1-1-2013

R.A. Fisher and the Foundations of Statistical Biology

Michael Dietrich
Dartmouth College

MICHAEL R. DIETRICH & ROBERT A. SKIPPER, JR.

8 R. A. FISHER AND THE FOUNDATIONS OF STATISTICAL BIOLOGY

INTRODUCTION

In July of 1951, Sir R. A. Fisher (1890–1962) used the occasion of his Bateson Lecture to reflect on statistical methods in genetics (see figure 8.1). Having made foundational contributions to both statistics and genetics beginning arguably in 1918, Fisher saw them as quintessential twentieth-century disciplines. While nineteenth-century antecedents could be easily found for both, statistics and genetics came to maturity as "distinct points of view" in the twentieth century.[1] Fisher played an important role in articulating the point of view of both modern genetics and modern statistics, but more importantly, he successfully managed their integration. In doing so, Fisher did more than bring his training in mathematics to bear on biological topics. He used his mathematical abilities to reconceive statistics, experimentation, and evolutionary biology.

Fisher came to biology as an outsider in the sense that his formal training as a student emphasized mathematics. His interests in genetics and eugenics, however, began early as well. Even though there were other well-known synthesizers of mathematics and biology at the time, Fisher's interests and training were not typical of the great majority of biologists of his day. Fisher's innovative work in statistical biology met both with indifference from many biologists, as a result, and often opposition from more established statistical biologists.

The opposition that Fisher's innovations faced led him to comment, "A new subject for investigation will find itself opposed by indifference, by inertia, and usually by ridicule. A new point of view, however, affecting thought on a wide range of topics may expect a much fiercer antagonism."[2] The statistical, genetic, and experimental "points of view" that Fisher championed developed significantly over the first decades of the twentieth century. The antagonism they faced only pushed Fisher to refine them further. A more passive advocate may have been pushed to the margins, but Fisher's persistence, personality, and patrons allowed him to redefine how mathematics could be brought into the core of biological thought and practice.

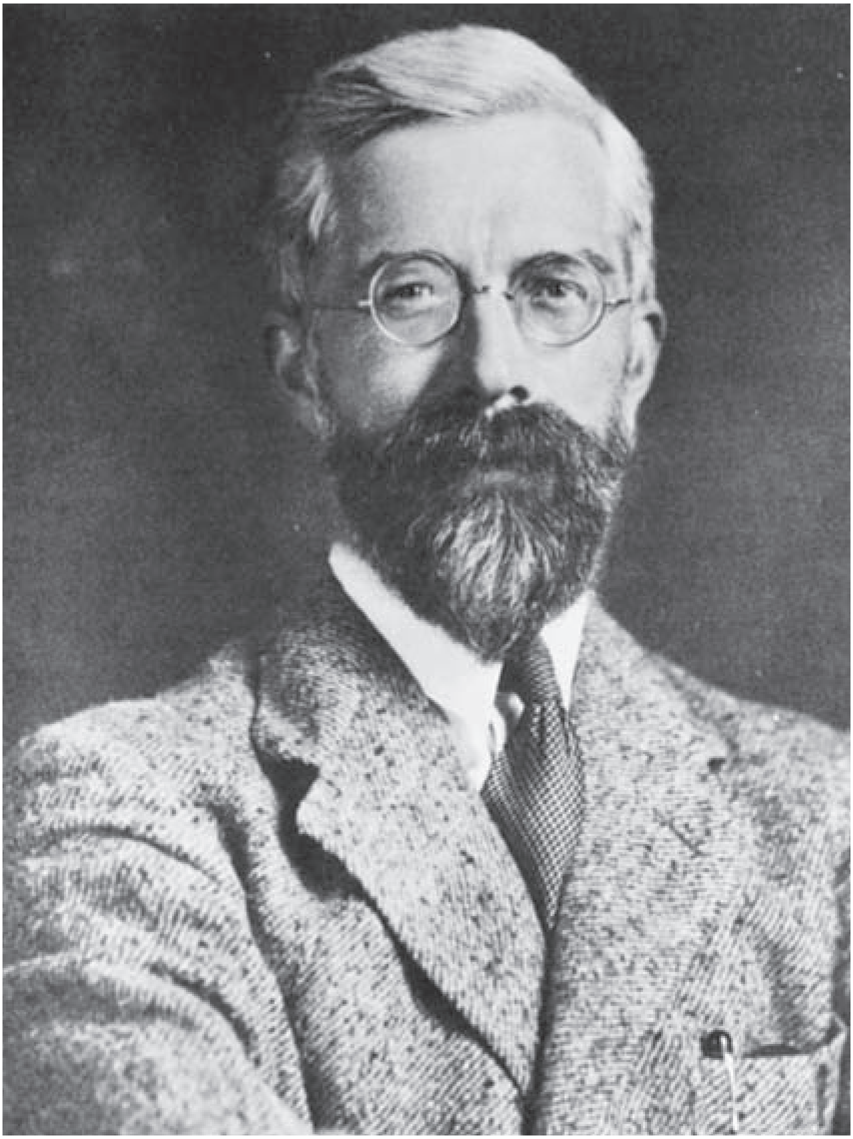

Figure 8.1: R. A. Fisher.
Photograph by SPL/Photo Researchers, Inc.

ARTICULATING NEW POINTS OF VIEW

In retrospect, Fisher claimed that genetics was a natural source for statistical thinking because of its heavy use of frequencies and the natural randomization of genotypes that make genetic experimentation easy. However natural the synergy between genetics and statistics may be, Fisher was initially drawn to their intersection through eugenics and biometry. As a student at Cambridge from 1909 to 1913, Fisher studied mathematics but developed broad interests. As he himself commented, he entered Cambridge on "the centenary of Darwin's birth and the jubilee of the publication of *The Origin of Species*."[3] William Bateson, Mendel's English champion, had been given a professorship the year before, and in 1912 it would be endowed as the Arthur Balfour Chair of Genetics.[4] Darwinism, Mendelism, and the debates over their differences would have been unavoidable for Fisher at Cambridge. Eugenics made them irresistible.

Sir Francis Galton inaugurated the English eugenics movement in 1869 with his book *Hereditary Genius*.[5] Galton's eugenics rested on the scientific management of human heredity by encouraging reproduction among those supposed to be "fit" and discouraging reproduction of the supposedly "unfit." In doing so, eugenicists hoped to direct the course of human evolution.[6]

Once introduced to the rising eugenics movement, Fisher's youthful enthusiasm led him to help form the Cambridge University Eugenics Society in 1911. Eugenics for Fisher was not a passing fad, however. His intellectual interest in the topic never faltered, and Fisher sought to live by the eugenic principles that he advanced by having eight children.[7]

Eugenics formed the natural bridge between Fisher's interest in mathematics and the new fields of genetics and statistics. In turn-of-the-century England, Francis Galton and Karl Pearson occupied this intersection, and it was to their work that Fisher turned as a college student. In order to understand the resemblance between generations, Galton had developed techniques of correlation and regression to represent heredity from a statistical point of view. Like his cousin, Charles Darwin, Galton believed that hereditary traits were continuous gradations of form best described by a normal distribution. An individual's heredity then was represented in terms of a law of ancestral heredity, in which an individual's ancestors each made a diminishing contribution.[8] Inspired by Galton's approach, Karl Pearson developed biometrics, a statistical approach to biology, which supported a Darwinian view of the gradual evolution of continuous traits under the direction of natural selection.

In 1911, Fisher addressed the Cambridge Eugenics Society on the intersection of Mendelism and biometry. At the time, Mendelism was cast by William Bateson in diametric opposition to the biometrical approach advocated by Karl Pearson. For Bateson, Mendelism supported his view of discrete hereditary characters and saltational evolution. However, as a dispute, the arguments between Mendelians and biometricians also raised questions about the appropriate role of probability and statistics. Fisher was taken by exactly these questions, and in his overview of both positions notes the role of probability in each and praises the power of the biometrical use of statistics to analyze biological observations without relying on theory or abstraction. Moreover, from a eugenic point of view, the biometrical approach convinced him that it was possible to create "a slow and sure improvement in the mental and physical status of a population" without the complications of the "experimental breeding" that Mendelism would require.[9]

While still an undergraduate, Fisher made his first foray into statistics with his 1912 paper, "An Absolute Criterion for Fitting Frequency Curves."[10] Inspired by work on the theory of errors in astronomy and mathematics, Fisher criticized the least squares method and the method of moments. John Aldrich claims that Fisher's true target here was Karl Pearson's 1902 essay, "On the Systematic Fitting of Curves to Observations and Measurements."[11] Pearson favored both of these methods, but Fisher found their justification to be arbitrary, and so their agreement with each other was problematic rather than reassuring. In their place, he championed an approach to error he had learned from the astronomers based on Gauss's least squares method.[12]

The dominant school of thought at the time was Bayesian and employed what was called the method of inverse probability. Named for Thomas Bayes, the Bayes theorem for any two events A and B claims that the conditional probability of event A given event B is:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

In this equation, the probability of event A, P(A), is called the prior probability, meaning that it is the probability of A before event B. The probability of A given B or after B has occurred is the posterior probability. The Bayes theorem tells us how to adjust our prior probability of A in light of a new event B; or the prior probability of a hypothesis, H, in light of evidence, E. In more direct terms, Bayes theorem addresses the problem of how new evidence can guide the revision of our previous beliefs—it addresses the problem of induction.

At the time Fisher was writing, it was common practice to recognize that if the prior probability was unknown, then one could assume that there was, in Pearson's words, an "equal distribution of ignorance" so all probability values of H are the same, or all beliefs about the probability of H are equivalent. This means that the probability of H given E does not depend on the prior probability of H, but on the remaining term, P(E/H)/P(E). Fisher was critically examining the claim that the curve that best fits the data was the one that maximized the posterior probability. What he was proposing was that maximizing P(H/E) and maximizing P(E/H)/P(E) were not equivalent. In fact, they weren't even both probabilities.

Fisher was a frequentist. He believed that probabilities described relative frequencies of events in a certain number of trials or experiments. The probability of a kind of occurrence was estimated by the ratio of the number of observed occurrences of an event for a given total number of trials. Moreover, given an infinite number of trials, the ratio of the number of events to the number of trials will converge on the true probability value. Fisher understood the Bayesian approach as trying to assign a probability value to something that was unique—not subject to repeated trials, and not subject to sampling as a result. For Fisher, it was legitimate to ask about the probability of observing an experimental outcome, but it did not make sense to speak of the probability of a hypothesis. To differentiate these approaches, Fisher distinguished the probability of a hypothesis from its likelihood. The likelihood of a hypothesis given some evidence is equivalent to the probability of the evidence given the assumption of the hypothesis. For Fisher, maximizing likelihoods was the more statistically sound method of estimation and curve fitting since it could be grounded in observed frequencies. In making his case for the distinction between probability and likelihood, Fisher would transform the foundations of statistics. Translating the resulting authority from statistics to biology and demonstrating the deep relevance of statistics to biology were crucial in Fisher's transformation from a mathematician to a biologist.

Fisher was drawn to the theory of errors by a paper by "Student," really W. S. Gosset who was not allowed by his employer, Guiness Brewing, to publish under his own name since the statistical test was considered a trade secret. Gossett addressed the problem of estimating error when using a small sample size. This led Fisher to consider how to estimate error when calculating a correlation coefficient for small samples. Moreover, Fisher's maximum likelihood method gave importantly different results from Gossett's. The fine points of Fisher's essay were not fully grasped by Gossett or Pearson, who read it and discussed it in correspondence, but they did appreciate that Fisher was a young talent in statistics.

After graduation Fisher spent some time working on a farm in Canada, presumably to rest his notoriously poor eyesight. However, Fisher seemed to genuinely love farm life and would settle on a farm when he married in 1917. By then Fisher had returned to England and, determined to do his part for the war effort, taught mathematics and physics at public schools since he was not eligible for military service.

In 1914, Fisher published a proof of Gosset's solution using multidimensional geometry. While other statisticians did not share Fisher's fondness for geometrical reasoning about distributions, Pearson did publish it in his journal *Biometrika*.[13] Pearson's group was also occupied during the war with distributions of correlation coefficients, and in 1916, his group published a paper critical of Fisher's approach, claiming that it depended on the Bayes theorem. The injustice of this claim stung Fisher and motivated him clarify his stance on likelihood as an alternative to Bayes. At the same time it signaled a declining relationship with Pearson, who declined to publish Fisher's 1916 paper on correlation with regard to Mendelian traits. Pearson probably thought of himself as the senior professor helping to sort out the younger Fisher, and sent Fisher extensive comments on another 1916 note regarding error and estimation. Fisher responded by seeking to put statistics on a firmer mathematical foundation that directly took aim at the substantial work of Pearson.[14] When Fisher published his own criticisms of Pearson, their enmity became mutual. However, as historian Stephen Stigler has convincingly argued, the antagonism that Fisher felt toward Pearson certainly spurred him to develop his method of maximum likelihood and articulate the grounds for the justification of statistical methods.[15]

Throughout this period Fisher had maintained his interest in eugenics and sought an active role in the Eugenics Education Society. This interest brought him into contact with Major Leonard Darwin, Charles' fourth son and an avid eugenicist himself.[16] Darwin became Fisher's mentor and advocate around 1914, when Fisher began publishing book reviews for *Eugenics Review*. While Fisher scraped by financially as a farmer and teacher, Darwin arranged for a stipend from the Eugenics Society. This undoubtedly helped fuel the fire that led Fisher to write over 200 reviews on eugenics over twenty years.

But the relationship was something more. Fisher revered Darwin and sought his approval. When he discovered that Darwin and Pearson had had a disagreement on the effects of natural selection on the correlation of hereditary traits, Fisher sought to vindicate Darwin, but Darwin counseled restraint.[17] Nevertheless, Darwin encouraged Fisher to continue his work on the intersection of statistics and heredity. Indeed, when Fisher's paper on the correlation of relatives was rejected by Pearson for *Biometrika*, Darwin sponsored its publication in the *Transactions of the Royal Society of Edinburgh*. This paper helped Fisher land two job offers the next year: one in the Galton Laboratory and one at the Rothamsted Experimental Station. He chose Rothamsted.

EVOLUTION FROM A STATISTICAL GENETICS POINT OF VIEW

When Fisher was at Rothamsted Experimental Station from 1919 to 1933, he engaged with a mass of biological data that allowed him to revolutionize statistics, and he become convinced of the value of engaging statistical analysis with real biological data. Researchers at Rothamsted had accumulated years of data on crop growth and yield under a wide range of different conditions. Fisher's challenge was to find something biologically interesting in that data. What he found and published in a series of papers was important both biologically and statistically. Biologically, Fisher was able to disentangle the effects of various fertilizer treatments from soil, weather, and cultivation conditions. At the same time, Fisher developed methods for statistical experimentation based on randomization and the analysis of variance that would radically change the way in which any researcher could conduct a statistical experiment in the future. These methods were published in 1925 as *Statistical Methods for Research Workers*.[18] Initial reviews were critical as primarily English statisticians struggled to make sense of this new empirical approach. Reflecting on this later, Fisher offered that recognition takes time when the revision of cherished beliefs bruises the feelings of their holders. Fisher's recognition came first from abroad, where experimental agriculture was established and appreciated in institutions such as the USDA and many land grant universities in the United States.[19]

Fisher's work on the analysis of variance as applied to biology had begun much earlier in his work on correlation in evolution. In his 1918 paper, "On the Correlation of Relatives on the Supposition of Mendelian Inheritance," Fisher considered the statistical consequences of dominance, epistatic gene interaction, assortative mating, multiple alleles, and linkage on the correlations between relatives. Fisher argued that the effects of dominance and gene interaction would confuse the actual genetic similarity between relatives. He also knew that the environment could confuse such similarity. Fisher here formally introduced the concepts of variance and the analysis of variance. He wrote:

> When there are two independent causes of variability capable of producing in an otherwise uniform population distributions with standard deviations σ_1 and σ_2, it is found that the distribution, when both causes act together, has a standard deviation $\sqrt{(\sigma_1^2 + \sigma_2^2)}$. It is therefore desirable in analyzing the causes of variability to deal with the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the normal population to which it refers, and we may now ascribe to the constituent causes fractions or percentages of the total variance which they together produce.[20]

Fisher used this new tool to partition the total variance into its component parts. He labeled that portion of the total variance that accurately described the correlation between relatives the "additive" genetic component of variance. The "nonadditive" genetic component included dominance, gene interaction, and linkage. Environmental effects, such as random changes in environment, comprised a third component of the total variance. In 1922, on the basis of his 1918 work, Fisher argued that the additive component of variance was the most important for evolution by natural selection. Indeed, he argued that, particularly in large populations, nonadditive and environmental components of the total variance are negligible. Selection would act most strongly on any factor with a large additive contribution to the total genetic variance, usually by eliminating them from the population.[21] Most of the time, evolution, and especially adaptation, proceed very slowly, with low levels of selection acting on mutations of small effect and in large populations holding considerable genetic variation. Where Fisher's 1918 paper defended the principles of Mendelian heredity against the criticisms of the biometricians, his 1922 paper defended Darwinism *using* the principles of Mendelian heredity. Fisher's aim was to respond to a set of criticisms that Darwinian natural selection cannot be the correct mechanism of evolution because the genetics of populations are such that there is not enough variation available for selection to act upon. During the course of the paper, Fisher eliminated from consideration what he took to be insignificant evolutionary factors, such as epistatic gene interaction and genetic drift, and placed his confidence in natural selection.

Fisher's synthesis of Mendelism and Darwinism within a mathematical framework culminated in his 1930 book, *The Genetical Theory of Natural Selection*, which became one of the principal texts, along with those of J. B. S. Haldane and Sewall Wright, establishing the field of theoretical population genetics.[22] Fisher begins his book with his case for the mutual compatibility of Darwin's mechanism of natural selection and Mendelian genetics. He ends it by exploring the eugenic consequences of this statistical and genetic understanding of the evolutionary process. Fisher considered the first two chapters, on the nature of inheritance and the "fundamental theorem of natural selection," the most important of the book. Indeed, these two chapters accomplish the key piece of the reconciliation by continuing the general argument strategy he had used in 1918 and 1922 of defending Mendelian particulate inheritance and then demonstrating how Darwinian natural selection may plausibly be the principal cause of evolution in Mendelian populations.

Fisher's mathematical approach is most fully developed in his second chapter of *The Genetical Theory of Natural Selection*. The arguments here are drawn largely from Fisher's 1922 paper "On the Dominance Ratio" and his 1930 paper "The Distribution of Gene Ratios for Rare Mutations," which was a response to Sewall Wright's correction of Fisher's 1922 paper. Three key elements may be distilled from Fisher's "heavy" mathematics in the second chapter of *The Genetical Theory*. The first is a measure of average population fitness, Fisher's "Malthusian parameter" (i.e., the reproductive value of all genotypes at all stages of their life histories). The second is a measure of variation in fitness, which Fisher partitions into genetic and environmental components (based on his distinctions from 1918 and 1922). The third is a measure of the rate of increase in fitness (i.e., the change in fitness due to natural selection). For Fisher, "*the rate of increase of fitness of any organism at any time is equal to its genetic variance in fitness at that time*" (emphasis in original).[23] This last element is Fisher's "fundamental theorem of natural selection," and it is the centerpiece of his theory of natural selection.

Interestingly, inasmuch as Fisher considered his fundamental theorem the centerpiece of his evolutionary theory, it happens that the theorem is also the most obscure element of it. The theorem was often misunderstood until 1989, when Warren Ewens rediscovered George Price's 1972 clarification and proof of it.[24] Fisher's original statement of the theorem in 1930 suggests that mean fitness can never decrease because variances cannot be negative. Price showed that in fact the theorem does not describe the total rate of change in fitness but, rather, only one component of it. That part is the portion of the rate of increase that can be ascribed to natural selection. And, actually, in Fisher's ensuing discussion of the theorem, he makes this clear. The total rate of change in mean fitness is due to a variety of forces including natural selection, environmental changes, epistatic gene interaction, dominance, and so forth. The theorem isolates the changes due to natural selection from the rest, a move suggested in Fisher's 1922 paper. The relative importance of the additive component of genetic variance was increasingly appreciated in the genetics community after the Second World War. Price and Ewens recognized this and clarified the statement of the theorem by substituting "*additive* genetic variance" for "genetic variance" (since genetic variance includes both an additive and nonadditive part). With the theorem clarified, however, Price and later Ewens argue that it is not so fundamental. Given that it is a statement about only a portion of the rate of increase in fitness, it is incomplete. The Price-Ewens interpretation of the theorem is now the standard one.

Fisher compared both his 1922 and 1930 exploration of the balance of evolutionary factors and the "laws" that describe them to the theory of gases and the second law of thermodynamics, respectively. Of the 1922 investigation, Fisher says,

> the investigation of natural selection may be compared to the analytic treatment of the Theory of Gases, in which it is possible to make the most varied assumptions as to the accidental circumstances, and even the essential nature of the individual molecules, and yet to develop the natural laws as to the behavior of gases, leaving but a few fundamental constants to be determined by experiment.[25]

He continues the analogy in 1930, adding that

> the fundamental theorem . . . bears some remarkable resemblances to the second law of thermodynamics. Both are properties of populations, or aggregates, true irrespective of the nature of the units which compose them; both are statistical laws; each requires the constant increase in a measurable quantity, in the one case the entropy of the physical system and in the other the fitness . . . of a biological population. . . . Professor Eddington has recently remarked that "The law that entropy always increases—the second law of thermodynamics—holds, I think, the supreme position among the laws of nature." It is not a little instructive that so similar a law should hold the supreme position among the biological sciences.[26]

The received view of these comparisons is that Fisher's interests in physics and mathematics led him to look for biological analogs.[27] No doubt this is part of the story. However, a more plausible interpretation of the comparison comes from treating Fisher's 1918, 1922, and 1930 works as one long argument. If we do so, we find that Fisher's strategy in synthesizing Darwinian natural selection with the principles of Mendelian heredity was to defend, against its critics, selection as an evolutionary cause under Mendelian principles. Following this argument strategy, Fisher built his genetic theory of natural selection piecemeal, or from the bottom up. That is, Fisher worked to justify the claim of his fundamental theorem by constructing plausible arguments about the precise balance of evolutionary factors. Thus, his piecemeal consideration of the interaction between dominance, gene interaction, genetic drift, mutation, selection, etc. led to his theorem. It was not, at least not primarily, the search for biological analogues to physical models and laws that underwrites the theorem.

No one has thought that Fisher's contribution to evolutionary genetics was less than groundbreaking. Rather, precisely what Fisher established, its nature and scope, and exactly how he did so, has been less than clear. With Fisher's work on variance in 1918, his work on the balance of factors in evolution in 1922, and his fundamental theorem of natural selection in 1930, we have a unified argument setting aside pervasive anti-Darwinism, originating a new mathematical approach to the evolution of populations, and establishing the very essence of natural selection.

The Genetical Theory sealed Fisher's reputation as a biologist. In 1933, he succeeded Karl Pearson as the Galton professor at University College London. His work on mathematics had earned him a place in the Royal Society in 1928. His 1925 book, *Statistical Methods for Research Workers*, was changing the way experimental biology could be conceived.

CONCLUSION

R. A. Fisher made extraordinary contributions to the mathematical foundations of statistics, statistical methods for experimentation, and the creation of population and evolutionary genetics. When combined with his strong belief in the social value of eugenics, the range of Fisher's interests pose a serious challenge to historians who would like to make sense of how one person could simultaneously pursue such disparate topics and make such important contributions to each. Fisher's daughter and biographer, Joan Fisher Box, offered separate chapters tracing simultaneous trajectories through his mathematics, statistics, genetics, and eugenics.[28] Other historians treat Fisher's interests in isolation from each other, and in doing so cast themselves in sharp contrast to historians who see Fisher's interests as mutually informed.[29] Having posed the question of how nonbiological training can foster innovation in the life sciences, we find ourselves seeking points of intersection in Fisher's work. These intersections are plentiful, but what were the historical conditions that allowed Fisher to find success in these intersections?

We claim that patronage and persistence played crucial roles in allowing Fisher to successfully bring his mathematical and statistical perspectives into biology. The patronage of Leonard Darwin early in his career facilitated the publication of Fisher's seminal 1918 paper on the correlation of relatives, and supported his work at the intersection of biology, eugenics, and statistics. At the same time, Fisher's personality allowed him to stubbornly persist in the face of Pearson's criticism and later opposition. Fisher's willingness to engage in a dispute and to oppose Pearson and the entrenched views on inverse probability were crucial in the history of his development of the foundations for theoretical statistics, on the one hand, and the methods of estimation and experimentation crucial to biology and evolutionary genetics, more specifically.

We do not wish to claim that the success of Fisher's contributions are solely the result of patronage or Fisher's personal advocacy. Their value and utility were recognized by many. In the case of experimental design and statistical inference, Fisher's work found an eager audience among agricultural experts around the world following the path of Mendelism in an age of faith in scientific progress.[30] However, intellectual worth alone did not overcome the barriers set by Pearson and other statisticians. Fisher's results were innovative, but the intersection of statistics, genetics, eugenics, and evolution was not an empty niche waiting to be filled. It took time for Fisher's early work to gain acceptance among the statistical and scientific communities. What allowed him to continue to innovate in statistical biology was the support of individuals, like Leonard Darwin, and institutions like Rothamsted, where his methodological insights more readily produced new results for field researchers.

FURTHER READING

Box, Joan Fisher. *R. A. Fisher: The Life of a Scientist*. New York: John Wiley, 1978.

Fisher, R. A. *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press, 1930.

Fisher, R. A. "Statistical Methods in Genetics," *Heredity* 6 (1952), 1–12.

Mackenzie, Donald A. *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge*. Edinburgh: Edinburgh University Press, 1981.

NOTES

1. R. A. Fisher, "Statistical Methods in Genetics," *Heredity* 6 (1952), 1–12.

2. Fisher, "Statistical Methods in Genetics," 4.

3. R. A. Fisher, "The Renaissance of Darwinism," *Listener* 37 (1947), 1000, 1009.

4. Joan Fisher Box, *R. A. Fisher: The Life of a Scientist* (New York: John Wiley, 1978), p. 22.

5. Francis Galton, *Hereditary Genius: An Inquiry into Its Laws and Consequences* (London: Macmillan, 1869).

6. For more on British eugenics, see Donald Mackenzie, "Eugenics in Britain," *Social Studies of Science* 6 (1976), 499–532; Lyndsay A. Farrall, *The Origins and Growth of the English Eugenics Movement, 1865–1925* (New York: Garland Publishing, 1985); Pauline Mazumdar, *Eugenics, Human Genetics, and Human Failings: The Eugenics Society, Its Source and Its Critics in Britain* (London: Routledge, 1992).

7. Box, *R. A. Fisher*; Donald A. Mackenzie, *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge* (Edinburgh: Edinburgh University Press, 1981).

8. William B. Provine, *The Origins of Theoretical Population Genetics* (Chicago: University of Chicago Press, 1971); Kyung-Man Kim, *Explaining Scientific Consensus: The Case of Mendelian Genetics* (New York: Guilford Press, 1994).

9. R. A. Fisher, "Mendelism and Biometry (1911)," in J. H. Bennett, editor, *Natural Selection, Heredity, and Eugenics: Including Selected Correspondence of R. A. Fisher with Leonard Darwin and Others* (Oxford: Oxford University Press, 1983), p. 57.

10. R. A. Fisher, "On an Absolute Criterion for Fitting Frequency Curves," *Messenger of Mathematics* 41 (1912), 155–160.

11. John Aldrich, "R. A. Fisher and the Making of Maximum Likelihood, 1912–1922," *Statistical Science* 12 (1997), 162–176, 163.

12. Aldrich, "R. A. Fisher and the Making of Maximum Likelihood," 162.

13. Box, *R. A. Fisher*, 75–76.

14. Stephen Stigler, "How Ronald Fisher Became a Mathematical Statistician," *Mathematics and Social Sciences* 44 (2006), 23–30; R. A. Fisher, "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society A* 222 (1922), 309–368.

15. Stigler, "How Ronald Fisher Became a Mathematical Statistician."

16. Box, *R. A. Fisher*, 49.

17. Box, *R. A. Fisher*, 51–52.

18. R. A. Fisher, *Statistical Methods for Research Workers* (Edinburgh: Oliver and Boyd, 1925).

19. Box, *R. A. Fisher*, 130–131; see Diane Paul and Barbara Kimmelman, "Mendel in America: Theory and Practice, 1900–1919," in *The American Development of Biology*, R. Rainger, K. Benson, and J. Maienschein, editors (New Brunswick, NJ: Rutgers University Press, 1988), pp. 281–310.

20. R. A. Fisher, "The Correlation of Relatives on the Supposition of Mendelian Inheritance," *Royal Society of Edinburgh* 52 (1918), 399–433, 399.

21. R. A. Fisher, "On the Dominance Ratio," *Proceedings of the Royal Society of Edinburgh* 42 (1922), 321–341, 334.

22. R. A Fisher, *The Genetical Theory of Natural Selection* (Oxford: Oxford University Press, 1930). Released in a revised second edition in 1958 from Dover Publications, and released in a variorum edition by Oxford University Press in 1999 edited by J. H. Bennett; J. B. S. Haldane, *The Causes of Evolution* (London: Longmans, 1932); Sewall Wright, "Evolution in Mendelian Populations," *Genetics* 16 (1931), 7–159; Sewall Wright, "The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution," *Proceedings of the Sixth Annual Congress of Genetics* (1932), 356–366.

23. Fisher, *Genetical Theory*, 37.

24. Warren Ewens, "An Interpretation and Proof of the Fundamental Theorem of Natural Selection," *Theoretical Population Biology* 36 (1989), 167–180; G. R. Price, "Fisher's 'Fundamental Theorem' Made Clear," *Annals of Human Genetics* 36 (1972), 129–140.

25. Fisher, "On the Dominance Ratio," 334.

26. Fisher, *Genetical Theory*, 11.

27. Provine, *The Origins of Theoretical Population Genetics*; Jean Gayon, *Darwinism's Struggle for Survival: Heredity and the Hypothesis of Natural Selection* (Cambridge:

Cambridge University Press, 1998); Jonathan Hodge, “Biology and Philosophy (Including Ideology): A Study of Fisher and Wright,” in S. Sarkar, editor, *The Founders of Evolutionary Genetics* (Dordrecht: Kluwer, 1992), pp. 231–293.

28. Box, *R. A. Fisher*.

29. Stigler, “How Ronald Fisher Became a Mathematical Statistician”; Mackenzie, *Statistics in Britain*.

30. Paul and Kimmelman, “Mendel in America.”