Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

Edited by James F. Crow and William F. Dove

R. A. FISHER, A CENTENNIAL VIEW

R. A. FISHER was born 100 years ago on February 17, 1890, in London. He was one of a pair of twin boys, the other being stillborn. What a tragedy that they did not both survive! And I would wish them to have been monozygotic. What would a replicate of FISHER's DNA have produced? Would he have had his brother's extreme nearsightedness? His urbane conversation and graceful prose? His witty and sometimes pointed sarcasm? His unpredictable temper explosions? His social idealism? Above all, his mathematical creativity and his astonishing geometric intuition? And how would two FISHERs have gotten on?

FISHER was an outlier, both scientifically and personally. His scientific work has often been reviewed, and we can expect more in this centennial year. This account is more personal, written by an admirer who knew him less well than some but better than most.

FISHER had an insight into multidimensional geometry that was little short of occult. He could answer questions that completely baffled others. He often arrived at an elegant answer, seemingly with no intermediate steps. Somehow, he found the pearl without opening the oyster. This meant that his papers could not be understood by most mathematicians, with the result that they were often not trusted. Long ago I saw a copy of a FISHER paper that had belonged to a mathematician; in the margin he had scratched "Fisher is fishy." Ultimately, many of FISHER's results were demonstrated in a more orthodox way, often by others.

Although FISHER had a talent for mathematics, his real interest was biology. Nevertheless, he sought and won a maths scholarship at Caius College, Cambridge. He chose mathematics for two reasons. One was that he had seen a mounted, disarticulated codfish skull and envisioned an arduous and futile exercise of learning all the bones. (I, too, was daunted by such a preparation.) The other reason was that he thought that, for a future biologist, "a mathematical technique with biological interests is a rather firmer ground than a biological technique with mathematical interests." His Cambridge tutor said later that he would have been a first class mathematician had he "stuck to the ropes." Yet, FISHER never regarded himself as a mathematician. The title of his biography (BOX 1978) is *R. A. Fisher, the Life of a Scientist*, surely the way he would have wanted it. Nevertheless, FISHER's great mathematical talent was predominant throughout his life.

As a schoolboy, FISHER had eyesight so bad that he was not permitted to read by lamplight and received his instruction, even mathematics, aurally and without visual aids. He developed a remarkable ability to solve problems in his head and acquired the geometrical insights that were so natural to him and so baffling to others. To what extent this was innate ability and to what extent a necessity brought on by poor eyesight, we shall never know. FISHER's development of significance tests for correlation and regression coefficients, for the t distribution, and for the analysis of variance were all done geometrically.

While still a student at Cambridge, he wrote a paper (FISHER 1912) in which he maximized the expression that he later called likelihood. Despite a degree from Cambridge and an evident talent, his next six years (1913–1919) were miserable. He was refused admission to the army. He worked in an office, taught school, and in his spare time tried subsistence farming. Yet, during this period he wrote two famous papers. One (1915) showed how to test the significance of a correlation coefficient, r, and introduced the transformation, $z = \tanh^{-1} r$, in which the distribution of z is nearly normal. The other (1918) reconciled biometry and Mendelism and laid the foundations for quantitative genetics. I have written about this remarkable paper before in this column (CROW 1988). The conclusions have hardly been changed in the 72 years since it was written, although they have been formulated and proved more precisely and rigorously by MALÉCOT (see NAGYLAKI 1989).

Finally, in 1919, FISHER was offered a temporary position to analyze agricultural data at Rothamsted Experimental Station. He accepted, and he stayed. There he developed the statistical procedures and experimental designs that are now universally used.

Genetics 124: 207–211 (February, 1990)

The early farm-crop influence is reflected in the retention of such words as plots and blocks in analysis of variance and of nitrogen, phosphorus, and potash in textbook explanations of factorial design. The enormous increases in crop yields in the past half century owes a great deal to reliable field testing that used these methods. Fisherian practices spread widely at about the same time that hybrid corn was introduced. Perhaps the inbreeding and hybridization technique should share some of the credit it enjoys with the efficient design of field trials.

FISHER's contributions to statistics are legion, and so well known that I shall mention them only in passing. Small-sample statistics, analysis of variance and covariance, experimental design, and statistical estimation are subjects that he founded. He straightened out the number of degrees of freedom for PEARSON's χ^2 test, he recognized the importance of STUDENT's[1] t test and demonstrated its correctness, and he pointed out the useful properties of the maximum likelihood method. His book *Statistical Methods for Research Workers*, despite being uniformly panned by reviewers, went through 14 editions and was translated into French, German, Italian, Japanese, Spanish and Russian. He was surely the greatest statistician of his time, if not of all time.

Something that is less fully appreciated is that FISHER was the first to employ nonparametric tests involving permutations of the observations. I think it is clear that he regarded randomization as primary, and tests based on normality assumptions as labor-saving approximations. This view is clearly set forth in, of all places, an expository paper on craniometry. Here he described how measurements on two groups of 100 individuals could be written on cards, then shuffled and divided randomly into two sets of 100 each. One could then ask what fraction of such random divisions would lead to a difference between the sets as large as or larger than the observed difference between the two measured groups. If this fraction were small, the groups could be regarded as differing significantly. Then FISHER (1936) wrote: "Actually, the statistician does not carry out this very simple and very tedious process, but his conclusions have no justification beyond the fact that they agree with those which could have been arrived at by this elementary method." One other relevant consideration: FISHER's methods were all devised with a view to minimizing complex computations. I believe that if high speed computers had been available, FISHER would have relied much less on normal distribution theory and much more on robust permutation tests.

This year is the 60th anniversary of another FISHER tour de force, *The Genetical Theory of Natural Selection*, arguably the deepest and most influential book on evolution since DARWIN. Each rereading of this classic brings something new. The book starts out by contrasting blending and particulate inheritance and emphasizing the remarkable variance-conserving properties of the latter, never before so clearly articulated. FISHER introduced what he called the "Malthusian parameter" as a measure of population increase. This was not new, but what was new was an extension, his "reproductive value." This is a weight to be assigned to each age group in proportion to the contribution of that group to the future population after age stability has been achieved, and thus it has an evolutionary as well as demographic significance. The idea has become popular with demographers (*e.g.*, KEYFITZ 1968). In FISHER's book the central idea was his "Fundamental Theorem of Natural Selection," that the rate of increase of fitness attributable to gene-frequency changes under selection is given by the additive component of the genetic variance. Although a cottage industry has grown up devoted to criticisms, exegeses and proofs, this succinct statement seems to me to capture the essence of the way selection works and to encapsulate a great deal of evolutionary insight in a simple formula.

FISHER not only asked important questions; he found answers. Some of his mathematical tricks were astonishing. He developed an ingenious method for finding the probability of survival of a mutant gene for a specified number of generations. He worked out the partial differential equation for gene-frequency change using a trigonometric transformation that made the variance independent of the allele frequency. He generalized HALDANE's formula, $P = 2s$, for the probability of ultimate fixation of a gene whose heterozygous selective advantage is s; it was further generalized by MALÉCOT (1952) and by KIMURA (1964). He gave the first coherent quantitative theory of sexual selection, mimicry, polymorphism, evolution of recombination rates, and supergenes. He explained why the sex ratio is nearly 1:1, even in polygamous species–one of the best illustrations that natural selection does not necessarily maximize fitness. In doing this he introduced the concept of parental expenditure, thereby precipitating a landslide of ecological literature. Rarely have so many new, and often deep, ideas been put into a single book. Curiously, although FISHER led the way to a quantitative theory of random genetic drift, he never regarded it as having much evolutionary significance. The evolutionary possibilities of random drift were advocated, with quite different emphases, by WRIGHT (1988 and earlier) and KIMURA (1983).

FISHER is never easy reading. The book is far from the explicit formulation and clear exposition that is

[1] "STUDENT" was a pseudonym of W. S. GOSSET, whose employers, the Guinness Breweries of Dublin, did not permit him to publish under his own name. Another employee, E. M. SOMERFIELD, published as "ALUMNUS."

the ideal of contemporary population genetics. He was partly poet. He was as much a master of elegant English as of elegant mathematics. But elegance and clarity are not the same. Fisher hardly ever made clear what his assumptions were, when and how he was approximating, and how to get from one equation to the next. I can empathize with GOSSET, who once wrote: "When I come to 'evidently' I know that means two hours hard work at least before I can see why" (Box 1978, p. 115).

The last five chapters of *The Genetical Theory* are devoted to human society. From his student days FISHER had been an ardent eugenicist, full of idealism and belief that mankind could be persuaded to reproduce so that the hereditary components of health, intelligence, character, and social conscience would increase. A much discussed topic of the time was the rise and fall of civilizations, about which FISHER read a great deal. His idea was that promotion of the gifted and industrious into a higher social class, where they would reproduce less, was a major factor in the decay of civilizations, and he discussed social and economic incentives that might forestall this. He advocated voluntary sterilization of the genetically impaired and family incentive payments proportional to income. As far as I can tell, his eugenic writings have had no lasting influence on either biologists or historians. In his later life Fisher did not write about these subjects, nor did he talk about them (to me at least). I don't think he had changed his mind, but simply tired of trying to get people to take his proposals seriously. At the same time, he was increasingly honored for his statistical and evolutionary work.

FISHER was part of the great trinity that included SEWALL WRIGHT and J. B. S. HALDANE. Together they founded and almost completely dominated the field of population genetics for its first quarter century. Each made important contributions, but in one way FISHER stands apart. HALDANE and WRIGHT formulated a problem and then doggedly ground out the results, come what might. FISHER was more likely to invent a new, neater approach. His work had elegance and grace, and flashes of insight and creativity, along with a touch of genius that can be fully appreciated only by those with mathematical insights deeper than mine.

During all of his active life–at Rothamsted, at University College London, and at Cambridge–FISHER always had genetic experiments going, often in collaboration with friends. He studied dogs, poultry, locusts, butterflies, sorrels, primroses, and especially mice. Many of the animals were kept in his home and he and his family took care of them. The presence of rooms full of mouse cages in the Professor's lodging is said to have been a deterring factor in the selection of his successor at Cambridge. What came out of all this experimental work was minor, certainly nothing comparable to what came out of his head. Yet, I think FISHER's constant touch with experiments and field observations guided his statistical and evolutionary work along practical lines. His most lasting contributions to experimental genetics are methodological. He showed how to measure linkage when simple backcrosses were impractical. His last paper (1962), lightweight by his standards, was on this subject. He exhaustively classified the gametic output of tetraploids, hexaploids and octoploids and showed how to allow for double reduction. He recognized ascertainment bias in human studies and examined the efficiency of various procedures designed to overcome it. He worked out computation-saving methods of detecting and measuring linkage in human pedigrees. Although his procedures have recently been superseded by computerized methods, his likelihood approach is the basis of most of them.

To FISHER, genetics was transmission genetics, strange as this seems today with the current emphasis on molecular approaches to gene action and development. Intermediate mechanisms were of secondary interest to him. Meiosis, for example, was a black box. In 1947 JOSHUA LEDERBERG and I sat together at the founding meeting of the Biometrics Society at Woods Hole. FISHER was elected president and gave a major address. He presented a model of recombination and interference that, among other things, permitted more than 50% recombination (for which he had some supporting mouse data). We were both taken aback by his not taking account of the four-strand nature of crossing over and exchanged whispered expressions of incredulity. Later, in response to LEDERBERG's question as to why he used a two-strand model, FISHER said: "Young man, it is not a two-strand model, it is a one-strand model." This epitomized FISHER's view of genetics. He developed the point more fully in the published paper and discussion (FISHER 1948). The geneticist's job, he said, is to develop a theory for predicting the frequencies of different genotypes from multiply heterozygous parents.

FISHER placed great emphasis on linkage analysis and chromosome mapping, and much of his mouse work was directed to this end. As soon as he had a formal position in genetics, he extended this interest to human genetics. He played an active role in gathering information on the rapidly increasing number of genetic markers, especially blood groups, with a view to mapping the human genome. Out of this grew his novel three-locus hypothesis for inheritance of the Rhesus factor (FISHER 1947), which at least notationally was a great advance. FISHER loved formal genetics; what a time he would have with human linkage analysis were he still alive, and how he would delight in the powerful computers and the plethora of reliable neutral markers!

FISHER enjoyed conversation and could be utterly charming. He could also be petty, quarrelsome, stubborn and outspoken. He fitted the classical definition of a gentleman: he never insulted anyone unintentionally. He was constantly involved in one or another controversy, often with other distinguished statisticians and geneticists, *e.g.*, JERZY NEYMAN and SEWALL WRIGHT. His sarcastic barbs could be amusing, except to their targets. FISHER was particularly bitter toward KARL PEARSON, who had misunderstood his early work and had treated him with arrogance. He was at has acerbic best (or worst) with PEARSON who, a decade after his death, elicited this: "If peevish intolerance of free opinion in others is a sign of senility, it is one which he had developed at an early age" (FISHER 1950, p. 29.302a).

In his later years FISHER visited the University of Wisconsin several times, mainly because a daughter lived in Madison. He always visited the Genetics Laboratory; we looked forward to his visits and saved problems for him. But it was necessary to engineer his coming and going so that he would not encounter SEWALL WRIGHT in the hallway. Their relationship had deteriorated to the point that neither wanted to see the other.

I shall finish this essay with two personal anecdotes. The first concerns my first meeting FISHER. It was during a statistics course at North Carolina State College in the summer of 1946. He gave an evening lecture to a large audience, composed almost entirely of statisticians, on his three-locus theory of Rh inheritance. This was new to me, and I was entranced. In the question period he was first asked how he did the χ^2 test, to which he gave a curt answer. Clearly it was the genetics that interested him so I asked some genetic questions, which pleased him and which we continued informally after the session was closed. He suggested a glass of beer at a bar across the street. (I then realized for the first time that in poor light Fisher was nearly blind.) This was a time of postwar shortages, and the bar had run out of both beer and wine. There was champagne, however, and we got a bottle, only to be told that North Carolina law prohibited drinking it on the premises. So we repaired to my dormitory room and began, over a shared bottle of champagne, a friendship that lasted through the remainder of his life.

The second anecdote concerns the famous paper of LURIA and DELBRÜCK (1943). I found its argument for the preadaptive nature of evolution of virus resistance in bacteria fully convincing, but thought that the mathematical treatment was shoddy and confusing. Taking advantage of my newly formed acquaintance with FISHER, I asked him how to find the distribution of mutant cells in an exponentially growing culture. He leaned back in his chair, thought for perhaps a minute, took a scrap of paper, and wrote a generating function. I took the paper and, not understanding it, put it aside to work on later–and then managed to lose it. The solution was published two years later by LEA and COULSON (1949). Unless that scrap of paper turns up, we'll never know whether FISHER was the first to solve this problem.

FISHER died in 1962. He had written several hundred reviews, comments, and letters. His major papers–294 of them–are included in five volumes edited by BENNETT (1971–1974), often with introductory comments and amendments by FISHER himself. The first volume also includes a biography, written by F. YATES and K. MATHER. Those interested in his personal life will enjoy the biography by JOAN FISHER BOX (1978). Written by a loving and admiring daughter, the book is touching as it brings out FISHER's blemishes along with his greatness. It is also scholarly, for BOX took the trouble to understand and explain the difficult conceptual points, especially in statistics.

A large number of people have read an earlier draft of this article and I am grateful for their comments. My greatest debt is to JOAN FISHER BOX and THOMAS NAGYLAKI, who provided numerous improvements in both content and style.

JAMES F. CROW
Genetics Department
University of Wisconsin
Madison, Wisconsin 53706

BIBLIOGRAPHY

Books by and about FISHER:

BENNETT, J. H. (Editor), 1971–1974 *Collected Papers of R. A Fisher.* University of Adelaide, Australia.

BENNETT, J. H. (Editor), 1983 *Natural Selection, Heredity, and Eugenics: Including Selected Correspondence of R. A. Fisher with Leonard Darwin and Others.* Clarendon Press, Oxford.

BOX, J. F., 1978 *R. A. Fisher, the Life of a Scientist.* John Wiley & Sons, New York.

FISHER, R. A., 1925–1970 *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh. 14th ed. (1971, 1973) Hafner, New York.

FISHER, R. A., 1930 *The Genetical Theory of Natural Selection.* Clarendon Press, Oxford. 2nd ed. (1958) Dover, New York.

FISHER, R. A., 1935–1966 *The Design of Experiments.* Oliver and Boyd, Edinburgh. 8th ed. (1971, 1973) Hafner, New York.

FISHER, R. A., 1949, 1965 *The Theory of Inbreeding.* Oliver and Boyd, Edinburgh.

FISHER, R. A., 1950 *Contributions to Mathematical Statistics.* John Wiley & Sons, New York.

FISHER, R. A., 1956, 1959 *Statistical Method and Scientific Inference.* Oliver and Boyd, Edinburgh. 13th ed. (1973) Hafner, New York.

FISHER, R. A., and F. YATES, 1938–1963 *Statistical Tables for Biological, Agricultural and Medical Research.* Oliver and Boyd, Edinburgh.

Cited articles:

Many of FISHER's papers were published in obscure journals. The

best source is the five-volume set, comprising 294 papers, edited by J. H. BENNETT and listed above.

CROW, J. F., 1988 Fifty years ago: the beginnings of population genetics. Genetics **119:** 473–476.

FISHER, R. A., 1912 On an absolute criterion for fitting frequency curves. Messeng. Math. **41:** 155–160.

FISHER, R. A., 1915 Frequency distribution of the values of the correlation coefficients in samples from an indefinitely large population. Biometrika **10:** 507–521.

FISHER, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinb. **59:** 399–433.

FISHER, R. A., 1936 "The coefficient of racial likeness" and the future of craniometry. J. R. Anthropol. Inst. **66:** 47–63.

FISHER, R. A., 1947 The Rhesus factor: a study in scientific method. Am. Sci. **35:** 95–102, 113.

FISHER, R. A., 1948 A quantitative theory of genetic recombination and chiasma formation. Biometrics **4:** 1–13.

FISHER, R. A., 1962 The detection of sex differences in recombination values using double heterozygotes. J. Theor. Biol. **3:** 509–513.

KEYFITZ, N., 1968 *Introduction to the Mathematics of Population.* Addison-Wesley, Reading, Mass.

KIMURA, M., 1964 Diffusion models in population genetics. J. Appl. Prob. **1:** 177–232.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

LEA, D. E., and C. A. COULSON, 1949 The distribution of the numbers of mutants in bacterial populations. J Genet. **49:** 64–285.

LURIA, S. E., and M. DELBRÜCK, 1943 Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28:** 491–511.

MALÉCOT, G., 1952 Les processus stochastiques et la méthode des fonctions génératrices ou caractéristiques. Publ. Inst. Stat. Univ. Paris **1:** Fasc. 3, 1–16.

NAGYLAKI, T., 1989 Gustave Malécot and the transition from classical to modern population genetics. Genetics **122:** 253–268.

WRIGHT, S., 1988 Surfaces of selective value revisited. Am. Nat. **131:** 115–123.