Rejoinder

Glenn Shafer

I am honored that *Statistical Science* has published my inaugural lecture, together with thoughtful comments from a very distinguished group of readers. These readers include Ian Hacking, who is probably the most prominent living philosopher of statistics, and six prominent members of leading statistics departments. I want to thank the editor, Carl Morris, for recruiting these readers, and I want to thank them for the thought and care they have devoted to reading and discussing my lecture.

My chair in the School of Business at the University of Kansas was endowed by Ronald G. Harper, a Tulsa businessman who designs and sells expert systems that combine statistical and artificial intelligence methods to help retailers locate stores, control inventory and set prices. In appointing me to this chair, the university was recognizing my work on probability in expert systems. My first thought, when I was asked to acknowledge the appointment by giving a lecture to a general audience, was that I should discuss the relation between probability and statistics on the one hand and artificial intelligence on the other. I soon realized, however, that my audience knew so little about these topics that I could not cover them all. Much of the audience would not even be aware that some universities have whole departments devoted to statistics. So I left out artificial intelligence. I talked about the intellectual history of probability and statistics and the why and how of statistics departments. I closed with some suggestions for reinvigorating statistics departments.

It takes some chutzpah for a business professor from the provinces, who has not taught in a statistics department for 14 years, to give prescriptions for reforming statistics, and I think it is a tribute to the openness of our discipline that my ideas have been taken seriously by prominent statisticians.

In their comments, the statisticians have focused on the role of applied statistics. These comments seem very important to me, for applied statistics is crucial to the future of our discipline, and the discussants have the stature to help shape this future. In this rejoinder, I hope to raise some questions that will encourage yet further discussion.

I will begin by responding to Ian Hacking's critique of my sketch of the history and philosophy of probability. Then I will turn to the central questions we should ask about applied statistics: What is its intellectual content, and how can we raise its perceived stature *vis-à-vis* the mathematics of statistics?

THE HISTORY AND PHILOSOPHY OF PROBABILITY AND STATISTICS

Ian Hacking's work on the philosophy and history of probability and statistics has been an inspiration to me since my graduate work in the early 1970s, and I have prized his personal encouragement of my own work in these fields. I am honored, therefore, by his serious and extended response to my lecture.

Hacking quite rightly takes me to task for my simplifications of history. It is true that we cannot compress the rise of frequentism into the years 1842 and 1843, that positivism was not a simple phenomenon and that hints of frequentism can already be found in Laplace and even more clearly in Poisson. He has passed over in silence some other equally egregious simplifications. I should mention in particular my simplified account of David Hilbert's ideas on the foundations of mathematics, for several people have protested to me privately that this account comes closer to popular caricature than careful analysis. (For careful analysis, see Benacerraf and Putnam, 1983.)

Hacking's most important point is that the rise of frequentism had complex roots, going beyond positivism. He sees the early nineteenth century's emphasis on measurement as something broader than positivism, and he points to the proliferation of published statistics as a source of frequentism. I find myself only half convinced by his arguments on these points. Why did people in the nineteenth century want to measure frequencies instead of beliefs? And why did statistics on suicides seem stable? The Ian Hacking of 15 years ago would have seen here the influence of theory on desires and perceptions and, intellectual fashion notwithstanding, I think that theory is still part of the story. A complete understanding of the rise of frequentism must take into account the mathematical possibilities of probability theory and the internal logic of the mathematician's drive to apply this theory. Bernoulli's motivation for the law of large numbers was his desire to apply probability theory to social and economic problems, and this same motivation encouraged nineteenth-century frequentism. Probability theory seems to have a very narrow scope of application if every probability must be both a frequency and a belief, as the probabilities in fair games are. It has a much wider scope of application if every frequency is a probability.

I agree with Hacking that probability is similar to many other concepts, in that it has unified prototypes that generalize in different directions. But when so much effort has been devoted to denying either the subjective or objective side of probability, it seems Polyannaish to say that the concept has always been unified.

The desire to apply probability widely still impels people to narrow its definition. Just as nineteenth-century frequentists saw opportunities to apply probability in situations where there are frequencies but no beliefs, twentieth-century subjectivists see opportunities in situations where there are reasons to believe but no frequencies. I point out the decline of positivism not simply to encourage my readers to send frequentism down with it. Rather, I am arguing that this decline permits us to escape from the idea that broadening the application of a concept requires us to narrow its meaning. We can now take a broader view of the relation between theory and application. We can describe a theory itself in a more unified way and see its diversity as a diversity of uses.

George Box's views about distinct ways of using probability in statistical inference fit here. In fact, his views have influenced my formulation. But there are more than Box's two ways of using probability. I listed several in my lecture.

The standard axioms for mathematical probability, formulated by Kolmogorov, do not presuppose any collection of data or structure of repetition; instead they appear to deal with the probability of a single event. As Hirotugu Akaike points out, this makes mathematical probability an awkward foundation for statistics. Yet Kolmogorov's axioms are not faithful, in this respect, to the original thinking of Pascal, Fermat and Huygens. My proposal for reunifying probability theoretically (or emphasizing its unity, if you prefer) involves bringing repetition back into the axiomatic framework. As David Moore notes, I have not given enough detail about the axiomatization I have in mind for my readers to evaluate it. I wish I could report that I have corrected this by completing my book on the unity and diversity of probability. Unfortunately, this book is still "forthcoming." However, I did present my axiomatization in one form in an appendix to Shafer (1985).

WHO SHOULD DO HISTORY?

Hacking makes a plea for leaving history to the historians. Philosophers make bad historians, he says, and so do statisticians. Historical research is best done by historians trained in archives.

Here I disagree. The conclusion is wrong, and the arguments are terrible.

Many historians (my wife, for example) wince at the suggestion that learning to do research in archives is the most important part of their training. Historians do need to be able to examine and judge evidence about the past, but they also need perspectives that will enable them to make sense of the past, and their task always involves relating the past to the present. Historians say that each generation rewrites history. Each generation must redo the French Revolution in light of its own understanding of politics and economics. Each generation must also redo Bernoulli in light of its own understanding of the law of large numbers.

It is disingenuous of Hacking to claim that his writings on the past are not history, for they are history, and very influential history at that. Historians need theses of the kind that Hacking formulates, and they must test these theses against the evidence. If his theses sometimes do not stand up to a careful examination of the archives, then Hacking must take his lumps; he cannot escape by taking the word "history" out of his books. Moreover, statisticians sometimes make very good historians. Stephen Stigler is one of the best historians of statistics ever—perhaps the very best. He is unsurpassed in the care and judiciousness with which he examines the archives.

Let statistics be done in many houses, Hacking proclaims. History, too, should be done in many houses. But we need ways to bring these houses together. The history of statistics, in order to be done as well as it should be, needs talents and training that we now find scattered in history, philosophy, statistics and elsewhere.

In my lecture, I suggested that statistics departments seek joint appointments with history departments, in order to strengthen our own understanding of our past and present. Historians also need these ties. The recent burst of work on statistics by historians is valuable, but it suffers from its exclusively externalist approach. (See my reviews of *The Rise of Statistical Thinking* and *The Empire of Chance* in Shafer, 1990a and b.) The next step in understanding the development of probability and statistics in the seventeenth, eighteenth and nineteenth centuries should be to integrate what these historians have done with a deeper understanding of the internal possibilities of the subject. As for the history of probability and statistics in the twentieth century, the historians have scarcely been able to get started by themselves, because their training in our discipline is not yet adequate.

In general, historians have not hesitated to reach out to other disciplines for new perspectives. History departments often have joint appointments with other departments and programs in the humanities and social sciences, including American studies, area studies, women's studies, African-American studies, sociology and economics. I believe that historians of science would welcome opportunities for similar interactions with statistics. Joint appointments between history and statistics could lead to a crop of historians capable to writing the history of twentieth-century statistics, a result that would serve both disciplines well.

THE STATURE AND INTELLECTUAL CONTENT OF APPLIED STATISTICS

A prominent statistician said to me recently that the central problem facing our discipline is how to raise the stature of applications *vis-à-vis* mathematics. This is reflected in the comments here. Hirotugu Akaike, George Box, Art Dempster and David Moore all demand more respect for applied statistics and the logic of statistics.

Why has this problem arisen? Why has prestige within our discipline moved in the direction of mathematics?

In truth, mathematics has always held the most prestigious position in the discipline. In Hotelling's model, mathematics was already the more or less exclusive intellectual foundation for statistics. Yes, the applied statistician needed "additional resources" (to use George Box's term), but these resources were to be obtained from experience. They were not the subject of graduate courses in the statistics department or of articles that would merit promotion in the discipline. In Hotelling's time, however, it was easier to combine mathematics flashy enough to gain prestige—within the discipline, at least—with applied concerns. As time has passed, an inevitable process of competition and theoretical progress has pushed up the requirements for mathematical recognition within the discipline, to the point where the applied statistician can no longer meet them in his free time.

In my lecture, I advance the thesis that we must make the other aspects of statistics part of its *intellectual* foundation. We must make Box's "additional resources" a subject of intellectual study—a proper topic for courses, books and articles, a specialty in which statistics departments recruit and promote faculty. I submit that unless we do this, applied statisticians will retain the second-rate status within statistics departments that Box bemoans.

Art Dempster and David Moore insist, and I agree, that there is a logic of applied statistics. How do we get it out of the heads of applied statisticians and into books, articles and graduate courses, where it can compete with mathematics for the attention of our brightest graduate students? We must approach the task with the same professionalism with which we approach the task of teaching mathematics. We cannot simply ask applied statisticians to bring their war stories to the classroom. We cannot simply require our graduate students to endure summer internships. We must develop scholars who teach and write about how applied statistics is done. And in order to develop such scholars, we must go outside for help. Just as we went to mathematics departments for help in developing our mathematical scholarship, we must go to history, the sciences and the professional schools for help in developing our scholarship about the realities of applied statistics.

I am convinced that we, more than the historians, need to learn the history of twentieth-century statistics. It is natural to assume that because we did it, we already know all about it. But we do not. Understanding requires effort. It requires observation, synthesis and criticism, an effort of many minds, even many generations. We have made a tremendous effort, over several generations, to achieve an understanding of the mathematics of statistics. We have not made a similar effort to understand, in Dempster's words, the "active logic of the process of doing statistics."

Instead of history—instead of sustained thinking about the complexity of contemporary applied statistics and probability—we now have only intriguing ideas without elaboration. Brad Efron, for example, asserts that the fractionation of statistical thinking is an evolutionary adaptation of the statistical point of view to different data environments. He is probably right, but whose doctoral dissertation has pursued the idea and subjected it to criticism?

In private conversations, some statisticians have reacted to my suggestion that we try to teach the logic of statistics with, "Oh no, not courses in methodology." One statistician buttressed his skepticism by observing that he has been impressed, over the course of his career, by how naive and vacuous prominent and accomplished statisticians—applied and theoretical—can be when they turn their attention to methodology. This I can believe. How could it be otherwise when, as statisticians, they are trained only in mathematics?

MISUNDERSTANDINGS

Our failure to pursue an intellectual understanding of applied statistics makes it difficult even for statisticians to talk to each other about the topic. It is difficult for applied statisticians to claim credit due, and it is difficult for theorists to avoid insulting applied statisticians. Since we know how to talk about the mathematical component of statistical reasoning and we do not know how to talk about the nonmathematical component, we can scarcely avoid unbalanced accounts—accounts full of holes that look like slights to the applied statistician.

In my description of how statistics departments use joint appointments, I mentioned that statisticians with such appointments often communicate new problems back to their statistical colleagues, together with "their own attempts at solutions." I wish I had written merely "their own solutions," for George Box has interpreted my words as a disparagement of those who actually apply statistics. I am surprised and chagrined by this, for I meant only to convey the standard idea that new applied problems are a stimulus to theory. It is no disparagement of those first working with such problems that those following in their footsteps can sometimes improve on their solutions.

Box is right to have wanted me to say more. It is often the new ideas that the applied statistician devises, not further mathematical elaboration of these ideas, that constitute the most important contributions to theory. But this is not the formulation that comes first to my mind. Mathematics is my intellectual framework for thinking about statistical theory, and I am apt to need reminding to distinguish between the two.

Just after having read George Box dressing me down, I read David Aldous on applied probability, and I shuddered at Aldous's failure to discuss the nonmathematical intellectual content of contributions to applied probability. I do not fault Aldous for this. None of us can discuss this nonmathematical intellectual content, for no one has given us the words with which to do it. But in the absence of such a discussion, contributions close to real problems seem to derive merit only from their location, not from their content. Figure 1 portrays Aldous's linear interpolation as it might feel to an applied statistician. As we go from real scientific problems to pure mathematics, the level of intellectual content required for publication rises from near zero ("routine mathematics") to its maximum.

In order to learn to talk to each other, we must make it part of the mission of any department (or journal) that houses applied probability to exposit the nonmathematical component of probabilistic modeling and reasoning in real problems. We need books, articles and courses that survey and organize the diverse applications of probability, describe the strategies at work and analyze their success and failures. When there are such courses in Aldous' department, applied applied probability, as opposed to pure applied probability, will no longer be as fantastic and indescribable as the unicorn, for both Aldous and I will have a vocabulary for praising applications that matches our vocabulary for praising mathematics.

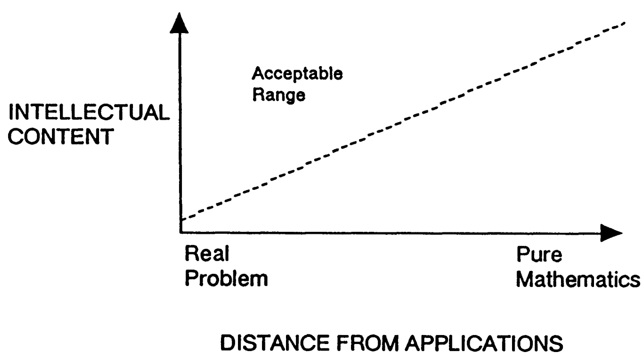

FIG. 1. *Aldous' linear interpolation.*

WHO SHOULD DO STATISTICS?

I agree with Ian Hacking and David Moore that it is healthy for teaching and research in statistics to flourish outside statistics departments. I do take very seriously, however, the leadership role of the statistics department.

Some years ago, in the course of an unsuccessful effort to establish a statistics department at the University of Kansas, I looked at the numbers of statistics courses offered, in and out of statistics departments, at several comparable universities. I found that the universities with statistics departments usually had more rather than less statistics taught in other departments.

I have spent my teaching career at Princeton, where statistics is now dormant, and at Kansas, where there never has been a statistics department. This gives me a different perspective than my colleagues at Stanford, Purdue and Berkeley. I am less confident than Efron that statistics will flourish even if statistics departments do not, less confident than Moore that conversations about probability will take place among varied scientific cultures even if statistics departments do not serve as the foci and more inclined than Aldous to believe that anyone who wants to influence the intellectual life of a university must pay attention to where administrative lines are drawn. I invite these statisticians to take a close look at Kansas and Princeton, to see how far universities can go in doing without advanced courses in statistics. The fact that there is a need does not mean it will be filled.

Having lived outside statistics departments for 14 years, I am also freer to fantasize about a statistics department with a broad mission. Indeed, I do not see how you could start to convince a university such as Kansas or Princeton that it should have a statistics department unless you were to promise more breadth than you find in existing statistics departments. A statistician at Berkeley, Stanford or Purdue may find my suggestions for broadening statistics more difficult. It might not be easy, for example, to convince some well-established statistics faculties that historians could deserve appointments in their departments.

Hirotugu Akaike has found it easier to bring people with diverse training together under the name "statistical science" than under the same "statistics." The name does not matter. What matters is leaders such as Akaike, who will bring ideas and people from outside statistics into statistics.

THE LIMITS OF PROBABILITY

A serious history of twentieth-century applied probability and statistics will be more than a collection of success stories. In truth, most problems of inference in science and the professions do not lend themselves to effective probabilistic or statistical treatment. An understanding of the intellectual content of applied probability and applied statistics must therefore include an understanding of their limits. What are the characteristics of problems in which statistical logic is not helpful? What are the alternatives that scientists, engineers, and others use? What, for example, are the characteristics of problems for which expert systems should use nonprobabilistic tools of inference?

Here is another way in which a better understanding of our own discipline requires expertise from other disciplines. If we are going to understand when to use probability and when to use alternatives, we must understand the alternatives. We must, for example, understand the nonprobabilistic methods of inference for artificial intelligence that have been developed in computer science. Contrary to Efron's assertion, there are computer scientists who think systematically about inference. We need to absorb their expertise into our understanding of the limits of probability and statistics.

THE QUALITY OF TEACHING

Both Aldous and Moore conclude their comments with a plea for more emphasis on the quality of teaching, especially the teaching of beginners. I am sympathetic with this, and I applaud Moore's putting his ideas into textbooks that bring the excitement of the history and logic of statistics into introductory courses. In the long run, however, I do not believe that we will convey this excitement to our beginning students unless we cultivate it in our research and advanced teaching. We are all proponents of the scholar-teacher model; we believe that in the long run, the best teaching is teaching based on the teacher's own life of the mind. At least we believe it when we want to explain to administrators and taxpayers why we should have time to do our research. We should also believe it when we make decisions about the allocation of resources within statistics departments. If we want our students to have teachers who can convey excitement about the history and logic of statistics, we must appoint scholars with serious research interests in these topics.

A CHALLENGE

I would like to end with a plea for more discussion. In my lecture, and again in this rejoinder, I have offered a rationale for strengthening our ties to other disciplines. Other statisticians, with more experience and insight than me, will be able to improve this rationale. Such improvement, and the consensus that can result form further discussion, are prerequisites for effective action.

There are many signs of an awareness of the need for new breadth in statistics. The very existence of the journal *Statistical Science* is one of the these signs. The new vigor of the applications section of the *Journal of the American Statistical Association* is another. The recent IMS report on cross-disciplinary research (IMS Panel on Cross-Disciplinary Research, 1990) is another. But there is not yet a consensus that the need for new breadth applies to the intellectual core of statistics. The applications section of *JASA* is used primarily by statisticians in other disciplines, and the very word "cross-disciplinary" deemphasizes the interdisciplinary nature of statistics itself.

I would like to challenge the contributors to this discussion, and other statisticians as well, to address these questions: What is the intellectual content of applied statistics, and how can we make this intellectual content as central to our discipline as the mathematics of statistics?

Art Dempster has been addressing these questions squarely, in his contribution to the discussion here, in his own publications on the philosophy of applied statistics (e.g., Dempster, 1990) and in his recent work on course development. I am grateful for his support and applaud the scope of his work.

I would particularly like to challenge George Box, perhaps the most accomplished living applied statisticians, to address these questions. We have much to learn from him about the intellectual content of applied statistics, and we deserve more than references to "additional resources." Box may not be sure what the originators of the idea of statistics departments intended, but he must know a lot about what the founder of at least one statistics department intended, and he could teach us what he learned from that experience.

Brad Efron also has the stature to contribute very influentially to this discussion. In his comments here, he modestly hopes that we are generating ideas that will secure our successor's place in the academy. As a leader in generating ideas that have stimulated new mathematical work, he can speak with authority about how we can achieve a balance between such mathematical work and broader understandings. I hope he will do so.

ADDITIONAL REFERENCES

AKAIKE, H. (1985). Prediction and entropy. In *A Celebration of Statistics* (A. C. Atkinson and S. E. Fienberg, eds.) 1–24. Springer, New York.

BENACERRAF, P. and PUTNAM, H. (1983). *Philosophy of Mathematics: Selected Readings*, 2nd ed. Cambridge, Univ. Press, London.

BERGER, J. O. (1984). The robust Bayesian viewpoint (with discussion). In *Robustness of Bayesian Analysis* (J. Kadane, ed.) 63-124. North-Holland, Amsterdam.

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer, New York.

BERGER, J. O. (1990). Robust Bayesian analysis: Sensitivity to the prior. *J. Statist. Plann. Infer.* **25** 303–328.

BOX, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness. *J. Roy. Statist. Soc. Ser. A* **143** 383–430.

BOX, G. E. P. (1983). An apology for ecumenism in statistics. In *Scientific Inference, Data Analysis, and Robustness* (G. Box, T. Leonard and C.-F. Wu, eds.) 51–84. Academic, New York.

BOX, G. E. P. (1984). The importance of practice in the development of statistics. *Technometrics* **26** 1–8.

BOX, G. E. P. (1990). Discussion of "Communications between statisticians and engineers/physical scientists" by Bruce Hoadley and Jon Kettenring. *Technometrics*. To appear.

BROWNIE, C. and KIEFER, J. (1977). The ideas of conditional confidence in the simplest setting. *Comm. Statist. A—Theory Methods* **6** 691–751.

DEMPSTER, A. P. (1990). Aspects and examples of uncertain causal assessment. Fifteenth International Biometric Conference.

GARFIELD, J. and AHLGREN, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *J. Res. Math. Ed.* **19** 44–63.

GARFUNKEL, S. and YOUNG, G. (1990). *Math Outside of Math: A Study of Mathematics Enrollments in Non-Mathematics Departments*. COMAP, Inc., Arlington, Mass.

GIGERENZER, G., SWIJTINK, Z., PORTER, T., DASTON, L., BEATTY, J. and KRÜGER, L. (1989). *The Empire of Chance*. Cambridge Univ. Press, Cambridge.

GRATTAN-GUINESS, I. (1970). *Joseph Fourier 1768-1830: A Survey of His Life and Work*. Harvard Univ. Press.

HACKING, I. (1990a). *The Taming of Chance*. Cambridge Univ. Press, Cambridge.

HACKING, I. (1990b). The disunities of the sciences. In *The End of Science? The Proceedings of the 25th Nobel Conference* (R. J. Elvee, ed.). Harcourt Brace Jovanovich, San Francisco. To appear.

IMS PANEL ON CROSS-DISCIPLINARY RESEARCH (1990). Cross-disciplinary research in the statistical sciences. *Statist. Sci.* **5** 121–146.

KIEFER, J. (1977). Conditional confidence statements and confidence estimators. *J. Amer. Statist. Assoc.* **72** 789–827.

KUHN, T. S. (1976). A function of measurement in modern physical science. In *The Essential Tension* 220. Chicago Univ. Press.

LAKOFF, G. (1987). *Women, Fire and Dangerous Things: What Categories Teach About the Human Mind*. Chicago Univ. Press.

MOORE, D. S. (1988). Should mathematicians teach statistics (with discussion). *College Math. J.* **19** 3–35.

NISBETT, R. E., FONG, G. T., LEHMAN, D. R. and CHENG, P. W. (1987). Teaching reasoning. *Science* **238** 625–631.

SHAFER, G. (1985). Conditional probability (with discussion). *Internat. Statist. Rev.* **53** 261–277.

SHAFER, G. (1990a). Review of *The Empire of Chance. How Probability Changed Science and Everyday Life*, by G. Gigerenzer, Z. Swijtink, T. Porter, L. Daston, J. Beatty, and L. Krüger (Cambridge Univ. Press, 1989). *Science* **247** 225.

SHAFER, G. (1990b). Review of *The Rise of Statistical Thinking, 1820–1920*, by Theodore M. Porter (Princeton University Press, 1986). *Ann. Sci.* **47** 207-209.

TURNER, J. (1990). More students take advanced math courses in departments other than mathematics. *Chronicle of Higher Education* March 21.

TVERSKY, A. AND KAHNEMAN, D. (1983) Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Rev.* **90** 293–315.