Statistical Science
2019, Vol. 34, No. 2, 334–348
https://doi.org/10.1214/18-STS691

A Conversation with Robert E. Kass

Sam Behseta

Abstract. Rob Kass has been been on the faculty of the Department of Statistics at Carnegie Mellon since 1981; he joined the Center for the Neural Basis of Cognition (CNBC) in 1997, and the Machine Learning Department (in the School of Computer Science) in 2007. He served as Department Head of Statistics from 1995 to 2004 and served as Interim Co-Director of the CNBC 2015–2018. He became the Maurice Falk Professor of Statistics and Computational Neuroscience in 2016.

Kass has served as Chair of the Section for Bayesian Statistical Science of the American Statistical Association, Chair of the Statistics Section of the American Association for the Advancement of Science, founding Editor-in-Chief of the journal *Bayesian Analysis* and Executive Editor of *Statistical Science*. He is an elected Fellow of the American Statistical Association, the Institute of Mathematical Statistics and the American Association for the Advancement of Science. He has been recognized by the Institute for Scientific Information as one of the 10 most highly cited researchers, 1995–2005, in the category of mathematics. Kass is the recipient of the 2017 Fisher Award and lectureship by the Committee of the Presidents of the Statistical Societies. This interview took place at Carnegie Mellon University in November 2017.

Key words and phrases: Statistical training, Bayesian statistics, statistics in neuroscience, academic life, statistical narrative.

EARLY YEARS

SB: Rob, where did you grow up and where did you go to school? Were you an enthusiastic student?

RK: I grew up in the suburbs of Boston. My father was a well-known professor at Harvard Medical School. He was an M.D.-Ph.D., which was unusual back then. His Ph.D. was in microbiology and he did research primarily in infectious diseases, and also in epidemiology of hypertension (high blood pressure). As for school, I was not a good student, but I did love math. I went to a private school, linked into a large network of New England private schools. They were trying to imitate the English school system. I just didn't have patience for a lot of the stuff. One of the subjects we had to learn was Latin. I remember one time I got a C on my Latin translation. The teacher wrote a one-sentence comment on it, "Much of this is gibberish"! Then I went to a wonderful high school. It was what they used to call a free school because it was small and progressive, meant for kids who would not do well in traditional schools. This would have been in 1967. I was there through tenth, eleventh, twelfth grade, 1967 through 70, the peak of the sixties.

SB: You must have done other things before graduating from Antioch College because when I was looking at your CV, there are other universities listed before your undergraduate degree, which is in 1975.

RK: Right. I loved math but a part of me was really a lab guy, which I got from my father. So in college I worked in several labs essentially as a lab technician.

SB: You did some work at the Harvard Medical Unit and Department of Pathology at the University of Geneva in Switzerland.

RK: The Geneva experience was interesting. I actually did my own experiment that I'm proud of, even though it never got published.

SB: OK. But what's your undergraduate degree in?

RK: Math.

SB: Mathematics! Even though you were a math major, you were doing all this lab work.

Sam Behseta is Professor of Mathematics, California State University, Fullerton, 800 North College Blvd., Fullerton, California 92831, USA (e-mail: sbehseta@fullerton.edu).

RK: Yes, it was a different part of me. I was thinking maybe I should go to medical school to do the lab stuff. But I really liked math better. Then I had a fairly a typical experience that led me to statistics. It was the summer before my last year of college and the scientist I worked for had a good-sized lab with maybe ten people working there. Some of them were doing statistics and were completely baffled. They could not figure out what was going on. I picked up one of the books they had. It was Erich Lehmann's elementary book. I read through it and I said, "I can help you." They were so appreciative. It was simple things like t-tests, but they thought I was a genius! So, I realized that might not be a bad thing to do: I could still work with lab people and I could enjoy it vicariously. That's what made me decide to go to graduate school in statistics and, of course, I came full circle in the sense that once I discovered neuroscience I was back to working with lab people.

SB: We'll get to that. But in your undergraduate work you took all these math courses. Did you take real analysis, geometry, algebra?

RK: Yes. Other than number theory I took everything that was offered.

SB: That must have helped later, when you were a graduate student in statistics.

RK: Yes, of course. I almost went back to math because I didn't understand what statistics was. To me, it just looked like math. I didn't get the connection with the real world. Steve Stigler was my probability and statistics teacher. He was great in many ways but in the class he made no real connection with the real world because nobody did. If it was a theory class, we were only doing theory. He did a fine job with the teaching but it was all math. Then I decided I needed a smaller environment because I was used to small schools so I went to Chicago.

SB: Stigler was your teacher at Wisconsin?

RK: Yes.

SB: He then moved to Chicago.

RK: After I moved! He moved when I was in my third year.

PH.D. WORK AT CHICAGO

SB: So you were a Ph.D. student at Chicago when Stigler moved there?

RK: Yes. And when he got there I asked him to be my advisor. I respected him as a teacher and also liked that he brought history into the class, at least a little. I'm not a historian but I always liked trying to understand where the concepts came from. For example, I remember in my first semester we were doing probability and I found a translation of Kolmogorov's monograph from 1933. I read it on the bus every day. It's not a long book, so I finished it within a couple of weeks. It is remarkable how similar it was to what is taught today.

SB: This is the one where he introduces the axioms?

RK: Yes. I also read some of Laplace at that time. And when my theoretical statistics class got to sufficiency, I went back to Halmos and Savage because they did this very nice measure theoretic treatment of sufficiency, which would now be called coordinate-free because it didn't have parameters. A beautiful paper. I was very excited to read that. But it led me back to Fisher. Of course, Stigler had talked about Fisher, at least briefly. This was all during my first and second semesters in graduate school, at Wisconsin. The fact that Stigler was there, and he was a resource for these kinds of conceptual and historical points, was really important to me.

SB: Right. That extra reading must have taken some time.

RK: It did, and I requested an under-load of courses in the Spring semester of my first year, one less course than was taken typically by first-year students. I was pretty adamant about wanting time to myself, and they allowed it. Perhaps Steve Stigler put in a good word for me, because he would have seen that I was a serious student. And I was in the process of transferring, so I would no longer be their problem!

SB: Was the program in Chicago fairly theoretical? I mean, did you have data analysis courses and computer programming?

RK: We totally did. And I eventually figured out that my favorite aspect of math is that in disciplines such as statistics, the math is about something. One of my complaints even today is that courses are too segregated: they're either theoretical or applied, one or the other, and it makes no sense to me. I mean, you have to do some theory to hone your skills so it makes sense to have a bunch of theoretical topics taught together, in sequential chunks. But the chunks don't have to be as long as a whole course. You can in fact interweave data and theory, which would get people more quickly into the frame of mind of a modern statistician.

SB: That's an interesting idea. So Chicago had both theoretical and applied courses?

RK: Back then theory and applications were considered, essentially, two separate tracks. Here's a story about that. Paul Meier was my academic advisor before I had a thesis advisor. I'd gone to Chicago partly because of his reputation and I thought I was going to work with him, but then I learned quickly that I did not like clinical trials.

SB: Important in biostatistics, but not for you.

RK: I was more of a lab guy. Clinical trials are completely different. Anyway, when I graduated, I called him up to ask advice about applying for jobs, and he asked, "Are you a theoretical statistician or an applied statistician?" He knew I was kind of both. There was a pause, and he said, "You have to choose." I said, "Ok, theory!" If I had to choose, my first love was math, so that was my answer. It wasn't until I became a full professor that I finally realized that I hadn't gone into statistics solely because it was a way to keep doing math. It was because it was "about something." I missed the applications. Plus, honestly, I wasn't so great at the theory anyway: I always felt I could understand other people's ideas and results, if I tried hard enough, but I was not especially quick nor as creative as the really good mathematical statisticians.

SB: Before we move on from grad work, what courses did you take at Chicago?

RK: They had this really great thing for Ph.D. students. They gave us a very thorough list of things we were supposed to know for our Ph.D. exams. It was probably five pages long.

SB: A list of topics you needed to know for the Ph.D. exam?

RK: Yes. They had every topic broken into detailed parts and every part had references, specific sections of a handful of well-known textbooks: Feller (1950, 1957), Rao (1973), Cramér (1973), Kendall and Stuart (1977), Ferguson (1967). And we were responsible for all of it. That was great because it told me what to study. So, I would say the curriculum at Chicago was theoretically oriented, but it had major topics that were foundational yet methodologically oriented, like analysis of variance. I wish someone, somewhere, would make a list like that for today's Ph.D. students.

SB: How did you decide on the topic of your thesis?

RK: I started on a problem that didn't end up in the thesis. We had a consulting unit at Chicago where people from across campus would come in and ask questions. One day a post-doc in molecular biology came in. He was examining multiple strains of *E. coli*. They had a theory that said basically two particular genes would act independently. He did standard two-by-two table chi-squared tests of independence in a whole bunch of strains he'd looked at. They were statistically significant, and then he found one strain where the results were not significant. That was what he was looking for, independence. But he came in to ask about it because he said he had paid attention in his statistics class when they said you cannot prove a null hypothesis, you can only reject it.

SB: He understood the logic.

RK: Yes. So he asked me, "What am I supposed to do?" My first thought was, OK, well, let's switch what's the null and what's the alternative hypothesis, and re-calculate a p-value to test a null of non-independence. In order to do that, you have to hypothesize a value for the odds ratio if it is not 1. Then I said well, wait a second, I don't really know exactly how big the odds ratio would be if it were different than 1. Let me put weights on the different possible values. But that amounts to selecting a prior: I was trying to reinvent the Bayesian approach to hypothesis testing!

SB: Did you know Bayesian statistics?

RK: Not well enough to recognize immediately what I was doing. In fact, David Wallace taught a course on Bayesian statistics. He was an inspiring lecturer, but he was very hard to understand. There were a lot of things that he'd do in class where we thought, "This seems so cool" yet we didn't really get it. The whole course ended up sort of like that. But then I realized it had permeated, and what I wanted to do with this data set was pretty similar to what Wallace had done in *The Federalist Papers*.

SB: I see. This was the famous study with Mosteller and Wallace (1964) where they found strong evidence of authorship among papers where historians previously weren't sure?

RK: Yes. They did a great Bayesian analysis. With a bunch of null and non-null data, you can view hypothesis testing as classification. In my case it was a little different but I worked on it and got some results, and showed the results to the scientist. He was happy. He repeated the experiment, which is very important. That is, he repeated the null case he had found and got the same answer. He published a paper on that.

SB: But your work wasn't part of his paper?

RK: No. The method I'd developed, which was an interesting application of hierarchical modeling and what we now usually call Bayes factors, ended up being in the 1989 paper that I published with Duane Steffey (Kass and Steffey, 1989) and I also used it in the 1995 paper I eventually wrote on Bayes factors with Adrian Raftery (Kass and Raftery, 1995). While I was trying to put together a thesis topic, in 1980, I did start thinking about making it more robust. There were things that I could have tried. I went to Stigler and said "I think I am going to make this my thesis, but there is this other thing I found in Jeffreys." I had started reading carefully in the Bayesian world and had gone back to Jeffreys, just as I had gone back to Fisher, previously. There is this thing I found in Jeffreys where he uses differential geometry. That's how he gets his prior. I told Stigler I wasn't sure what that's about, but I had already spent significant time in graduate school, with another student, learning differential geometry. When I told Steve Stigler about my two interests it was just before Spring break. He said, "Over break take a look at the differential geometry thing you're curious about, and when you come back, decide what you want to do." It turned out that a little knowledge of differential geometry and reading Jeffreys got me going on my thesis topic.

SB: I see. Did material from the thesis turn into the *Statistical Science* paper (Kass, 1989) and the book with Paul Vos (Kass and Vos, 1997)?

RK: Yes. The paper was published 9 years after I defended my thesis, and the book was published another 8 years after that! I was not great at finishing things for a long time in my life. At some point, I was at a social event with someone who was laughing about how long I had been, supposedly, working on this book. He said, "You need a co-author." It was like a light bulb went off. I thought about it and I knew who I could get: this guy, Paul Vos, who was four or five years behind me at Chicago. I figured he could probably do the stuff that I was having trouble with. The reason I wasn't finishing it was that I'm a bit of a perfectionist and I wanted to get certain results, yet I didn't know how to prove them. I figured he could do it, and I was right. The idea of having co-authors who could do things I couldn't, or wouldn't, really changed my life.

ON TO ACADEMIA

SB: How did you choose where to go for work?

RK: I had become convinced that the Bayesian approach was really important. I had been exposed not only to David Wallace and Arnold Zellner, but also Dennis Lindley who gave an extended set of lectures at Chicago. I would say it might have been a whole semester. I got to know him a little bit, which was great.

SB: Lindley is an avid Bayesian!

RK: Absolutely. Much more avid than I would be, because for example he thought that p-values were nonsense. To him they didn't make sense but I never felt that way. I don't have a problem with p-values. They may not agree with the Bayesian version, with the Bayes factor, but that doesn't mean they are wrong. I had a more in-between opinion. I thought that if we go out twenty-five years (which would be 2006), Bayes is going to be a big deal. It's going to take a prominent place. But in 1981 there were only two places in the U.S. that were really hospitable to Bayes. One was Wisconsin, because of Box, and one was Carnegie Mellon because of DeGroot. Those were the only places I looked at.

EARLY CHALLENGES AT CMU

SB: When you arrived, the statistics department at CMU was very young.

RK: Very young. Steve Fienberg had just won the COPSS award.

SB: Was everyone doing Bayesian work?

RK: Well, Morrie, Steve and Jay were the core, and Mark Schervish converted to Bayes soon after his arrival. John Lehoczky was doing applied probability. Bill Eddy worked on a lot of different things, with an emphasis on computation and I would say he was agnostic. I would characterize Diane Lambert also as agnostic, but she worked mainly on frequentist problems, and she ended up leaving.

SB: So, you arrived at CMU. Then it's time to publish papers as an academic, which is the non-trivial part of the story always.

RK: Highly non-trivial for me.

SB: Was there pressure on you to produce?

RK: Sure. But I didn't know how to do it. One thing about Chicago is it was this brilliant ivory tower environment, but the faculty there didn't publish much. The most extreme was David Wallace, who hardly published at all. He had his famous book on *The Federalist Papers*, and he had one really famous paper, maybe a couple of papers. Yet everybody knew he's like the smartest guy around. He read constantly. From a student's point of view he seemed to understand everything, and deeply. I would say the big exception was Leo Goodman who was a prolific publisher in categorical data.

SB: Lots of good papers, indeed.

RK: Paul Meier was a pretty good publisher too, in the medical literature. I remember David Wallace once making a remark about publishing in volumes created to honor people, and whether it's worth doing. He said his feeling was, "How many third-rate papers do you have?" There was a kind of snobbishness about it, which got in his way. But once I started taking academic life seriously I realized that if you want to have an impact, you shouldn't keep these ideas to yourself.

SB: You should publish.

RK: Yes. If you want to get other people to see what you see, you have to publish. That was not taught to me as a student at all.

SB: But you had some publications, early on.

RK: Very few, actually. The first real one was a note in *JRSSB* about different ways of parameterizing distributions. That appeared four years after I received my Ph.D. I had also written a short note about the evolution of Jeffreys's philosophy, which was a commentary on "Is Jeffreys a Necessarist?" by Zellner at *The American Statistician* in 1982. And I had another little paper with James Fu in *Statistics and Probability Letters* and something about education that was with several other people. When I came up for promotion to associate professor (at CMU, without tenure) I really had only one small publication in a top journal. To me, it's miraculous that I survived, and it was because I had people who believed in me, which was really great.

SB: They didn't give up on you.

RK: They didn't give up.

A SERIES OF INFLUENTIAL PUBLICATIONS ENSUE

SB: Then I guess you suddenly started to publish. What happened?

RK: That's right. I think most importantly, Luke Tierney had come across Laplace's method for computing posterior expectations and had written a nice paper with Jay Kadane about it. I already knew about it from Jeffreys and also from David Wallace's technical work on *The Federalist Papers*. So I started collaborating with Luke and Jay. Jay was very publication oriented. He was very methodical and product driven. I learned especially from him what that meant. Laplace's method turned out to be a very rich area. We wrote seven papers on that.

SB: In a short span of time.

RK: In a short span of time, yes.

SB: In top journals.

RK: Yes. Those were good papers. That's what solidified my reputation. Before that, people had known I was doing differential geometry. They thought it was kind of interesting and I also think a lot of people felt like I must be able to do some math. Which was sort of true. I could do it up to a certain point, but I made a lot of mistakes in calculations, and sometimes in reasoning. In my favor was that I was pretty persistent about details and, when I read other people's work, I was very capable of finding errors, or flaws of reasoning, or gaps. Anyway, Laplace's method provided a straightforward path, and we published a bunch of papers, plus I finally finished that long overview paper on differential geometry in statistics, the one that appeared in *Statistical Science*, and I got tenure.

SB: That was before MCMC became popular.

RK: Yes. At first it looked like Laplace's method could make Bayes practical on a large scale, but then MCMC came along and quickly took over.

SB: So when did you realize that posterior simulation would be a big deal?

RK: Actually, before MCMC. I had a problem I was interested in that involved a correlation matrix. I wanted to estimate features of the matrix, such as the first eigenvalue and eigenvector (principal component). I was doing it Bayesianly. I realized, because I had read some of the bootstrap stuff, that simulations were very powerful and, furthermore, we could do them from a Bayesian point of view. So I did that, and it worked out pretty nicely. I ended up submitting a paper on it to *Biometrics*. It came back with comments that in retrospect were just revise and resubmit. I was so headstrong that, instead, I started arguing with the referee about some of the things he or she had said. It was a lengthy back and forth. I remember Dan Solomon, the editor, was trying to help me out. He was basically trying to convince me to revise and resubmit.

SB: So you could get it published.

RK: And I didn't do it. That would have predated the Tanner and Wong paper and the Gelfand and Smith paper on Gibbs sampling.

SB: Sounds like you and others were considering simulation at that time.

RK: The interesting thing, to me, is that when people talk about the history of MCMC I think they generally forget to give credit to Efron, who very deliberately and effectively changed how people were thinking. Efron had pushed the idea that simulation was very flexible and could do things in analytically intractable problems. That was his high-level message, and I heard it, as did others in the Bayesian world. After I had written about posterior simulation I did talk about it at a conference. Arnold Zellner was running this Bayesian conference and I talked about it there. He was wildly enthusiastic about it. Luke Tierney said to me, "You know there's a method out there that would do this in general. It's called the Metropolis algorithm. That would be the right way to do it." His comment went right past me. Luke eventually had this really great paper, showing how Gibbs should be considered a special case, and how the theory of Markov chains gives important results. That took a long time. My unpublished technical report was in 1985 and I believe Luke's paper wasn't published until the early nineties (Tierney, 1994).

PERSONAL LIFE AND FAMILY

SB: So you published your research on Laplace's method, and also the long paper on differential geometry, and you got tenure. Did you have a family during this time?

RK: No, I was single. From the time I was hired at CMU until I got tenure I worked very long hours, and it would have been difficult for me at that stage of my life to have divided my time.

SB: But at some point, you were the head of the department with two small children!

RK: That was a lot of work. But here's the thing: when I was single, I was not strategic in managing my time. Pretty much everyone I know, in all walks of life, they get more efficient when they have children. They realize how precious the time is. You can waste a lot of time when you have a lot time to waste! Also, my wife Loreta helped me enormously by being in charge of the household and leaving me second-in-command. What was I thinking about in my spare moments? I wasn't thinking about what we were going to have for dinner the next three days, I was thinking about the latest statistics problem.

SB: That can make a difference. Have you had important insights during those down times? Stepping onto the bus, so to speak?

RK: Absolutely. My theory is that extended concentration on a problem somehow engages powerful processes that can continue even when we are not aware of them. "Incessant contemplation," was Newton's explanation for his success, and there are great accounts of eureka moments, such as when Loewi woke up in the middle of the night with the idea of an experiment, which led to the first discovery of a neurotransmitter, acetylcholine (for which he got the Nobel Prize).

SB: Did you ever have experiences like that?

RK: Yes, occasionally, though obviously not such consequential ones as that. They have been quite startling to me, nonetheless. When fundamental discoveries are made, the feelings they generate are unlikely to be due to greatness of the scientist. Rather, it seems to me, they must have to do with the process. Even those of us with modest capabilities can have our own versions of such experiences and feelings.

SB: That's a nice way to put it. Can you give any examples?

RK: When I was in college, taking real analysis, one night I was thinking very hard, for a long time, about a homework problem. I couldn't get the idea of how to solve it. I went to bed, and in the morning, the instant I woke up, I knew the answer. That was startling enough that I still remember it. Another one I remember well was when, after being frustrated on-and-off for many years, I had an insight about point process modeling of neural synchrony. I remember I was in a hotel room and my mind went back to that problem and suddenly it was obvious! Something that had stumped me became obvious in that moment.

SB: And the key is to keep thinking?

RK: Yes, and here's a related story about me and my wife Loreta, who is a physician. Sometimes at home I would sit quietly, thinking, basically just staring straight ahead, concentrating on some problem. The first time Loreta saw me that way she got very agitated, loudly saying, "Rob! Are you OK? What are you doing?" I turned slowly and looked at her, and said, "I'm thinking." She rolled her eyes and said, "I thought you were having an absence seizure!"

FROM THEORY TO CONCEPTS, AND THE BEST BOOK EVER WRITTEN IN STATISTICS

SB: In your chronology we had gotten up to tenure. How would you describe your work after that?

RK: Over the years I've put a lot of effort into concepts, as opposed to methods. This is what I like about writing review papers. And I've always liked writing commentaries on articles, as well. There are some people who never wanted to do that. They think the real work is in the novel mathematical results. That's fine. But we need many kinds of research. I have put substantial effort into the conceptual side. And, when I do develop a method, typically with co-authors (like you!), try to make sure we really understand its purpose and properties, and that we have explained these as clearly and precisely as we can.

SB: Conceptual understanding of the ideas is also important in teaching. To what extent should we aim for teaching concepts and to what extent methods and theory?

RK: This goes back to the bifurcation of theory and applications. It seems to me that when you are teaching theory, you have to emphasize why we need the theory, which is based on the concepts behind the theory. What are the questions you're trying to answer? What does the theory accomplish? I think very good teachers do this.

SB: But not textbooks.

RK: No. The way books are set up, they tend to be dry, with too little motivation and conceptualization. An interesting contrast is what I consider to be the best book ever written in the field of statistics, namely Feller's *An Introduction to Probability Theory and its Applications*.

SB: Two volumes.

RK: Yes. Feller's is an amazing book because it is rigorous and conceptual at the same time. He has wonderful motivation. He has digressions of a very applied nature, and it really broadens your notion of probability when you see how it can have consequences about things in the real world. There are chunks of his text that are informal and heuristic, but he can switch gears anytime he wants and give a theorem and proof. It's great! I think that this is still a useful model for us in statistics and I'm disappointed that our field has not taken up this model.

SB: You wrote a textbook in 2014, *Analysis of Neural Data* (Kass, Eden and Brown, 2014), with Emery Brown and Uri Eden. What style did you adopt for that?

RK: At first it was strictly applied, but Emery convinced me we had to cover some theory, too, and once I got going it became much more balanced. Still, I tried hard to keep it conceptual. Eventually, after it was finished, I realized that I was, in a sense, all these years later, trying to imitate Feller! Feller, of course was a great mathematician, and our book is relatively elementary, but certainly there's an attempt there to be very conceptual and yet also be rigorous in appropriate places. There are theorems and proofs in this book about neural data, even though it's as easy to follow as we could possibly make it. We talk about things like consistency and efficiency, not because they need to know these things for some progression in their courses, but because if they want to understand how statistics works, they need to know that there's theory that backs up the methods. So there's a lot of discussion of these concepts and then we show how you make it rigorous. That's the model that Feller really exemplified.

SB: It's interesting because you can contrast his book with Doob's book and I think they were contemporaries. But Doob's is very rigid and it's hard to decode.

RK: Well, there is a whole tradition in math of trying to make things as elegant and concise as possible, and that's I think a good thing except it has the unfortunate consequence in almost every implementation of eliminating the motivation and the ideas behind the way people thought of these things. Actually, another favorite math book of mine, Spivak's *Calculus on Manifolds*, is both conceptual and mathematically elegant. It shows how a really important development, namely Stokes' theorem in physics, can be made at once more general, more rigorous, and simpler.

EDITING STATISTICAL SCIENCE

SB: In 1992 you became the editor of *Statistical Science*!

RK: Right. Morrie DeGroot had started the journal with Ingram Olkin. I have a feeling that Ingram was really pushing hard for it and Morrie was maybe one of the best people that he could think of to edit it. Morrie was very well respected, but he cared about the conceptual side of the field, and he also wanted to make it fun. Morrie was a kind of fun loving guy. He wanted to have interviews and all this. So it immediately became one of the most widely read journals in the field. From my perspective, there was nowhere better you could publish a paper than *Statistical Science*.

SB: In your first issue you wrote an editorial about the importance of review papers.

RK: Yes. Sometimes a review is mostly novel, meaning people are putting things together in a way that no one had put them together before, and that's when it's most exciting. Or when they open up an area for research that no one before had realized is an interesting area. Of course, they can be more mundane. They can just be a summary of literature. I don't think those are the best ones. I think synthesis of literature is important, but a really good review does a lot more: it can evaluate the ideas, and identify the most interesting ones, and it can not only say here's all the things people have done, but also put everything in perspective in a way that guides the field. I was really thrilled to take that on.

BIBLE CODE CONTROVERSY

SB: The Bible code controversy happened while you were the editor of *Statistical Science*. Why did you publish this paper?

RK: I got the paper from Herman Chernoff. It had been submitted to the *Proceedings of the National*

Academy of Science. And Herman felt like it was totally not appropriate for that but maybe it was appropriate for *Statistical Science*. And this journal, you should remember, had already published Jessica Utts' article on ESP. People enjoyed reading it and seeing the controversy. *Statistical Science* was always supposed to be open-minded in a certain sense. In brief, the authors took what they call equidistant letter sequences in the book of Genesis, to spell names and birthdays of famous rabbis who were not alive at the time the bible was written. They found matches in close proximity, within the text, between the name and the birth date, as I recall. They had a metric, they devised a permutation test, and they rejected a null hypothesis, because they got p-values that were like ten to the minus six. And of course the implication was that this was divine prediction! I sent the paper out for review and the reviewers immediately wrote back that if the authors do the things reviewers suggested, it will be clear that the findings are an artifact. But when the authors did those things the effect persisted.

SB: What did your editorial board say?

RK: I had a good editorial board. One of the editors on the board very clearly said you shouldn't publish this no matter what happens because it will be viewed as affirmation for what these crazy guys believe. I felt like that wasn't enough of a reason to shoot it down. So, in the end I published it with a note in the front of the journal saying that this is a challenging puzzle for the readers to figure out. It took several years before finally a computer scientist and some statisticians in Israel got to the bottom of it and figured out that essentially what had happened was the authors had tuned their method to the data by allowing it to use different forms of the names and the dates written in different ways in Hebrew. And there were actually many combinations of such things. Especially when you go through all the one hundred and fifty names of the rabbis across this list.

SB: Did you get any negative reaction from the statistics community from publishing the article?

RK: No, I never did. There was a very interesting podcast about this a couple of years ago (http://israelstory.org/en/episode/28-on-the-outs/). Let me add that I have a purist view of our discipline. I actually think that statistics is a higher calling, in a way. What we do in statistics can, in some ways, be more important than what the rest of the world is doing. We are in it for the long haul. And for us to make progress we have to examine these cases and we have to see how they play out.

SB: Let me stop you there! You just said we have a higher calling in statistics. Could you elaborate?

RK: What I meant by that is things like the Bible Codes might make certain people believe things they wouldn't otherwise believe. They'll have some effect in convincing the gullible that there are strange things in the world that are not scientifically explained. On the other hand, I would like to believe that civilization will continue. Statistics is a slowly evolving discipline. We try to get things right. It will be with us in the long-term future and anything you do that improves our understanding of statistics and our ability to communicate statistics will have very long-term value. In that sense, it's a higher cause. We are in it for a much longer period than the splash effects of the applications of statistics.

BAYES FACTORS AND RULES FOR SELECTING PRIORS

SB: Recently you added something to your website that talks about how you view your contributions in different categories, one of which being review papers.

RK: Right. When I took over *Statistical Science*, I had published one review, which was the differential geometry one. As I think I said, I spent a lot of time on that. And it had a lot of different parts to it that I felt involved novel synthesis of what's out there, putting things together in a new way. Plus, it contained ideas that led to other papers. I had also already started on two review papers before I took on *Statistical Science*. One was with Adrian Raftery. And the other was with Larry Wassermann.

SB: Your review with Raftery became very highly cited.

RK: Yes.

SB: How did you know him?

RK: I didn't know him! What happened was, I read a paper by Jim Berger and Mohan Delampady on the Bayesian approach to hypothesis testing (Berger and Delampady, 1987). As I said, back then it seemed like some of the most influential articles were in *Statistical Science*, and that one certainly got a lot of attention. The emphasis that Jim was taking, there and elsewhere (Berger and Selke, 1987), went back to a paper by Edwards, Lindman and Savage in 1963 on the relationship of p-values to posterior probabilities for hypotheses. Savage showed there that t-statistic p-values, at traditional levels like 0.05, were not strong evidence against the null when you calibrate them by Bayes theorem. Jeffreys had similar calculations in his book. Jim and his colleagues nicely generalized the fundamental story, but when I read their papers I remember having this visceral reaction: I felt they were missing Jeffreys's main point, which was that Bayes factors could be useful. They were saying p-values and Bayes factors give different answers, but there is a place where Jeffreys says, "Fisher and I rarely disagree on the actual outcome of an analysis, and when we do, it is probably because of a failure of assumptions". I wanted, instead, to develop further the idea that Bayes factors could be useful, and modernize it. Jeffreys wrote so long ago, before computers. I wanted to see how Bayes factors played out in contemporary practice.

SB: And Raftery?

RK: I wrote to him because he was the only person who was doing practical problems with Bayes factors. We started in 1988. The paper was published in 1995 and that's how long it took. We went through a lot of revisions to that paper ourselves, and with reviewers, and worked up all these case studies, which I think was a rare thing to do in statistics. We had five small case studies in there of real problems.

SB: Have you used Bayes factors a lot in your applied work?

RK: No. The main place where Bayes factors are likely to be of use is in getting evidence in favor of the null hypothesis. There is no good way to do that without Bayes factors. But early on, just after my Ph.D., when I was more favorably inclined toward Bayes factors, I worked a multidimensional problem with Jeffreys's geometrical method, and got a Bayes factor giving astronomical odds on the null, which didn't make sense. I realized that when you go to multidimensional problems, it's really hard to come with some good useful prior probability on the alternative. That's the challenging part of Bayes factors, getting a prior under the alternative. You don't usually have data for that, and I came away realizing that unless you have data that are relevant to the alternative, the Bayes factor is going to be somewhat arbitrary. In all my years doing neuroscience applications, I've only applied them a couple of times.

SB: So when you were working on the Bayes factor paper, were you anticipating it becoming such a popular highly cited paper?

RK: No! The citations came as a complete shock.

SB When the Institute for Scientific Information came out with their top ten list of most-cited authors in mathematics, you were number four!

RK: I felt like I won the lottery. It still feels that way because it seems so random. It sounds very impressive, especially to people outside the field, yet, as statisticians we can appreciate that it's pretty arbitrary: high citation counts depend on many things, and should not be considered indicators of importance. Still, the citations do indicate that people are reading your work, and feeling that, for whatever reason, they need to cite it. There are certainly things to learn from citation patterns. I sometimes look at my Google Scholar cites to see which papers are cited a lot. It's interesting to try to speculate why some and not others. And some continue to be highly cited for a long time. Our Bayes factors paper has had over 3000 citations in the past three years.

SB: More than 20 years after it was published.

RK: Yes. That's an interesting fact. And I'm now very conscious that some papers I'm working on may have a long shelf life, and should be written that way.

SB: Then there is the paper on selecting prior probability distributions, with Larry Wasserman (Kass and Wasserman, 1996).

RK: Yes. That also has had a pretty long shelf life. And it's also a good example of work I never would have done without my co-author. I had thought about writing that paper years earlier, but it was a huge effort, and Larry was really great at reading things quickly and summarizing the main ideas. We distilled and categorized the ideas, and in the end we concluded that there's not going to be any one compelling rule for selecting priors. It's never going to work. That's what the literature tells us. And I came out of it thinking that it's the asymptotics that carry the day, which is the same as what Jeffreys said: the Bayesian framework is great for thinking about problems, but Bayesian estimation is essentially the same as maximum likelihood. In practice, my collaborators and I use Bayesian methods for the sake of convenience, not principle. And if you're not in a large sample situation, it's problematic. You don't want to be relying too much on the prior!

FROM CASE STUDIES TO NEUROSCIENCE

SB: During the time you were concentrating on Bayesian methods, you also ran a series of international meetings called Case Studies in Bayesian Statistics. What was its purpose?

RK: The best articulation of the purpose of case studies comes from Mosteller and Wallace's book on *The Federalist Papers*. To paraphrase them, case studies can be a great way to learn about statistics, but they should be "real" in the sense that the statisticians have to be committed not only to the statistical methods but also to the substantive conclusions. That's the only way we can see how statistical methods really work in practice. And commitment to conclusions is the big distinction between examples in a paper, and case studies.

SB: They were good meetings. How many were there?

RK: We had eight, and they were excellent meetings, but they never succeeded in the big goal of informing us about the way the art of statistics-in-practice can become more of a science. The problem was that in each meeting we had several case studies, but they were all on different subjects. In fact, I think we've succeeded better in the other series of conferences I've run, in neuroscience.

SB: Statistical Analysis of Neural Data (SAND)? The ones you've organized with Emery Brown.

RK: Yes.

SB: When did they start?

RK: 2002. They are every other year, with one gap year. In 2019 we will have SAND9. They've been terrific meetings partly because they are focused within one domain, so the audience is much better equipped to ask pertinent questions. And the field has evolved dramatically, from the most basic problem definition in 2002, to state-of-the-art statistics and machine learning in the most recent one, in 2017. We can help advance neuroscience research, but it's also a great way to better understand statistical methods.

SB: When did you first get involved in neuroscience?

RK: I joined the Center for Neural Basis of Cognition (CNBC) in 1997. As you know, that's a wonderful, large umbrella organization, across many departments at both Carnegie Mellon and the University of Pittsburgh.

SB: Was that when the CNBC started?

RK: No, it was started by Jay McClelland (one of the major figures in neural network modeling in psychology) in 1995, so I joined 2 years later.

SB: It was pretty new.

RK: Yes. What happened was, as soon as I became a full professor I realized I was missing applications and found neuroscience. It was very slow at first. I was Department Head at the time, and couldn't devote serious effort to it. Fortunately I again had a wonderful collaborator, Valérie Ventura, in our department at CMU. She did a lot of the detailed statistical work, and originated many of the data analytic ideas. We wrote many papers together, and still collaborate. But also, in 1998, Emery Brown and I agreed to work on a review paper on statistical methods in neurophysiology.

SB: You wrote papers on neuroscience with Emery, including the review paper, but also a paper about statistical pedagogy.

RK: Yes, we had many conversations. The first thing we did was we thought we should write a review paper on statistical methods in neurophysiology. That was in 1998. The paper did not appear until 2005! (Kass, Ventura and Brown, 2005). But the reason it took so long in this case was that the methods we thought we should be reviewing didn't exist. So we kind of had to create them, and we each pretty much separately published on stuff, and then we could kind of write a review of what we were doing. Emery and I had a lot of conversations about how statistics was not being used to its full capability in neurophysiology, which was the motivation for us to write this review, and we did a few other things together, including the book, eventually. But at the same time, he was looking around and saying, "Why aren't there more statisticians in neurophysiology?" And part of his complaint was that statisticians would not jump in. They're very reticent compared to engineers and physicists who are amazingly brave. They'll go right into a lab and do the recordings and are willing to learn new stuff on their own. It's rare that a statistician would do anything like that.

SB: I can't think of any student who did that.

RK: This was his fundamental concern. And I shared it, but with a slightly different perspective because it's easier to follow a middle road, in which a student could get solid statistical training, and remain a collaborative statistician without actually doing the experiments, while taking the science really seriously. At any rate, our point was the whole process would be a lot more efficient if you introduced students to real-life problems in graduate school, and it would be more connected to academic research as well.

SB: Have you tried to bring some of those ideas to the graduate program at CMU?

RK: Our department has been unusually good at this, especially because of our Advanced Data Analysis project course, which is the single most important thing that we've done.

SB: I totally agree. As a former CMU student I can attest to that. You work on a real-life data project, and experience the whole gamut of statistical responsibilities from formatting data to presenting it to the entire department, and writing it up as a paper.

RK: I think something like this should be part of every Ph.D. program in statistics.

SB: We jumped ahead a little. Let's come back to your experiences in neuroscience. You worked a lot with Andy Schwartz, right?

RK: Yes. For many years, Andy was my main collaborator, starting when I met him while he was being recruited to Pittsburgh. Andy is a primate motor physiologist, studying hand movement, and he's become one of the best people in the world at brain-computer interface (BCI), where the subject moves a cursor on a screen, or a robot arm, just by thinking. His was the first lab to get a monkey to feed itself using signals recorded from electrodes in the brain. Recently there have been a number of successful studies on humans, who are quadriplegic, meaning unable to move their own arms and legs, due to accident or illness.

SB: Fascinating.

RK: Yes, and there are some great videos on the web, including President Obama fist-bumping with a quadriplegic patient, here in Pittsburgh. We did a number of basic studies to help develop the technology. The biggest innovation was pretty simple. Emery had shown how state-space models, similar to the Kalman filter, could be used to predict where a rat would run based on signals recorded from the rat's brain (from the hippocampus). Our first step was to use this approach for getting the monkey's intended hand movement. We tested (and published) a number of variations on this theme. The use of state-space models for BCI is now pretty standard.

SB: How much physiology did you need to know?

RK: Some basics, though I also learned much more general background material and that turned out to be super-useful as I talked with lots of other neuroscientists, who were doing very different things. I read things, and I sat in on two classes. But, as I said, while I was department head things went slowly, especially because it's such a different field, and the knowledge base is so vast it can be hard to know what to prioritize.

SB: Did things pick up when you stepped down from being department head?

RK: Yes. My biggest advance in knowledge came immediately after, when I took a year-long sabbatical devoted to neuroscience.

SB: When was that?

RK: It was the 2004–2005 academic year. Actually, I don't think I mentioned it, but a really big year for me previously had been 1987–1988, when I also was on sabbatical. You can see it in my 1989 publications. In this second sabbatical I read a lot, and visited several labs across the country, just to learn about what they were doing. Then, during the spring, I wrote a grant proposal that marked a kind of turning point for me, where I really became more seriously committed to the science.

SB: What happened?

RK: The "eureka" moment occurred during a conversation I was having with an excellent neurophysiologist, David Redish, at the University of Minnesota. He was unusual because he had been a Ph.D. student in computer science at CMU, but learned how to do experiments.

SB: He was one of those fearless non-statisticians Emery talked about.

RK: Yes, and he became a very well known experimentalist. So I was excitedly telling him the kinds of statistical analyses I thought we could do, with the right data, and he said to me, "So what experiment should I do?" The question stopped me cold.

SB: It wasn't what you'd been thinking about?

RK: No. A couple others had asked me this question, too, and I had kind of shrugged it off as not part of my job, but for some reason, this time I realized, oh yeah, that's what I should be thinking about: experiments. How to get the data I'd like to analyze.

SB: So you wrote a grant proposal about getting data?

RK: Yes, exactly. With help from Andy, I designed a set of experiments. But they were designed to tap into the statistical methods I wanted to develop, which of course was also part of the proposal. A neurophysiologist tries to find a scientific question they can answer using a particular experimental setup, a setup they are already using or perhaps something they haven't used but feel they can develop. Then they can design their experiments. I wanted to find a scientific question that would take advantage of powerful statistical methods, and work from there.

SB: How did you do that?

RK: Powerful statistical methods can find effects with much smaller data sets. My collaborators were running experiments with dozens, sometimes hundreds of repetitions, known as trials, and a lot of the analyses were based on some kind of averaging of neural responses across trials. Of course, averaging ignores variation across trials. But I figured that with better statistical methods we could see any trends that might be occurring across trials. So I asked myself, and this was the new step, why might we care about such trends?

SB: You mean it was new scientifically?

RK: No, but it was a new kind of question for me, as a statistician, to ask. I designed a set of experiments that became the core idea in a grant proposal, and it got a superb score and got funded.

SB: What happened in those experiments?

RK: Well, they never got run in the form I designed! Instead, a postdoc in the lab introduced a variation on the same idea which turned out to be very successful, leading to a paper in the *Proceedings of the National Academy of Science* (Jarosiewicz et al., 2008), and then a series of further papers.

SB: This is a major contribution statistics can make, bringing knowledge of data analysis into experimental design.

RK: Exactly.

STATISTICAL MODELING OF SPIKE TRAIN DATA

SB: I know a lot of your work on neural data has involved point processes.

RK: Yes, the large majority of it, though only a little of the motor physiology and BCI work has used point processes.

SB: When you started, what drew you to point process models?

RK: Well, a neural spike train is a sequence of times at which a neuron fired, and the times are usually pretty irregular, so point processes are an obvious possibility, but when I started I didn't know the first thing about point processes! I knew that spike trains were not Poisson processes but I didn't know how to deal with that.

SB: How did you know that they are not Poisson processes?

RK: The first time we analyzed the data it was apparent that there were small but clear deviations from Poisson. Plus, I was aware of some theoretical arguments saying they shouldn't be Poisson. My colleague Emery Brown strongly felt that it should not be Poisson for theoretical reasons. Valerie Ventura and I used what seemed to be the simplest possible non-Poisson model, and all of our subsequent work has descended from that. If you look at our 2001 paper, it's really simple. But I have to tell you, and it's a little embarrassing, I was very confused about how to set things up for quite a while.

SB: You talked about point processes in your Fisher lecture https://www.youtube.com/watch?v=P_KV1ONiLE4.

RK: Yes, I mentioned that there were interesting complexities, and I also talked about the synchrony problem.

SB: What is that?

RK: Synchrony refers to the phenomenon that two or more neurons may fire nearly at the same time, more often than would be predicted by chance alone. When we build point process models, we get for each neuron the probability of firing in any particular small time bin and, assuming independence (conditionally on relevant variables that are included in the model), we can then get the probability that two (or more) neurons will fire. In my Fisher lecture I mentioned that we had done a lot of work on this, but I instead described some work by other statisticians, which illustrated beautifully the basic point that statistical principles matter. In this case, the statisticians' method differs in a subtle way from a method devised by physicists, and the statisticians' method has compelling properties that the physicists' method doesn't share.

SB: But your work went in a different direction.

RK: Valerie and I had a couple of papers, with one of her students, on synchrony for two neurons. We developed simple methods for detecting excess synchronous firing based on GLM type point process representations. Right away, however, I became puzzled because I didn't know any theoretical foundation for what we were doing, in terms of point process theory.

SB: What was the issue?

RK: The theoretical foundation was based on going from data in time bins (where we can use binomial or Poisson GLMs, for example) to point processes, which are in continuous time. For single neurons it's easy to show that the point process PDF is the limit of the discrete PDFs, as the width of the time bins goes to zero. But a standard point process regularity condition is that there can be at most 1 event at any point in time. So, in continuous time, how do we understand synchrony, where 2 or more neurons fire together?

SB: Sounds challenging.

RK: Everyone I asked about this said, "Use marked processes," the idea being that you can mark every spike as coming from neuron 1 or neuron 2 or both. However, that by itself doesn't solve the problem. I was stuck for several years until I had that flash of insight in the hotel room I mentioned earlier. And it's so simple! In a point process model for a spike train, such as a Poisson process, the probability of a spike in a small time bin is approximately proportional to the length of the bin. So if you have 2 independent processes, the probability that both will spike is approximately proportional to the square of the length of the bin. My sudden realization was that this condition was typically not satisfied in marked processes, so we had to modify the processes, forcing them to satisfy that condition. Once we did that, we were able to get a theoretical foundation that made sense. We published this, with some interesting data analysis, in the *Annals of Applied Statistics* in 2011.

SB: And you won a "best applied paper" prize for that, from ASA.

RK: Yes, but I like the story especially because it shows how it can be worth thinking about problems many, many times. I can't tell you how many times I thought about it before that basic property popped out as crucial to making it work.

MACHINE LEARNING AND ITS RELATIONSHIP WITH STATISTICAL RESEARCH

SB: As you were working on the neuroscience problems, computation had advanced significantly, and of course we are at CMU, where machine learning is very strong and you are affiliated with the Machine Learning Department yourself. How would you say this computational revolution has affected what we do as statisticians, and in brain sciences in particular?

RK: Well in multiple ways. I think honestly computational issues have been part of statistics for a long time. You can certainly go back to the time of Fisher and Yates. You know Yates' algorithm was the precursor to the Fast Fourier Transform and that would have been in the 1920s.

SB: Right. Frank Yates was doing design of experiments back then.

RK: Even before desktop computers, computation was important in statistics. In the 1970s and 1980s there were statisticians who worked on things like optimization and non-linear least squares, developing algorithms to make computation stable and reasonably fast. Computation was certainly part of statistics when I was a student. The things that are different now are the dramatic increase in computational speed and storage capacity and, of course, the World Wide Web. There's tons of available data and we have the ability to answer much more complicated questions. That is the heart of the big data revolution. In science, it's often less about the size of the data set, and much more about specificity (or "granularity") of the questions and, thus, the complexity of the problems. And absolutely in the brain sciences we have complicated problems. The challenge has mostly to do with the way we should conceptualize problems, re-define them, so we can simplify and get a handle on the phenomena that are driving the data.

SB: In your Fisher lecture you said the brain sciences are underserved by statistics.

RK: We could have dozens and dozens of statisticians go into this area and I think it still would be underserved. To anybody who might come across this interview and be interested in getting involved, I'd say there's a place for you. There is a lot to learn, and it's hard because, at least at first, it is not obvious what the outstanding problems are that you can really help with. But there is a demand (being filled currently by physicists and engineers) and in the rest of my working years I'm making it a priority to try to figure out what I can do to help all these people who might want to bring more statistical thinking into the brain sciences.

SB: Machine learning has also affected how we do or at least view statistics. I mean if you teach regression these days, you also include something about cross-validation.

RK: Yes, but cross validation goes back to Mosteller and Tukey (1968). They are the ones who pushed cross validation. They talked about k-fold cross validation and cross validation in general. It's true that in the old days everybody was stuck with very small data sets and you kind of had to be parametric. You couldn't fit something flexible if you only had thirty observations. So yes, the world has changed. There's no question. I just think that conceptually I don't see it as revolutionary. I mean, I lived through it, so I see how it all progressed and it seems fairly incremental to me and it took a very natural progression.

SB: That's very interesting.

RK: As you know, we have the only academic department that's devoted to machine learning, and it was started by Tom Mitchell in computer science and Bill Eddy and Steve Fienberg in statistics. I have a favorite story about this from the first time we had a retreat to try to figure out this new machine learning department. One of my colleagues in computer science, Roni Rosenfeld (who is now the head of that department) said, "I have finally figured out what the difference is between statisticians and computer scientists: statisticians try to solve problems with ten parameters and get it right, computer scientists try to solve problems with ten million parameters and get an answer!" And what I like to say is now it's evolved so that we're trying to solve the really big data problems and get it right. And in fact machine learning, you could argue, is the intersection between statistics and computer science. There's essentially nothing in machine learning that is not statistics. There's also nothing in machine learning that's not computer science. It really is the intersection. The one thing I don't like is the world at large has this notion that if you just sprinkle the magic machine learning dust on your data, wonderful things will happen.

SB: So how should the statistics community respond to that or react to that?

FIG. 1. *Rob with Paul Meier, circa* 1999.

FIG. 2. *Rob with Loreta*, 2007.

RK: Constructively. Statistics has been incredibly successful. When it first started, machine learning wasn't statistics, but now it is. I should add, though, that there is often a difference in perspective: computer science comes mostly from engineering, while statistics comes mostly from math. So the computer scientists tend to want to do stuff, build stuff. This is at the root of Roni's characterization. Right now, for example, deep learning involves turning knobs in some clever way to make things work well. I'm not very fond of it because, I mean, you know, I have this favorite quote from David Blackwell which I paraphrase as, "I never really wanted to do research. I just wanted to understand, and sometimes in order to understand you have to do research." (DeGroot, 1986). Statisticians want to understand. I think with deep learning we will get there. But we're a ways away from that right now. Deep learning is interesting because it works. So you kind of have to pay attention.

FIG. 3. *Rob with Sam Behseta in his office*, 2017.

ON THE VALUE OF APPRECIATING UNCERTAINTY IN SCIENCE

SB: In your Fisher lecture you also talked about the need to better educate our citizens.

RK: We in statistics are painfully aware that most people have great trouble appreciating variation and grappling with uncertainty. One conspicuous result is that discussions of science aimed at non-scientists are unable to include the kind of qualification and nuance that is routine within science, whether through statistical methods or informal description at the end of a scientific paper. There's tons of this stuff in the brain sciences of course. People love to say here's how our brain works; here's what's going on. But it's usually overreach. The interpretation is not justified by the data. It's speculative and it's passed off as factual. It gives people a very distorted sense of how science progresses and what we might mean by factual. In science, we have a loose yet pragmatic notion of "fact." This, unfortunately, feeds into our current arguments over what's fact and what's not fact.

SB: But science does have facts.

RK: The most important feature of experimental science is that results should be replicable. Interpretations are also crucial, but we know that interpretations are not the same as the data results themselves. And there's always uncertainty.

SB: This is what you were getting at in your Fisher lecture.

RK: Yes. I ended the Fisher lecture with a discussion of what Fisher said. He was concerned about the rise of totalitarianism around the world where the process of science seemed to be drastically undervalued and underappreciated by the citizens. So I used a quote from Fisher to repeat my fundamental point that one of the most important things we could do is get students, all of them, to see the variation in the world, and the way variation creates uncertainty. It has the consequence that we all have to be humble about what we claim because there's always uncertainty there. At the same time, we in statistics also have to emphasize the importance of principles. We should be teaching elementary statistics students that there is value in theory. We can show them the value of theorems without stating them precisely, and without showing the students proofs. There are things we can do that give them the idea that one method can be better than another method in a certain situation. The high-level point is that even though principles apply imperfectly to the real world, the process of combining principles with empirical data is the key to progress. That's what we have to get across to our fellow citizens.

SB: Partly we haven't communicated well with the broader audience. Not just in the classroom. I can't think of many statisticians who are good public speakers and show up frequently on TV to comment on events from a statistical point-of-view.

RK: True. I think, ultimately, though, it's about education more than anything else. In the U.S., it has to get into the middle schools and high schools. We have to figure out how to get everyone to appreciate that we can make progress in the face of uncertainty.

SB: That's a good way to end this. Thank you, Rob.

RK: Thank you!

REFERENCES

BERGER, J. O. and DELAMPADY, M. (1987). Testing precise hypotheses. *Statist. Sci.* **3** 317–352. MR0920141

BERGER, J. O. and SELKE, T. (1987). Testing a point null hypothesis: irreconcilability of P values and evidence. With comments and a rejoinder by the authors. *J. Amer. Statist. Assoc.* **82** 112–139. MR0883340

CRAMÉR, H. (1973). *The Elements of Probability Theory and Some of Its Applications*. Wiley, New York. MR0067379

DEGROOT, M. H. (1986). A conversation with David Blackwell. *Statist. Sci.* **1** 40–53. MR0833274

FELLER, W. (1950). *An Introduction to Probability Theory and Its Applications. Vol. I.* Wiley, New York, NY. MR0038583

FELLER, W. (1957). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. Wiley, New York. MR0270403

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach. Probability and Mathematical Statistics, Vol.* 1. Academic Press, New York. MR0215390

JAROSIEWICZ, B. CHASE, S. M. FRASER, G. W. VELLISTE, M. KASS, R. E. and SCHWARTZ, A. B. (2008). Functional network reorganization during learning in a brain-machine interface paradigm. *Proc. Natl. Acad. Sci. USA* **105** 19486–19491.

KASS, R. E. (1989). The geometry of asymptotic inference. *Statist. Sci.* **4** 188–234. With comments and a rejoinder by the author. MR1015274

KASS, R. E., EDEN, U. T. and BROWN, E. N. (2014). *Analysis of Neural Data. Springer Series in Statistics*. Springer, New York. MR3244261

KASS, R. E., KELLY, R. C. and LOH, W.-L. (2011). Assessment of synchrony in multiple neural spike trains using loglinear point process models. *Ann. Appl. Stat.* **5** 1262–1292. MR2849774

KASS, R. E. and RAFTERY, A. E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795. MR3363402

KASS, R. E. and STEFFEY, D. (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *J. Amer. Statist. Assoc.* **84** 717–726. MR1132587

KASS, R., VENTURA, V. and BROWN, E. N. (2005). Statistical issues in the analysis of neuronal data. *J. Neurophysiol.* **1** 8–25.

KASS, R. E. and VOS, P. W. (1997). *Geometrical Foundations of Asymptotic Inference. Wiley Series in Probability and Statistics: Probability and Statistics*. Wiley, New York. MR1461540

KASS, R. E. and WASSERMAN, L. A. (1996). The selection of prior distributions by formal rules. *J. Amer. Statist. Assoc.* **91** 1343–1370.

KASS, R. E., CAFFO, B. S., DAVIDIAN, M., MENG, X.-L., YU, B. and REID, N. (2016). Ten simple rules for effective statistical practice. *PLoS Comput. Biol.* **12** e1004961.

KENDALL, M. and STUART, A. (1977). *The Advanced Theory of Statistics: Distribution Theory. Vol.* 1, 4th ed. Macmillan Publishing, New York. MR0467977

MOSTELLER, F. and TUKEY, J. (1968). Data analysis, including statistics. In *Handbook of Social Psychology*, 2nd ed. (G. Lindzey and E. Aronson, eds.) **2**. Wiley, New York.

MOSTELLER, F. and WALLACE, D. L. (1964). *Inference and Disputed Authorship: The Federalist*. Addison-Wesley, Reading, MA. MR0175668

RAFTERY, A. (2001). *Statistics in the Twenty First Century*, 1st ed. CRC Press, New York.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York. MR0346957

SELLKE, T., BAYARRI, M. J. and BERGER, J. O. (2001). Calibration of p values for testing precise null hypotheses. *Amer. Statist.* **55** 62–71. MR1818723

TIERNEY, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1701–1762. MR1329166