint that you would ever want to assess and that the underlying assumptions are absolutely valid. You definitely want *effective* designs, ones that are sufficiently robust to departures from the assumptions.

Our statistics and biostatistics departments need to reemphasize design. I don't mean to imply that everyone needs to know A-optimality and D-optimality by heart, but we do need to broaden what has been quite good for clinical trials to other applications. I don't want to be quoted as saying we're not doing it, but we do need to have a renewed emphasis.

L: I think that's a good point and that we're at an interesting technological stage where there is rapid breakthrough in what we *can* measure, but not as much follow up on how we *should* measure it to address the questions that come up.

T: One of my favorite quotes is that "Space-age statistics will not rescue stone-age data."

L: Even if the data come from a space-age machine.

T: In the spirit of public health, treatment is fine but prevention is even more important.

L: You've discussed the individuals you interacted with, ideas you came across and computing you came across, all of which fed into your development as a statistician, and you've mentioned some advice from some individuals. Are there particular developments that have come along that influenced you? You mentioned the zeroth draft and you've mentioned to me in some settings the "level 0" design. What are some of these tidbits of advice you've picked up that have influenced you?

T: The level 0 design is my phrase, I'm not sure it's very artful, but the idea is that the construction of a research team involves finances, hardware/software and certainly personware. If you don't get these at least approximately right, you are fated to fail. We, as statisticians, have information to offer on how to construct an effective team. Many of us are proactive on that, but we have to keep it up because if we don't, the project won't work well, and it is perfectly possible that collaborators who have not had the opportunity to work

with good statisticians on a good team, will conclude that statistical input isn't worth it. They may be right, because the resourcing may not have been done properly.

So, level 0 design is these components along with some political concerns. A few years ago, there were two major studies looking at whether tissue plasminogen activator treatment (tPA) was effective in reducing the damage due to stroke. These were big studies funded by the U.S. National Institute of Neurological Disorders and Stroke (NINDS) with tPA coming from Genentech. Both studies showed, at least in general terms, that tPA worked well, especially if given within 90 minutes of the event. However, there was a broad community that disagreed strongly with those findings, either for scientific reasons or because they were concerned that some investigators had conflicts of interest. So, NINDS decided to have a group of statisticians independently review the studies, and it ended up being Mike O'Fallon, Vicki Hertzberg, and I. The first thing we communicated to the NINDS was that we wouldn't do the review unless there are also three clinicians on the team, and one of them must be a person who is identified with the critics. This person did not have to be an extreme critic, but had to be identified with that camp. We required the clinical input because though it was thought the issues were statistical, in fact they would inevitably have considerable clinical content. The NINDS pushed back, and we made it clear that without those team members, we won't do the project. Eventually they agreed to our recommendation, and also that we would pick the team. Then we communicated that we need to control data analysis, and that it couldn't be done by NINDS internal statisticians. We wanted to work with our experienced staff, so, for example, we could say, "Fit a random effects model, with ..." and not have to worry about it being properly conducted. But, as important, we wanted arm's-length distance from NINDS on the analysis. All of the foregoing was to make sure that whatever our findings, they would not be and would not be perceived to be influenced by the NINDS. For credibility on all accounts, the NINDS needed to take the chance.

Our results (see Ingall et al., 2004) primarily supported the studies' conclusions with a few critiques. If we hadn't had that distance and if we hadn't had the three clinicians on the team, it would have been very ill advised to have done the review, because we wouldn't have known enough about the clinical issues to meld those with the statistical, and we wouldn't have had arm's length credibility.

L: That's an excellent point that the effort you put into setting up your team is often just as important as the effort the team puts into the project.

BAYESIAN THINKING

L: Let's change topics a bit and discuss how having a Bayesian viewpoint influences your interactions with students and collaborators.

T: In the classroom and in the field, I like having my collaborators and me think of full distributions. You might say this is because I'm a Bayesian and that certainly is relevant, but it's more that I want my collaborators to think stochastically and consider the distribution of attributes of interest. Or, in Bayes-speak, to consider the full, joint posterior distribution of the attributes, and formulate scientific or policy questions to ask of it; what inferences to make. Frequently, it will come out that the center of that distribution isn't of primary interest, it might be the tail. By thinking in terms of full distribution processing, one doesn't automatically focus on the mean and variance; rather you ask, "What is it I'm after and is it supported by this (joint) distribution?"

I haven't done this yet, but the next time I'm teaching a basic course, when I introduce the normal distribution rather than stating that it's characterized by its mean and standard deviation, I'll say it's characterized by its 10th and 90th percentiles or some other pair of percentiles, to emphasize that it's the distribution that matters and not necessarily the mean and standard deviation.

L: I think that we do lose sight of that.

T: One more recommendation that comes from this blending of theory and practice is to have a point of view when you're approaching a problem. Mine is, at least generally speaking, Bayesian, but in live applications you can't be too doctrinaire. Your goal is to get the job done well. I use my viewpoint as an aid to navigation rather than a straightjacket. It's helpful to have a viewpoint, but it shouldn't be too rigid; you need to be open to the reasonableness of the views of other statisticians, of collaborators, and attentive to limitations.

L: You do have to have a point of view to know the tools you are going to use but if your point of view drives the solution and a different point of view yields a different solution, we really wouldn't have much of a profession.

T: In situations where the data aren't terribly informative, and maybe two different points of view *do* come up with different solutions, the most important

finding is that they are different to motivate a diagnosis of why.

L: That's right, there is more to a point of view than the calculations you do. There's the contribution of the data in any statistical procedure but there are a lot of nondata contributions lined up in your assumptions and data collection. One thing I've learned from you and the Bayesian point of view is, at least in my experience, Bayesians tend to list all of those.

T: What I like about it is being able to consider what if we actually *knew* the underlying truth (it might be a distribution), what action would we take or conclusion would we draw? Of course, frequentists are also permitted to engage in this thought process, but in complex situations the Bayesian formalism facilitates linking the unknown truth (we could rank the schools if we knew their true performance), and we can use the posterior distribution to structure the inference using the rules of probability. This approach is especially important for complex goals. My principal point is to consider goals and actions as though you knew the truth and then use data to address the goals and inform the actions.

It is interesting that asking, "What would you do if you knew the truth?" usually generates silence, and the request, "Couldn't you put in some uncertainty to make it easier?" is latent but palpable.

BIG IDEAS IN STATISTICS

L: What other ideas and concepts and individuals have caught your attention and propelled you on over the course of your career?

T: Well, one of the unifying and overarching ones is sample reuse. Start with the earliest form of sample reuse: computing the sample variance, using the very same data set you used to estimate the sample mean to provide an estimate of its uncertainty, then confidence intervals, etc. It's grown to be everything from general cross-validation to jackknifing and bootstrapping. To me, it's one of those big ideas that have allowed expansion of model-based approaches to inference and supported inference from algorithms. This leads to the developing harmonization of model-based and algorithmic approaches. Recall that the classic Breiman (2001) article emphasized the divide, but let's face it, every model-based procedure we apply manifests itself as an algorithm. And, good algorithms are, at least in some ways, informed by model-based thinking. So, my view is to encourage creativity, but make sure that as statisticians we are promoting evaluating properties, either via mathematics or simulation. With increased understanding, some algorithms will end up being model-based, some model-informed, some purely creative approaches.

Regarding some more sociological developments, as a field, certainly as exemplified by the American Statistical Association and other associations, we have and communicate increasing pride in our field. In my International Biometric Society Presidential Address in 2006 (see Louis, 2007), I stated that the respect we have for statisticians and statistics will be no greater than the respect we transmit to others. Why should anyone respect us more than we respect ourselves and our field? So, let's keep it up and amp it up.

L: Do you think this raises our own personal standards?

T: It raises our personal standards and at the same time we have to be careful not to make them so high that nothing is good enough. One of the arts and crafts of getting involved in applications is to do a "good enough" job. I am not promoting mediocrity, but perfection may not be possible or we may not even know how to recognize it. Also, we are engaged in a variety of activities, so you have to evaluate for each when to hold and when to fold; to decide how much effort is required in each instance.

L: I agree, there may be effective solutions that may not necessarily be optimal and solutions that are optimal under a particular measure of optimality that misses some of the other components that are actually part of the question.

T: A few other developments, and I mention computing for two reasons. One is, obviously, we get things done via computing, including impressive graphics. But the ability to deal with large databases and activate high-CPU cycles required has opened up imagining new methods that we might have been able to think about before there were these capabilities, but there would have been no reason to do so. We wouldn't be considering them, if all we had was an abacus! So, there is a sociology of knowledge, or a technology of knowledge where new measurement systems and new technology (computing in this case) activate the imagination and implementation of new statistical procedures. These, in turn, energize technical innovations and there is exponential growth in all domains.

L: It's interesting to look back in the literature and see some assumptions of computing capabilities at their time and you'll get these glimpses of people saying, "It would be possible to do this, but why would anyone want to do that?"

T: Yes, you can back to Quenouille (1956), who split a sample in half and did the jackknife to bias-correct. But, one split doesn't support good variance estimation, and the "leave one out" jackknife does it well. Pulling it off required additional computing power and affordability. These are prerequisites for the bootstrap and all computer-intensive procedures.

L: Yes, I believe the first discussion of Monte Carlo tests says, "suppose you could repeat this ten times..." (Barnard, 1963 in discussion of Bartlett, 1963). I believe Julian Besag in 1975 points out that MCMC could work, but seemed impractical (Besag, 1975). Linking theory and computing is often a function of time and the capabilities available.

FAVORITE COLLABORATIONS

L: You mentioned some of your favorite collaborations, are there any others you would like to highlight? You've worked on some important problems, some big problems, but maybe those aren't the favorites.

T: Here's a smallish sample in chronological order: Working at Boston University with Art Albert and others on statistical methods for cancer screening; a nice blending of methodological development and some relevant applied findings. That was an exciting and wonderful project in which I was primarily a mathematical statistician, but was motivated by and dived into the application including estimating a disease natural history from screening data (see Louis et al., 1978).

Then, at Harvard, there's no question that Nan Laird's and my many articles and other collaborations on empirical Bayes and Bayes are a highlight.

I enjoyed and benefitted from collaboration with Steve Lagakos, Louise Ryan and others working on the "Carcinogen Bioassay in Small Rodents" methods grant funded by the U.S. National Institute of Environmental Health Sciences. We produced several methodological developments with applied impact, and most definitely, we had a lot of fun. The project led to a successful training grant in Environmental Biostatistics.

The *New England Journal of Medicine* project: generated many articles in the *NEJM* and a book. The project was led by Fred Mosteller, and included John Bailar and several other faculty, pre-docs and post-docs. Part of its importance to me was watching Fred at work, and also learning a great deal about how to write to and for clinicians about statistical concepts, and how we have to be likewise educated by clinicians.

L: Can you tell me a little more about the *NEJM* project?

T: Funded by the Rockefeller Foundation, the goal was to write a series of articles in the *NEJM* on various issues in statistical practice, for example, what are the big issues in doing a crossover study. The focus was to be on concepts and examples, not the math. Other topics included longitudinal studies, studies without internal controls, the Moses/Louis Two-way Street, etc. In each of these, the approach was to base content on articles in the *NEJM* and other medical journals, but we also got access to some submissions that didn't make it to publication along with the reviews, so we could see the filtration process. The *NEJM* published our articles (after extremely rigorous review!) and also the *Medical Uses of Statistics* book that consolidated our work. There have been, I believe, three editions. The work wasn't technical, not theorem/proof. Rather, it focused on communicating best practice by mining the *NEJM* and other journals for examples of good and poor practice, but with the emphasis on good practice. Channeling Mosteller, it's fine to communicate the illness, but we should primarily focus on health, "Here's how to do it well" not, "Look at all of these bad examples."

L: How about collaborations at the University of Minnesota?

T: At Minnesota, due to the mission-orientation and the broad inclusion of collaborators, serving as co-Director of the Statistical Center for Community Programs for Clinical Research on AIDS was the big highlight of my research life from the early 1990s through early 2000s. Working with Jim Neaton and many others, we were the "other AIDS clinical trials network." Our studies operated in community health centers, methadone maintenance clinics and primary care hospitals; not in high-tech, university-based centers. These latter were the province of the AIDS Clinical Trials Group (ATCG). We spent a lot of time working with community advocates, and had statistical retreats with project clinicians and epidemiologists, community representatives and persons with AIDS. We gave workshops on trial design and analysis, these leading to group exercises. The goals were inclusion and ensuring that those participating on trial design committees would be able to speak the language, know the concepts and co-own the statistical issues and be comfortable stating things like "Don't you think we ought to design-in one more follow up so we get a little less censoring?" Successful co-ownership was exciting in that it served the mission of dealing with a very deadly disease, and because, as statisticians, we did help a great deal. I got to work with several people with AIDS and many community advocates who were very committed

to saving lives, and to learning enough about research to help with it.

Writing the Bayes book with Brad Carlin was enjoyable, with Carlin and Louis (2009) the most recent edition. It started with Brad and I co-teaching the inaugural Bayes course. Almost every day, the students would witness my stating something and Brad disagreeing!. Then he'd make a statement and I'd disagree. Finally, we thought it best to write a book and sort out our disagreements.

There were several other projects including the Lung Health Study, and the environmental justice study with you. All were was productive and enjoyable.

L: The environmental justice study turned out to be foundational to my own thinking, showing me how to think hard about what the questions were and to use distributions to answer these questions.

STATISTICAL LEADERSHIP AND NATIONAL COMMITTEES

L: Can you comment on your time at the University of Minnesota, RAND and Johns Hopkins, as well as your participation in National Academies' committees and projects?

T: I was recruited to Minnesota as chair of Biostatistics and am quite proud of that service. I could talk at length about it, but will focus on a few highlights, with the first, of course, being hiring you! In addition, when I arrived, the Biostatistics and Statistics departments were completely separate with almost no joint teaching and research. When I left in 2000, the educational programs were pleasingly integrated, and faculty were engaged in a fair amount of joint research. Administratively, I learned that running a small department, as it was when I arrived, was considerably easier than running a large one, as it was when I left. Fortunately, my management skills grew along with the department. From the start, I ensured that junior faculty careers were structured for success, with responsibilities and expectations that lined up with advancement. Of course, the faculty member had to generate the successes, but the environment was empowering.

T: In my never-ending quest to move around, I left Minnesota and helped give the statistics group at RAND a Washington D.C. presence. At RAND, I engaged in several interesting projects, including value added modeling to rate and rank teachers and schools (see McCaffrey et al., 2004). Our technical work was accompanied by cautions on not putting too much faith in the ranks. The research blended technical and applied evaluations, and operated in the context of a host of political considerations.

Then, I was off to Hopkins, carrying with me my hierarchical models grant that included collaboration with my RAND colleagues. Also, at Hopkins I collaborated on "Healthy Pathways for Students and Schools," a follow-up study of grade school kids, with assessments of their school performance, their home life situation and health status; all linked to grades and other aspects of school performance. I was involved in two clinical studies, one on treating eye disease (Kempen et al., 2011), and another on interventions for weight loss (Appel et al., 2011). I continue to be involved with the International Center of Excellence for Malaria Research, which has sites in sub-Saharan Africa and projects that include population-based epidemiology, designed interventions for malaria control (Sutcliffe et al., 2012), and mosquito genomics.

L: One thing I'm struck by is that your experiences and the things you've worked on certainly cover a lot of ground. There are common themes such as ranking and rating with different sources of variability, and linking clinical trials to observational studies and back again.

T: My activities with the National Academies and the Health Effects Institute have been exciting and pleasurable, but also very educational. I've learned about a wide variety of applications, and worked with a broad range of statisticians, other scientists and policy types.

L: Some of those lead directly to your interest in taking your detail at the Census Bureau.

T: Yes, some of the National Academies projects were survey related; these gave me some credibility to take the three-year, Associate Director of Research and Methodology/Chief Scientist position after Rod Little's term.

L: You were on several National Academies panels such as Gulf War Syndrome, CNSTAT (the National Academies Committee on National Statistics), etc. Any thoughts on those experiences?

T: I was on CNSTAT for six years and served on panels for other divisions of the Academies include the panel on the Health Effects of Service in the Persian Gulf War, which was administered by what was then the Institute of Medicine and is now the National Academy of Medicine. The project started in the mid-1990s and was related to Desert Shield (see Committee 1996). We were the first of many National Academy panels to study Persian Gulf War Illness, to review the evidence and though the syndromes were very real,

we couldn't come up with convincing proof that they were or were not service-related. Our reports communicated what information needed to be collected to have a chance of making a causal connection. I also learned about the penetrating properties of projectiles made from depleted uranium, and the health effects of exposure to sarin and to oil well fires.

L: Sometimes I think the best outcome of a panel, or the most helpful, is what you would need to know in order to make that case. What evidence is available, what evidence is required, what can you say with what you do have and what can't you say with what you do have.

T: Service on the Persian Gulf War panel provided an opportunity to interact with the press and public on a hot topic. I was one of the panel members designated to be on radio shows after our initial report was released. I learned to focus on at most three messages and stick with them. A Gulf War veteran would call-in and state that they were ill, and wonder (not calmly) how could we not conclude it was due to his service? I would respond in all sincerity that I have no doubt that you are ill and should be receiving care (message 1), but that we can't yet link your illness to your service (message 2); however, we have not excluded a possible connection, and are continuing our study (message 3).

L: In working in those policy areas and in your leadership roles in ENAR and IBS over time, what changes do you see, not just in the statistics profession and how we do it, but in how the statistics profession interacts with the broader policy and science worlds. Have you seen a change in that?

T: Yes, I do see a change in that. As a population of statisticians, we are just as adept as we ever were at methodological development and application. I don't want us to get away from that, but we have become much better at projecting and promoting ourselves and our profession in policy and other areas. A sufficient number of us are paying attention to how and what to communicate, over and above, for example, that the t-test for a sample mean has $(n - 1)$ degrees of freedom! Furthermore, we have an increasing number of excellent examples of how and why statisticians (with proper resourcing and roles), add considerable value.

For example, say 20 years ago, the idea that the ASA would have a staff person primarily engaging in policy wouldn't have even crossed its corporate mind, and now I would say that the ASA has become a go-to place for people providing commentary to the press and to the government. The Joint Statistical Meetings and other ASA-sponsored meetings have an increasing number of sessions that address topics and target an audience far broader than card-carrying statisticians.

L: I agree with that, and I think that, as a profession, we always assumed people would come to us for the answer but I think one thing we've learned is that there is some effort involved in being recognized, and I think it has paid off well.

T: I can remember back 30 years, the guideline was that we can't have opinions, that our role is to provide technical input, do the computations and provide some explanation of results, but to go no further. However, my view, and I believe that of our profession, is that we have an obligation to go further, to collaborate in interpreting and translating findings to a broad set of stakeholders—that it is irresponsible to shy away from this.

L: That's a long way from the Statistical Society of London's original motto which was a Latin phrase meaning "Let the others thresh it out." (*Aliis exterendum*, Hilts, 1978.)

T: We should help others thresh it out, and we should thresh it out as well.

L: You mentioned big data having a component in this data science. any comments on these more recent developments?

T: As many have observed, we no longer have a near-monopoly on access to data. It used to be that if we collaborate with a, say, clinical group it would be that they and we had data access and that would be it. But now, data are coming in via fat pipes from all sources and to a wide group of receivers. If we aren't where the pipes spew out, we're probably not going to be involved, and surely not in a leadership role. Happily, we are involved, but we need to keep up our energy and not be complacent. I'm not worried at all about statistics and statisticians remaining relevant. Yes, there are many other groups with equal access, but with care our roles can only increase. I liked what Jeff Leek wrote in one of his Simply Statistics blogs on data science. His closing line was something like, "The most important word in 'data science' is 'science'." We're not *uniquely* the ones who bring science to the analysis of big data, but we are the best in understanding the selection effects, the sampling plans and issues of confounding, and we are the inference professionals. A big issue is that big sample size doesn't necessarily imply big information or big validity; that care is needed. We'll be collaborating, doing analyses and will hang on, by our fingernails if necessary, to traditional values.

STATISTICS TODAY

L: What do you find most interesting about our field today?

T: It's the same thing I found interesting about it from the start, at least from the start of getting involved in both methods and applications—the broad relevance of statistical concepts and technical approaches; their applicability to virtually any area, and the excitement of being able to participate in challenging and important projects. Some of us concentrate in a small number of areas, take a deep dive and stay there. Others, and I'm one, behave a bit like an oil slick, spreading widely (hey, I'm on the scientific advisory board of the NIEHS GuLF Study of health effects of the Deepwater Horizon oil spill!). It is this transportable set of approaches and concepts that continues to be valuable. That is, to me, the most exciting aspect on the societal and personal levels.

Continuing, we are increasingly collaborating with computer scientists. Record linkage is one example, wherein statisticians have been instrumental in developing probabilistic linkage. For example, the question, "Is this the same Joe Smith?" has been replaced by "What is the probability that this is the same Joe Smith?" There are considerable technical challenges, but also ethical ones related to disclosure limitation. Government agencies, the biomedical community (e.g., disclosure of gene signatures) and beyond are confronting the trade-off between the usefulness of information and the consequent disclosure risk. Statistical approaches have a central role in quantifying these trade-offs and improving the operating characteristics.

L: To quote Spiderman, with great power comes great responsibility. And this played directly into your position at the U.S. Census (current as of the interview), correct?

T: Yes, one thing to mention is that I served (in 2013–2015) as the Associate Director for Research and Methodology coincident with being on the faculty at Hopkins so the Census position is really a perfect example of the elevation of statistics by having that position be of co-equal status to the Decennial Census, Economic Programs, Demographic Programs, IT, Field, Communications and Administration.

L: It's not small potatoes anymore.

T: One more comment on the associate directorship: Rod Little (University of Michigan) was the first of the newly constituted Associate Directors; I served as the second, and John Abowd from Cornell served as the third. Rod and I were both chairs of Biostatistics departments and, although the Census context is a little different, the mandate of that Directorate is to collaborate on projects, and be a driving engine for methodologic development with many of the new methods migrating back into production. That's really at least two of the three missions of a Biostatistics department with the third being education. As of 2014, the Census bureau is also getting involved in education, for example, by offering Big Data short courses. So, the application's a bit different but the administrative and sociological issues are very similar and it's not completely illogical that Rod and I were the first two associate directors.

L: There's a certain vibrancy in both worlds where you are encouraging the creativity in both the collaborative and methodological research component and being in a place that values both is really critical to the success of what the Census wants to do and what a Biostatistics department wants to do.

ADVICE FOR NEW STATISTICIANS

L: What advice would you have for new statisticians?

T: Enjoy is point one. Next, focus early on in your career, then broaden. Be aware of the mathematics of commitments, for instance, five times twenty far exceeds twenty times five. That is, twenty 5% obligations far exceeds five 20% obligations. Go for the five twenties!

Be sure you're in a context with excellent mentoring, with senior people around to advise you on these decisions and support you. That's a win/win, because it will be best for them in the long run and best for you too.

Literally make appointments with yourself, make sure of that. If you are in any kind of applied context, meetings with others will *feel* like they are more important, but you have to elevate the ones with yourself for reading the literature, and working on research, whether it be directly connected with the project or not, to have at least co-equal status. The only way to make it feel that way is to put them on your calendar. It's not like you would never modify them, just like you would certainly modify the other appointments, but they should have essentially equal standing so you don't look up in a month and realize that those Friday mornings that you were going to save for both research and perusing the literature have evaporated.

L: That's something I'm still learning.

T: I am too, that's why it made my list.

L: Do some service, whether you are in academe or other settings, but be careful that you don't do too much early on. Through service, you get to know a lot more about, say, being on a review committee, and you also *get known*, and assuming you are known to be good, you create further opportunities.

Travel but not too much. Each person's situation is different, but some professional travel gets you out and exposed to new environments, and new viewpoints. However, you need to be in your home office enough to keep up the traction and action there. Proximity breeds the spontaneous drop-ins and discussions that create a local culture and generate ideas.

Recognize that applications are the principal idea-generators, they have and continue to empower most of the important methodological research, so engage. Without question, our field also needs excellent mathematical statisticians, who devote most of their careers to developing innovative approaches and evaluating them in a deep mathematical way. But, statistics doesn't benefit a great deal from subtle refinements of asymptotics (probability theory does benefit!).

Here's one example of an application driving research. In the late 1970s, I was talking with Jack Wennberg about his research on variation in surgical rates in Maine, New Hampshire and Vermont. He was comparing histograms of hospital-specific hysterectomy rates and other procedures within each state. He asked something like, "Tom, what do you think of these, could I use a Bayesian approach?" I said, "Well, the ones you are producing are too spread out (they were histograms of the hospital-specific MLEs) because they incorporate both the variation in the true underlying hospital-specific rates, and the sampling variance." He replied, "Oh, I can see that, so what should I do?" I said, "You should do the histogram of posterior means." During my drive back to Boston, I thought, "Oh my, using posterior means isn't correct, and I don't know what is correct." That episode motivated my research on ensemble/histogram estimation (see Shen and Louis, 1998). I don't know if I would have ever gotten involved in that line of research, if I hadn't had that conversation with Jack.

One of the most nerve wracking and exciting things in a new statistician's professional life is to get involved in meaningful applications, for example, serving on a clinical trial monitoring board, or involvement in any study where the results truly matter. Results might change environmental policy, or determine use of a clinical intervention. Results that matter, with consequences far greater than getting a poor grade on a homework assignment, definitely focus the mind on getting it right.

Early on, develop a three and a five-year plan. What would you like your professional profile to be in three years, in five years and how do you get there? One aspect is evaluation of job opportunities, whether a post-doc, a faculty position or one in industry or government. You should try to go to a place that is good to be *from*. I don't mean that you are planning to leave, but that if you take advantage of the context, you will grow professionally, technically, conceptually and politically. If so, it is a good place to have been, and of course you may stay forever.

There are two styles, I've been a "mover," others are "stayers" who stay at a place and become an institution in their institution. I moved with good reason each time, but occasionally I wonder what my career would have been like if I had stayed, for example, at Boston University. There are the two styles; you don't have to decide early in your career between the two, but have a strategic plan to help with evaluations.

My last piece of advice on this list is that sociology and psychology are very important. They don't replace being a good technician or being good at understanding and communicating statistical concepts, but to be effective at collaboration you need to co-own goals and work well in a group, occasionally with difficult personalities. We are the statistical experts, but should avoid messaging the collaborative equivalent of "only historians are allowed to reminisce." We should be sufficiently comfortable with our expertise to invite collaborators to make statistical proposals; frequently they come up with comments and questions that are revealing, prompting the thought question, "why didn't I think of that?"

ADVICE FOR SENIOR STATISTICIANS

L: Any advice you want to add for the more experienced among us in our field?

T: Well, I love the careful wording, since I am now three days past my 70th [now 73.5] *birthday*. So, what do I want to say to the old fogeys of the world? One point is, at a minimum, learn about and appreciate the 21st century statistical world. It seems that I can still add value as a 20th century statistician. Definitely, reinvent yourself a bit. It's difficult to do a complete makeover, but be involved enough to know what is going on and help. And, don't grumble by stating that, "Things aren't like they used to be," it's likely they never were! Be positive. There are a lot of exciting

developments, many aren't traditional statistical activities, but they are the future of our field.

The last one is that traditional values still apply. Design, conduct, analysis and reporting are still very important. For example, the sampling plan determines much of what you can do with data. We should pay attention to these issues, not to stymie creativity, but in fact to generate it by understanding threats to validity and ways to ameliorate them. So, to those of us who are "experienced," I recommend remembering that those traditional values are still the core of our field, and it is pleasing that they manifest in an exciting environment.

L: It's been great running through a lot of topics. Any final parting thoughts?

T: I'll mention just a couple of things. One is that the exciting developments in our field have moved us back toward our origins based on deep collaboration in important applications. As I've already remarked, it continues to energize methodological development, and also generates the need to publish methods in the subject matter journals. Though that is both necessary and good, it's also true that we should still be evaluating and generalizing our applications to the point where statisticians engaged in other applications can learn about them. Of course, publish articles in *Human Genetics*, but also for a subset of methods, spend a little more time exploring properties, possibly enhance performance, make it a bit more generic and submit to a statistical journal. There needs to be a balance between the two publication types. We need to support publishing in domain-specific journals, to embrace the excitement of leading edge collaborations that must publish quickly. But, we also need to nurture the statistical commons. That may appear self-serving, but I'm convinced it serves society.

That comment segues into something we discuss often. What are the implications of big data, data science, computing and other developments for our training programs? What should be our core requirements? What should be on the qualifying exam?

L: I think it keeps your discipline alive to ask and answer those questions.

T: What if we had not changed from the 1940s? We would not be in good shape at the moment!

L: We should recognize that we are constantly changing a little bit and sometimes we change a lot, but that, getting back to your point about things not being what they used to be, I don't think our field has *ever* been static.

T: No it hasn't. There have been game-changing developments and maybe some of these are too. I'll close by saying that I've enjoyed my profession, my colleagues and my students, I can't imagine having chosen a better field, and as long as my brain keeps operating, I plan to be participating.

FIG. 2. *Tom Louis, on the water.*

REFERENCES

APPEL, L. J., CLARK, J. M., YEH, H.-C., WANG, N.-Y., COUGHLIN, J. W., DAUMIT, G., MILLER, E., DALCIN, A., JEROME, G., GELLER, S., NORONHA, G., POZEFSKY, T., CHARLESTON, J., REYNOLDS, J., RUBIN, R., LOUIS, T. A. and BRANCATI, F. L. (2011). Accomplishing weight loss in clinical practice—A comparative effectiveness trial. *N. Engl. J. Med.* **365** 1959–1968.

BARNARD, G. A. (1963). Discussion of "The spectral analysis of point patterns". *J. Roy. Statist. Soc. Ser. B* **25** 294.

BARTLETT, M. S. (1963). The spectral analysis of point processes. *J. Roy. Statist. Soc. Ser. B* **25** 264–296. MR0171334

BESAG, J. (1975). Statistical analysis of non-lattice data. *Statistician* **24** 179–195.

BREIMAN, L. (2001). Statistical modeling: The two cultures. *Statist. Sci.* **16** 199–231. MR1874152

CARLIN, B. P. and LOUIS, T. A. (2009). *Bayesian Methods for Data Analysis*, 3rd ed. Chapman & Hall/CRC Press, Boca Raton, FL.

COMMITTEE TO REVIEW THE HEALTH CONSEQUENCES OF SERVICE DURING THE PERSIAN GULF WAR (1996). In Health Consequences of Service During the Persian Gulf War: Recommendations for Research and Information Systems. National Academies Press, Washington DC.

HILTS, V. L. (1978). Aliis exterendum, or, the origins of the Statistical Society of London. *Isis* **69** 21–43. MR0495096

INGALL, T., O'FALLON, M., ASPLUND, K., GOLDFRANK, L., HERTZBERG, V., LOUIS, T. A. and CHRISTIANSON, T. J. H. (2004). Findings from the reanalysis of the NINDS tissue plasminogen activator for acute ischemic stroke treatment trial. *Stroke* **35** 2418–2424.

KEMPEN, J. H., ALTAWEEL, M. M., HOLBROOK, J. T., JABS, D. A., LOUIS, T. A., SUGAR, E. A. and THORNE, J. E.

(2011). Randomized comparison of systemic anti-inflammatory therapy versus fluocinolone acetonide implant therapy for intermediate, posterior and panuveitis: The multicenter uveitis steroid treatment trial. *Ophthalmology* **118** 1916–1926.

LOUIS, T. A. (2007). Our future as history. *Biometrics* **63** 1–9, 309. MR2345569

LOUIS, T. A., ALBERT, A. and HEGHINIAN, S. (1978). Screening for the early detection of cancer III: Estimation of disease natural history. *Math. Biosci.* **40** 111–144.

LOUIS, T. A., FINEBERG, H. V. and MOSTELLER, F. (1985). Findings for public health from meta-analyses. *Ann. Rev. Public Health* **6** 1–20.

MCCAFFREY, D. F., LOCKWOOD, J. R., HAMILTON, L., KORETZ, D. and LOUIS, T. A. (2004). Models for value-added modelling of teacher effects. *J. Educ. Behav. Stat.* **29** 67–102.

MOSES, L. and LOUIS, T. A. (1992). Statistical consulting in clinical research: The two-way street. In *Medical Uses of Statistics*, 2nd ed. (J. C. Bailar and F. Mosteller, eds.). NEJM Books, Waltham, MA.

QUENOUILLE, M. H. (1956). Notes on bias in estimation. *Biometrika* **43** 353–360. MR0081040

SHEN, W. and LOUIS, T. A. (1998). Triple-goal estimates in two-stage hierarchical models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 455–471. MR1616061

SUTCLIFFE, C. G., KOBAYASHI, T., HAMAPUMBU, H., SHIELDS, T., MHARAKURWA, S., THUMA, P. E., LOUIS, T. A., GLASS, G. and MOSS, W. J. (2012). Reduced risk of malaria parasitemia following household screening and treatment: A cross-sectional and longitudinal cohort study. *PLoS ONE* **7** e31396.