Available online at www.sciencedirect.com
SCIENCE DIRECT®
Stud. Hist. Phil. Biol. & Biomed. Sci. 37 (2006) 59–82
Studies in History and Philosophy of Biological and Biomedical Sciences
www.elsevier.com/locate/shpsc

What was Fisher's fundamental theorem of natural selection and what was it for?

Anya Plutynski

Department of Philosophy, University of Utah, 260 South Central Campus Drive, Orson Spencer Hall, Salt Lake City, Utah 84112, USA

Received 29 March 2004; received in revised form 27 June 2005

Abstract

Fisher's 'fundamental theorem of natural selection' is notoriously abstract, and, no less notoriously, many take it to be false. In this paper, I explicate the theorem, examine the role that it played in Fisher's general project for biology, and analyze why it was so very fundamental for Fisher. I defend Ewens (1989) and Lessard (1997) in the view that the theorem is in fact a true theorem if, as Fisher claimed, 'the terms employed' are 'used strictly as defined' (1930, p. 38). Finally, I explain the role that projects such as Fisher's play in the progress of scientific inquiry.

Keywords: Fundamental theorem; Natural selection; Average effect; Average excess; Genetic variance; Malthusian parameter

1. Introduction

> No practical biologist interested in sexual reproduction would be led to work out the detailed consequences experienced by organisms having three or more sexes; yet what else should he do if he wishes to understand why the sexes are, in fact, always two? The ordinary mathematical procedure in dealing with any actual problem is, after abstracting what are believed to be the essential elements of the problem, to consider it as one of a system of possibilities infinitely wider than the actual, the

E-mail address: plutynski@philosophy.utah.edu (A. Plutynski).

1369-8486/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsc.2005.12.004

> essential relations of which may be apprehended by a generalized reasoning, and subsumed in a general formula, which may be applied at will to any particular case considered. (Fisher, 1930, p. ix)

This brief description of his method encompasses what is both most admired and (more often) disliked in Fisher's work. Fisher proceeded in almost every problem he treated on a level of abstraction that simultaneously made otherwise intractable problems soluble, and otherwise simple claims incomprehensible. Fisher was trained as a physicist; he was used to solving problems involving idealized gases, infinite populations, and other 'weightless elephants upon frictionless surfaces', as a commentator on one of his early papers put it (Punnett, 1916). This made his approach to biology both innovative and yet often too abstract to have the impact for which he had hoped. Debates over the interpretation of Fisher's fundamental theorem of natural selection—both what Fisher intended to demonstrate, and what was in fact demonstrated—continue to this day (Lessard, 2000; Ewens, 1989; Price, 1972). Many of the mathematical reconstructions of his theorem depart in different degrees from the original 'derivation', which was characteristically elliptically stated:

> *The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time*. (Fisher, 1930, p. 37, emphasis original)

While there have been many different mathematical reconstructions of the theorem, there has been to date no attempt to determine how the theorem fits in to Fisher's general project. Why did Fisher deem it necessary to derive this theorem? What problem—empirical, conceptual, or biological—was it intended to solve? In what sense was it 'fundamental'? Fisher compared the theorem to the second law of thermodynamics. Was this claim indicative of the typical bravado of this famously over-confident biologist? Or, ought the theorem indeed to have had the same import in the history of biology as the second law did in physics? On what grounds ought the theorem to have such import, if, as so many biologists and philosophers have claimed, it strictly speaking holds true for a range of conditions that are never instantiated?

Attempts at deriving general theorems in biology, such as those made by Fisher in his 1930 book, are often the source of conflict in the biological sciences. In contrast, in the physical sciences, such general projects are not universally regarded with suspicion. Why is this the case? What is it about the science of biology that makes the enterprise of deriving general laws less than respectable?

I argue that Fisher's theorem, correctly understood, is true (this has been most recently argued by Lessard, 2000, and Ewens, 1989). Indeed, it is necessarily true, given the sense of 'partial change in mean fitness' and 'genetic variance' that Fisher intended, and the restrictive conditions he described. Of course, this is what it means for something to count as a theorem. There is an important point to be made here about the role of 'theorems' as opposed to laws in science. The debate concerning the fundamental theorem is not over whether it is predictively accurate or supports counterfactual conditionals. Rather, the debate is over whether and how the theorem explains something fundamental about how selection works. Fisher himself noted that the restrictive conditions he sets out are never met in the actual world. Nonetheless, it is a true *theorem*. Why ought one to be concerned about a 'fundamental' relationship that holds only under idealized conditions? Why does it matter that the additive genetic variance is proportional to the partial change

in mean fitness, given all the caveats necessary to derive this result? There are two reasons. First, for Fisher, finding such a relationship was enormously significant to his lifelong projects of vindicating the theory of natural selection, and reconciling Mendelism and Darwinism in a rigorous mathematical theory of evolution. Unfortunately, by the time he derived the theorem, these projects were well underway. Indeed, I will suggest that this is part of the reason his accomplishment may not have been as valued as he had hoped.

Second, setting out such mathematical relationships, however idealized, helps us understand a basic fact about the conditions on the possibility of evolutionary change via natural selection. Natural selection requires a 'reservoir' of additive genetic variance. The project of deriving general theorems—or, rather, finding relationships between different measures (population size and the effects of drift, for example), however idealized, is and should be properly part of the biological sciences no less than the physical sciences. This is a fact to which philosophers of biology ought to be more attentive in their analysis of the aim and practice of biology (see also Sober, 1984, on what he calls 'source' and 'consequence' laws).

Idealized models or theorems, such as the fundamental theorem, may serve to explain the conditions on the possibility of evolution. Only by abstracting away from the multiple and often multiply interacting causal pathways involved in evolutionary change can we arrive at a general characterization evolution, or, for example, how selection depends upon available variation. Haldane has written, in defense of 'beanbag genetics':

> In the consideration of evolution, a mathematical theory may be regarded as a kind of scaffolding within which a reasonably secure theory expressible in words may be built up. I ... show that without such scaffolding, verbal arguments are insecure. (Haldane, 1964, p. 350)

Given certain initial conditions, classical population geneticists can, in Haldane's words, 'deduce the evolutionary consequences of these facts' (1964, p. 343). In other words, the work of classical population genetics is the making precise of verbal claims about evolutionary dynamics by construction of mathematical models and deductive demonstration of the consequences of those models. Without these demonstrations, we would not know, generally speaking, how evolution works.

2. Historical context of the theorem

To understand the significance of the fundamental theorem for Fisher, it is important to understand the scientific context he entered at Cambridge in 1909. The kind of science Fisher did and the reason he believed his theorem to be significant is a product of his ancestry: the view of science he inherited from Pearson, Galton, and Darwin. His vision for biology was informed by the biometrical and Mendelian traditions, and his lifelong project was to synthesize these traditions and show how they were not only compatible with, but also supported, a Darwinian view of life. Thus, it is important to briefly trace this history.

While at the University, Fisher excelled in mathematics, but also developed interests in a wide range of subjects, including physics, astronomy, and especially statistics. It was toward the end of his schooling that Fisher took up the subject of biology. Fisher's reading of Pearson's 'Mathematical contributions to the theory of evolution' (1894), and his exposure to Punnett, Bateson and Mendelism, were each to have lasting influence. Fisher's aims

and methods as a scientist are in some sense continuous with those of scientists like Galton, Pearson, Weldon, Punnett, and Bateson. These men founded the schools of biometry and Mendelism, which were very much at odds at the turn of the century, exactly when Fisher first encountered the theories of heredity and evolution.

2.1. Biometry vs. Mendelism

Francis Galton (1822–1911) had an important influence on twentieth-century theories of heredity and evolution. Paradoxically, his ideas served as the template for both the biometrical and Mendelian schools of thought, which were at odds at the turn of the century over whether evolution was Darwinian—the result of gradual selection on continuous characters, or discontinuous—consisting in 'jerks' or major shifts from one 'stable' state to another. The paradox of this dual influence may in part be resolved by a close look at tensions in Galton's own thought on heredity and evolution via natural selection. In the 1890s, Galton argued that Darwin's conception of selection acting on small individual differences could not permanently alter a 'race'. On the other hand, Galton was a great champion of Darwin, especially insofar as evolution seemed to support eugenics. How could he consistently criticize gradual selection and espouse eugenics?

Galton believed that there were two types of variation: 'individual' variation, which yields the normal distribution of traits in a population, and discrete variation, or 'sports'. Galton thought that the only way in which eugenics could modify the human race was by analogy with artificial selection. Artificial selection at that time consisted primarily in the selective breeding of sports. Galton believed that, by analogy with artificial selection, evolution must be a process of discontinuous 'transilience' from one 'position of organic stability' to another:

> No variation can establish itself unless it be of the character of a sport, that is, by a leap from one position of organic stability to another, or as we may phrase it, through 'transilient' variation. If there be no such leap the variation, so to speak, a mere bend or divergence from the parent form, towards which the offspring in the next generation will tend to regress; it may therefore be called a 'divergent' variation. I am unable to conceive of the possibility of evolutionary progress except by transiliences. (Galton, 1894, p. 368)

Bateson and DeVries, among others, were to take up these ideas of Galton's and initiated a Mendelian school of thought with respect to the major causes of evolution. This gave rise to a divide between two schools of thought, and a dramatic struggle at the turn of the century as to the mode of evolution.

Galton and Weldon formed a committee of the Royal Society whose aim was to make 'statistical inquiries into the measurable characters of plants and animals' (Weldon, 1895a, p. 360, cited in Gayon, 1998, p. 201). The group was subsequently renamed the 'Evolution Committee of the Royal Society', though ironically, what was meant by the term 'evolution' was a hotly contested issue among several members of the Committee. With the exception of Galton himself, the founding members of the Committee—Weldon, Pearson, and Poulton—were all adherents of the gradualist conception of evolution. Building upon the statistical tools Galton had invented (such as correlation and regression), Weldon and Pearson began investigating the following questions: What is the distribution of characters in a population? Can variation in a specific range be correlated with death rate? These

empirical investigations grew alongside the science of statistics, which, at that time, fell under the heading 'biometry'.

Weldon and Pearson were the main co-founders of the science of biometry. Weldon's empirical investigations provided Pearson much of the impetus to treat biological problems statistically. Pearson brought to the collaboration an instrumentalist, proto-positivist philosophy of science, which was to shape biometry and influence those who followed its methods and aim (notably Fisher). That aim was to demonstrate Darwin's theory of evolution by natural selection empirically. However, Weldon and Pearson's theory of demonstration was unique. Theirs was a 'purely statistical' or phenomenological project—both refused to speculate on causes (hereditary, physiological or functional), instead focussing their inquiry entirely on correlations. The biometricians stressed exact measurement and description, without theory, of the observable phenomena of evolution. Correlations could be measured and described; causes could not. Heritability, for Pearson, was a matter of correlation:

> The true measure of heredity is the numerical correlation between some characteristic or organ as it occurs respectively in parent and offspring. (Pearson, 1897, p. 65)

Weldon adopted a similar philosophy. In responding to R. Lankaster's criticism that he had not given any reason *why* the frontal breadth of the crab showed a selective destruction, (i.e. what the functional significance might be), Weldon cited lengthy passages from Hume emphasizing that it was not the role of the scientist to speculate on causes, but only to describe correlations. In his 'Remarks on variation in animals and plants' (Weldon, 1895b), directly following his paper on crabs, Weldon sets himself apart from those 'naturalists' (the Mutationists, such as DeVries) who make 'assumptions' about the value of major variations without systematically measuring the effects of small variations. He refused to make any such assumptions, either about the physiological basis of variation, or the functional importance of specific traits (1893, p. 371).

Weldon wrote in 1893:

> It cannot be too strongly urged *that the problem of animal evolution is essentially a statistical problem*: that before we can properly estimate the changes at present going on in a race or species we must know accurately: (a) the percentage of animals which exhibit a given amount of abnormality [i.e., variation from the mean] with regard to a particular character; (b) the degree of abnormality of other organs which accompanies a given abnormality of one; (c) the difference between the death rate per cent. In animals of different degrees of abnormality with regard to a particular character; (d) the abnormality of offspring in terms of the abnormality of parents, and *vice versa*. These are all questions of arithmetic; and when we know the numerical answers to these questions for a number of species we shall know the direction and the rate of change in these species at the present day—a knowledge which is the only legitimate basis for speculations as to their past history and future fate. (Weldon, 1893, p. 329, my italics)

Weldon's aim was to set out the measurable conditions for testing a case of natural selection and to determine where those conditions existed in the natural world. For this, his only requirements were determining the correlation between parent and offspring in the measure of a specific character, and the correlation of the 'degree of abnormality' of that character with death rate. One can see the influence on Fisher's thinking. For Fisher,

seeing evolution as an essentially statistical problem was enormously attractive, exactly because it was quantitative and had a ready analogue with the physical sciences. To defend Darwinism and make evolution a science, one must develop a quantitative model of the essential or 'fundamental' processes underlying evolution, by analogy with statistical mechanics.

A direct predecessor to the fundamental theorem was Pearson's 'Fundamental theorem in selection' (1898). Pearson's theorem was a multiple correlation equation that enabled one to use tables of correlation to predict the composition of a population if some individuals were prevented from reproducing. However, the theorem was not very fruitful, since it required vast tables of genealogical and demographical data that were as yet not available. Also, it was not a theoretical principle so much as a way of summarizing and describing hereditary data. It was purely phenomenological, based on data at the phenotypic level. Biometry lacked a theory of heredity. Pearson and Weldon refused to speculate on the physiological causes of heredity, and dealt only with the measurable effects of heredity and selection. Although Pearson built significantly upon the statistical foundation provided by Galton, without a theory of heredity, biometry's scope was relatively limited.

2.2. The mutationists

In contrast, the Mendelians, or 'mutationists', posited a particulate theory of inheritance. DeVries and Johannsen claimed to have found evidence that evolution proceeded via selection on major mutations, or 'discontinuous' variation. After the rediscovery of Mendel's 1865 paper in 1900, Bateson, Punnett, and DeVries claimed Mendel as their predecessor, and thus came to be called the 'Mendelian' school. The Mendelians held that variation was of two sorts, 'fluctuating' and 'discontinuous', and that it was the latter that was important for evolution. Evolution must ultimately rely on mutation as the source of new variation, where by mutation Bateson meant new genetic factors that caused genuinely novel characters, or major morphological shifts. However, the biometricians rejected Mendel, ostensibly because it was unclear how a Mendelian system of inheritance could account for continuous heredity.[1]

This was the environment Fisher entered as a student at Cambridge. Fisher developed an interest in heredity and evolution in part because of his interests in eugenics. Norton has written, 'eugenics was the tail that wagged the dog of population genetics' (1983, p. 19) and this is especially true for Fisher, as the final chapters of the Genetical Theory of Natural Selection attest, with subheadings including "*The decay of civilizations*" (Fisher, 1930, p. 174). Fisher was concerned to the end of his life with what he understood to be one of the greatest threats to modern civilization, the greater contribution that 'less fit' individuals make to the human gene pool (i.e., genetically unfit individuals having more

[1] It is important to note that the divide between the biometricians and Mendelians was essentially a British phenomenon. Marga Vicedo (1992) has demonstrated that the antagonism between biometry and Mendelism did not exist in the United States. In fact, American agriculturists and scientists such as Davenport, Castle, and East adopted both the statistical methods of biometry and the Mendelian understanding of heredity as complementary. In the case of the agriculturists or 'breeders' this was largely a matter of pragmatics—they found the methods of both schools to be extremely effective in perfecting crops. In the case of the scientists, they saw that both biometry and Mendelism were at the cutting edge of the new science of biology, and believed the experimental pursuit of the consequences of both theories to be essential to the advance of biology.

children than the purportedly more fit). His career in genetics began with joining the eugenics society as an undergraduate, and some of the talks he presented to that society are uncompromising in their suggestions for a policy that encourages the better sort to reproduce through various public incentives, or more problematically, prevents the worse sort from doing so:

> Suppose we knew, for instance, twenty pairs of mental characters. These would combine with over a million pure mental types, each of these would naturally occur rather less frequently than once in a million, or in a country like England may occur in 20,000 generations; it will give some idea of the excellence of the best of these types when we consider that the Englishmen from Shakespeare to Darwin have occurred within ten generations; the thought of a race of men combining the illustrious qualities of these giants, and breeding true to them, is almost too overwhelming, but such a race will inevitably arise in whatever country first sees the inheritance of mental characters elucidated.A large number of rare defects among men are now known to be Mendelian dominants, colour blindness, Brachydactyly and the form of insanity known as chorea are among these; the inheritance of these is easily traced, since half the offspring of any affected person will be affected ... These would all be stamped out in one generation by prohibiting affected persons from pairing. (Transcript of Fisher's talk 'Heredity' at the Cambridge University Eugenics Society, Norton & Pearson, 1976, pp. 158–159)

Notice that in the above, Fisher is already employing mathematical and statistical thinking in the service of eugenics. The influence of Pearson and the biometricians is clear. Fisher adopted a 'statistical viewpoint' in the following sense. His aim was to give a quantitative statistical representation of the conditions necessary for selection to operate in a Mendelian population. What distinguished Fisher was his willingness to endorse the Mendelian model of inheritance and consider whether that might underpin the biometrical observations. Fisher was less encumbered by the biometricians' scruples. What formed the core of Fisher's research throughout his life (eugenic motivations notwithstanding), was the attempt to vindicate Darwinism—or, to supply a formal theory of the operation of selection compatible with Mendelism.

3. Fisher's research program

3.1. 1918 and before

Fisher's 1918 paper 'On the correlation between relatives on the supposition of Mendelian inheritance' is an attempt to demonstrate how a multifactorial theory of inheritance can serve to underpin observed correlations between relatives discovered through biometrical methods. It is widely regarded as the seminal paper that successfully synthesized biometry and Mendelism. Biometricians such as Galton and Pearson found first that the vast majority of continuous traits (such as height or weight) exhibit a normal distribution, or if not, can be broken down into sums of two or more normal distributions. Second, they discovered correlations between relatives that decrease by one-half relative to ancestors.

Fisher demonstrated that, given a number of assumptions, the observations of biometry could be made compatible with a Mendelian, 'particulate' scheme of inheritance. The 1918

paper illustrates Fisher's persistent method of beginning with a great deal of abstraction, and then gradually showing how, even with the added complication of biological reality, the effects still 'average out'. It is an enormously difficult and abstract mathematical paper, and one that struggled to reach the light of day. Fisher withdrew the paper after unfavorable reviews from Pearson and Punnett when it was first submitted in 1916 to the Royal Society, and it had later to be published under the sponsorship of Major Leonard Darwin in the *Transactions of the Royal Society of Edinburgh*.[2]

Fisher entered the dispute between the biometricians and Mendelians with faith both in the power of selection and the explanatory power of Mendelism. Fisher demonstrated how the observed correlations between various pairs of relatives—father–son, cousin–cousin, sibling–sibling, among others—can be explained on the presupposition that traits such as height are derived from a multitude of Mendelian, or discrete, independent factors. Fisher's demonstration was possible because he relaxed some of the assumptions that others were unwilling to give up (for more detail, see Moran & Smith, 1966; Sarkar, 1998).[3]

[2] Fisher was not the first to address the compatibility of biometry and Mendelism. In 1902, Yule had shown that it was possible for normally distributed characters to be susceptible to a Mendelian explanation (Yule, 1902). Pearson subsequently addressed the same question in a paper published in 1904, his twelfth 'Mathematical contribution to the theory of evolution' titled 'On a generalized theory of alternative inheritance, with special reference to Mendel's laws'. Pearson was able to show that on the Mendelian assumptions of some trait depending linearly and additively on n Mendelian loci, the distribution of stature if n was large would be close to the normal, and there would be a linear regression between relatives. However, the value predicted for the correlation of son and father for Pearson's model was 0.333, which disagreed with the observed correlation which was in the range of 0.45–0.5. Thus, Pearson concluded that the Mendelian model was 'not sufficiently elastic to cover the observed facts' (cited in Norton, 1975), or, the laws of Mendelism could not serve to underpin or explain the observations of biometry. Pearson thought that while there was no 'essential repugnance' (cited in Norton, 1975) between the results of biometry and Mendelism, there seemed to be a lack of compatibility between the two.Yule replied in 1906 in a paper addressed to the Third International Congress of Genetics that Pearson had made an important assumption in his paper that could reasonably be suspended. Pearson's model assumed complete dominance. With intermediate or no dominance, Yule was able to show that the correlation between relatives would rise to 0.5, and thus 'the theory of the pure gamete, as applied to compound characters is much more flexible than would appear from Professor Pearson's work, and can hardly be summarily dismissed as inapplicable to cases in which the coefficients of correlation approximate .5' (Yule, 1906). Pearson resisted Yule's assessment; his claim was that complete dominance was one of Mendelism's own tenets, and until the Mendelians were willing to give it up, clearly Mendelism would be incompatible with the results of biometry. In 1909, Pearson himself showed how the anti-Mendelian results he reached in 1904 depended upon a number of assumptions whose modification would make the two views compatible. Namely, only if one allowed a Mendelian schema with no dominance, or if the heterozygote was exactly intermediate in phenotype between the two homozygotes, would the discovered correlation between father and son result. However, Pearson thought that the assumption of dominance was essential to Mendelism, and as long as this was retained, the two schools were incompatible.Pearson had philosophical objections to the very suggestion of Mendelism, as is made evident in his positivistic *Grammar of science* (Pearson, 1957). Here he is quite explicit that the object of science is description, and that the positing of unobservables was unacceptable. Moreover, Pearson objected to the very idea of positing *causes* acting in the world. All one could do as a respectable scientist was observe and document correlations: 'No phenomena are causal; all phenomena are contingent, and the problem before us is to measure the degree of this contingency' (Pearson, 1957, p. 174). In other words, Pearson had metaphysical objections to the Mendelian school insofar as they suggested unobservable particulate causes as giving rise to the phenomena of heredity. Thus Pearson was unwilling to accept Mendelism, despite his own mathematical demonstration of its compatibility with the results of biometry. It is ironic that the very same man who rejected Fisher's paper in 1916 had shown in his own work that the difficulties of explaining continuous variation on a Mendelian system of inheritance could be overcome.

[3] For a discussion, see Morrison (2002).

Given, however, that Yule and Pearson had themselves demonstrated that the correlations between relatives were derivable from a Mendelian system of inheritance, what exactly was Fisher's contribution to the biometrician–Mendelian debate? In my view, Fisher did not so much add to the number of facts explained as reconceive the object of explanation. A population of organisms was, in a sense, reduced to a cloud of points with no interactive forces and independent actions. The normal distribution and the correlations between relatives follow from these assumptions. He thus shows how a very attenuated sense of Mendelism is compatible with the results of biometry (Sarkar, 1998). Fisher's work in the 1918 paper defies the notion of scientific progress as mere accumulation of facts or laws. Sometimes an explanation consists not in providing new laws or deductions, or in bringing to light new facts, but in rethinking what it is that requires explaining, or in synthesizing two research programs previously regarded as at odds, in however an idealized or abstract fashion.

Perhaps Fisher himself was aware of the very limited nature of his achievement, though of course we cannot know. Weak compatibility is hardly a grand unification in the spirit of physics like Maxwell or Boltzmann, Fisher's heroes. It was in the fundamental theorem that he sought a more rigorous demonstration of the effectiveness of selection and a quantitative relation between the presuppositions of Mendelism and the Darwinian view of life.

3.2. 1930 and beyond: Fisher's program

Fisher's 1930 book, *The genetical theory of natural selection*, is generally regarded as his central contribution to evolutionary theory. Many commentators have seen this book, more than others that contributed to the synthesis, as a treatise to the 'beanbag', 'genic', and 'panselectionist' approach to evolution (Gayon, 1998, pp. 325–327; Depew & Weber, 1995, pp. 249–250), according to which evolution is primarily a deterministic process, as if driven by natural selection acting on genes directly. Moreover, many see Fisher's fundamental theorem of natural selection as a statement that evolution is somehow 'progressive' (Ruse, 1997; Gayon, 1998): some have gone so far as to suggest that this was Fisher's solution to the problem of how to find purpose, or "Natural Selection was the immanent working of the Creative will through the medium of Natural Law" (Turner, 1985, p. 337), in an otherwise "random, purposeless world" (Turner, 1983, p. 159). While Fisher makes no secret of his view about the significance of natural selection for evolution, the text, properly understood, is by no means a tribute to *genic* selectionism *per se*. Fisher is not endorsing an argument for selection's direct action on genes. Nor is the fundamental theorem aimed at showing why evolution must be progressive, or fitness maximized, or so I will argue.

Fisher's main object in the *Genetical theory* was not to offer an empirical view of how evolution has in fact gone forward. He is not advancing an empirical theory, a 'Large Size' theory. While it is true that Fisher believed that evolution occurs for the most part in populations of large size, defending this claim was not his central focus. Fisher was first and foremost a theoretician, not an empirically minded biologist. Most central to Fisher's project was to show that it was possible for a gradualist, Darwinian model to explain the diverse phenomena we see in nature (dominance, sex ratio, etc.) His aim was to vindicate Darwinism and demonstrate its compatibility with Mendelism—indeed, its necessity given a Mendelian system of inheritance. He is explicit that he is advancing a mathematical theory and a set of arguments for how evolution 'must' proceed. Fisher wrote, 'my primary

job is to try to give an account of what natural selection *must be doing*, even if it had never done anything of much account until now' (**date, page number**). Working from a model of physics, Fisher's aim was to give a rigorous mathematical treatment to the relationship between the fundamental factors at work in evolution. He even suggests that we treat selection as a 'vector', akin to force functions in classical physics (1930, p. 79). He viewed his object as following out the mathematical consequences of a Mendelian scheme of inheritance for Darwin's theory of natural selection. In this way, Fisher hoped to answer long-standing objections to Darwin, and demonstrate that selection indeed was sufficient to explain the diversity of life.

In all this, he was following in the tradition of the biometricians, Pearson in particular. Fisher does not concern himself with the details concerning the 'why' of evolution, so much as the 'how', and the quantitative relationship between the different measures of change. Fisher's book, and the theorem in particular, is best understood as a continuation of his attempt to breach the divide between biometricians and Mendelians concerning the nature of heredity and the effectiveness of Darwinian selection. His motivation in almost all his work was to explain how it was possible to resolve this conflict, and to vindicate Darwinian selection as both a plausible and necessary cause of evolutionary change.

This is evident from the preface and first chapter of Fisher's book. The opening discussion provides an excellent overview of the problems and issues to be resolved in biology, as Fisher understood them. Fisher begins with a discussion of the problems encountered by Darwin and how they were resolved by the particulate theory (see Hull, 1973). In particular, he discusses Jenkins' (1867) (and, subsequently, Galton's) famous objection to Darwin, namely, that variation will be 'swamped' or revert to the mean in every generation after selection. He makes mention of his opponents, and of the theoretical and biological context to which he sees himself as replying. Moreover, he explicitly rules out explanations for the modification of species involving mutationist, Lamarckian or orthogenic principles. The text is an argument for the effectiveness of natural selection. There is no necessity, Fisher argues, given available knowledge about rates of mutation, to involve mutationism in evolution. This brief chapter is thus Fisher's articulation of his research program—his statement of where he saw himself in history and where he saw the important work in biology to be heading.

Significantly, the first subject Fisher discusses in his (1930) is Darwin's blending theory of inheritance. According to the blending theory, variation would be halved and the effects of selection would be swamped in every generation. Fisher explains why it was thus necessary for Darwin to assume that variation came from either the 'conditions of domestication' (food or environment) or use or disuse, in order for there to be new variation for selection to act upon. In contrast, according to the particulate theory of inheritance, variation is not lost. Viewing genes as singularly acting, independent factors thus solved the problem with Darwin's theory of natural selection that prompted Galton and the Mendelians to advocate mutationism.

Fisher believed that it was important to treat the frequencies of alleles in a particulate fashion in order to vindicate Darwin's theory. Fisher was not a literal believer in the view that selection acted on genes *per se*. Fisher was well aware that it was the organism as a whole that was facing the environment. Fitness is, after all, predicated of 'organisms' in his (1930), and 'natural selection' is acting on 'organisms', not genes (see p. 42, for example). In any case, it is anachronistic to read a Dawkins-style argument back into Fisher (whether or not Dawkins himself actually endorses this view, the 'genic' view of selection

is often attributed to him). The entire debate between 'genic' and 'organismal' selectionists was not one that Fisher or any of his contemporaries would have recognized as a problem. Changes in gene frequency were a measure of the effects of selection—it is the individual organism that faces the environment or that is adapted. What is relevant to the argument here is that Fisher saw the particulate theory as an important solution to a long-standing problem for the theory of evolution by natural selection, and an answer to Galton and the Mutationist school. Fisher comments:

> For any evolutionary tendency which is supposed to act by favouring mutations in one direction rather than another, and a number of such mechanisms have from time to time been imagined, will lose its force many thousand-fold, when the particulate theory of inheritance, in any form, is accepted; whereas the directing power of Natural Selection, depending as it does on the amount of heritable variance maintained, is totally uninfluenced by any such change. (Fisher, 1930, p. 10–11, see also p. 20)

Fisher is here carving out a distinction between mutationism and Mendelism. One can be a Mendelian, he explains, without being a mutationist. Moreover, one can be Darwinian and embrace the particulate theory. Indeed, a particulate theory is necessary, or so he claims, for evolution via natural selection to proceed.

Fisher elucidates the difference between the blending and particulate theories of inheritance via analogy with the kinetic theory of gases:

> The particulate theory of inheritance resembles the kinetic theory of gases with its perfectly elastic collisions, whereas the blending theory resembles a theory of gases with inelastic collisions, and in which some outside agency would be required to be continually at work to keep the particles astir.' (Fisher, 1930, p. 11)

Such comments should not induce us to attribute to Fisher the thought that genes are mere random associations of particles. From the above context, it is clear that Fisher uses the analogy with physics *only to emphasize the contrast of the particulate with the blending theory*. He is not arguing that in fact, genes are mere random clouds of particles, or that the theory of selection is a theory akin to statistical mechanics, but rather, that it is better to think of genes as having 'inelastic' as opposed to 'elastic' collisions! The analogy, in other words, is serving the narrow purpose of explicating how and why Mendelism, or the particulate theory, resolved one of the 'main difficulties felt by Darwin' (Fisher, 1930, p. 12).

Fisher viewed the theory of natural selection as the only remaining explanation for evolution once one has abandoned mutationism; essentially, his emphasis on natural selection was not because of some dogmatic insistence on the deterministic character of evolution (Turner, 1985), but because he sees Darwinism as only alternative, and thus by default the most progressive research program. Once the alternatives are eliminated:

> The sole surviving theory is that of Natural Selection, and it would appear impossible to avoid the conclusion that if any evolutionary phenomenon appears to be inexplicable on this theory, it must be accepted at present merely as one of the facts which in the current state of knowledge does seem inexplicable.' (Fisher, 1930, p. 21)

It is in contrast to the theories that Fisher saw as competing (Lamarckism, mutationism, orthogenesis), that natural selection is the best available explanation.

3.3. The derivation of the fundamental theorem

Fisher regarded the fundamental theorem of natural selection as a significant contribution to evolutionary theory. This held, he wrote, 'the supreme position in the biological sciences', analogous to the second law of thermodynamics (Fisher, 1930, p. 39). Nevertheless, this is one of the most contested of Fisher's claims. Starting with Wright's (1931) review of Fisher's *Genetical theory of natural selection*, mathematical biologists have been debating the meaning and significance of Fisher's theorem (Edwards, 1967; Kimura, 1958; Price, 1972; Ewens, 1989. Several commentators, starting with Wright's (1931) have taken W to represent the mean fitness of the population, and thus criticized the fundamental theory on the grounds that this fitness can decrease. Thus, Wright 'corrects' the theorem as follows:

> The total variance in fitness of a population is ascribable to the variance in fitness due to natural selection, which excludes the effects of dominance, epistasis, mutation, migration, change in environment, and drift. (Wright, 1931, p. 272)

However, the only real correction that would need to be made concerns mutation or change in environment. Subsequent commentators, and indeed the majority of textbooks in population genetics through the 1970s (for example, Crow & Kimura, 1970), interpreted Fisher's theorem along the same lines. The 'received' interpretation thus came to be that 'the increase in mean fitness of a population is approximately the current additive genetic variance in fitness, and this is non-negative' (Edwards, 1994). This takes the theorem to refer to the mean fitness of the population, and to be an approximate result. However, Lessard (2000), Ewens (1989), and, originally, Price (1972) have explained that the theorem refers not to the total change in mean fitness of the population, and moreover, that it is not a merely approximate result.

As with much of Fisher's work, in part this debate was due to the opacity of his own presentation of the theorem. Fisher's definitions of his terms are often obscure, and he himself does not provide a full derivation of the theorem. The not insignificant task of filling in the intermediate steps has been left to his interpreters. As we saw earlier, Fisher's statement of the theorem is as follows:

> *The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.* (Fisher, 1930, p. 37, italics original)

Careful attention to Fisher's intended meaning shows that this is a true theorem, but that its significance is perhaps more circumscribed than Fisher implies in his comparison with the law of entropy.

As Fisher's derivation of the fundamental theorem is very compressed (the derivation itself, as opposed to its explication, is only one page), it is worthwhile reproducing in full (see appendix). The text is obscure, but becomes clearer once one unpacks Fisher's terms. One way of understanding the fundamental theorem is as a statement about what will necessarily set the 'speed limit' of evolution by natural selection—namely, additive genetic variance. Alternatively, some suggest that it is an investigation into what and how natural selection acts to maximize some quantity in evolution (Gayon, 1998). As we will see, this is a mistaken interpretation. I support Ewens (1989) and Price (1972), who argue that the fitness to which Fisher refers is best understood as a 'partial change in mean fitness'. First, I will unpack the opening lines of the derivation:

> Any group of individuals selected as bearers of a particular gene, and consequently the genes themselves, will have rates of increase which may differ from the average. (Fisher, 1930, p. 37)

In other words, for any individual bearing one of several alternative alleles at a locus, their survival and fecundity will differ from the population on average, in that they will either be at an advantage (have a positive selective advantage), or at a disadvantage. Fisher continues:

> The excess over the average of any such selected group will be represented by *a*, and similarly the average effect upon *m* of introducing the gene in question will be represented by a. (ibid.)

Here Fisher appeals to two different quantities, the average excess (a) and average effect (α) of an allele substitution on the measure of some trait. In the *Genetical theory* Fisher explains that the average excess is the difference in average measurements of some phenotypic trait (such as height) due to a gene substitution. He explains that the average excess is measured by 'the population divided into two portions, each comprising one homozygous type together with half the heterozygotes, which must be divided equally between the two portions' (Fisher, 1930, p. 30) and simply subtracting the difference between the average height of the two sets. More precisely:

$$\text{Average excess} \quad a = \frac{\mathrm{Pi} + \mathrm{Qj}}{\mathrm{P} + \mathrm{Q}} - \frac{\mathrm{Qj} + \mathrm{Rk}}{\mathrm{Q} + \mathrm{R}}$$

where i, j, and k are the mean phenotypic measurements (such as of height) of the genotypes GG, Gg and gg; and P, 2Q and R are their frequencies in some population. Note that average excess is relativized to a particular population at a particular time. Average excess due to allele substitution is then the difference in measure of some quantitative trait in some population associated with that substitution. Fisher remarks that in most populations (except for the ideal case of a population breeding perfectly at random, or, in Hardy Weinberg ratios, or, by extension, when other loci are considered, in linkage equilibrium):

> This difference is not . . . to be ascribed to the single gene substitution, as it were necessarily an effect of it, and of it only; for by reason of such common phenomena as homogamy, or the mating of like with like, or a variety of similar causes, it may well be, and is probably the case in all real populations, that the two moieties, of which the averages are compared, differ not only in the one gene substitution selected, but in the frequencies of a number of other genes affecting the measurement in question. (1941, p. 53)

In other words, Fisher notes that the average excess of some allele substitution is not necessarily due to the substitution of that allele alone, but may depend upon other genes that travel in company with it, either because of assortative mating or linkage, for example. In other words, the *effect* is brought about by the replacement of an allele by another, while changing nothing else. The *excess* includes correlated changes at other loci or inbreeding, etc. The excess can be measured directly from observable quantities in the population. The effect requires something else—in practice, least squares. Fisher was aware that there are both linkage and epistatic effects between genes. This is exactly why Fisher makes the

distinction between average excess and average effect. To isolate the effect of a single allelic substitution, and rule out the effects of interaction, Fisher defines a different quantity, average effect (α).

Average effect (α) is the 'actual increase in the total measurement of a population, when, without change in environment and mating system, the gene substitution is brought about, as if by experiment'(ibid.). In other words, the average effect is the amount of difference produced in the measure of some trait in a population due to the allele substitution alone. Of course, one cannot assess this directly. Average effect is thus found by the partial regression of the measurement of 'of the population' on the numbers 0, 1 and 2 of some allele in each genotype. Picture a graph where the genotype is on the x-axis and the measure is on the y-axis. 'gg' 'Gg' and 'GG' on the x axis correspond to the positions 0, 1 and 2, or, when there are 0, 1 and 2 'G' alleles at a locus. On the y-axis, i, j and k correspond to the mean phenotypic measurements of individuals with these three genotypes. Average effect describes the slope of the regression. Fisher defines $\mu-\alpha$, μ, and $\mu+\alpha$ as the 'expected' values of the genotype. Thus, μ and α are determined by minimizing the sum of squares:

$$P(i-\mu-\alpha)^2+2Q(j-\mu)^2+R(k-\mu+\alpha)^2$$

This procedure, (minimizing the sum of squares), is used to rule out effects 'acting through the genetic environment', i.e., effects due to anything other than substitution of the gene in question. In other words, Fisher uses partial regression in order to rule out that part of the 'environment' which includes the rest of the genome. Fisher writes that 'α (average effect), unlike *a* (average excess), is a true average of the measurable differences i–j, and j–k, produced by substituting G for g in homozygotes and heterozygotes respectively' (1941, p. 55). Average effect is thus a sort of ideal quantity, in that it could never be directly measured but can only be estimated. It measures what the exact effect of a single allele substitution will be, absent all the complicating factors. More precisely:

$$\alpha=\frac{P(Q+R)(i-j)+R(P+Q)(j-k)}{PQ+QR+2PR}$$

Compare with:

$$a=\frac{P(Q+R)(i-j)+R(P+Q)(j-k)}{(P+Q)(Q+R)}$$

Note that average effect and average excess will only be equal if there is perfectly random mating, or when $PR=Q^2$ (Fisher, 1941). According to Fisher, this is because if mating is random, linkages and interactive effects will be broken up in every generation, such that the exact effect of substitution of an allele will soon equal its effect as measured in a real population.

Fisher also introduces a third quantity, *m*, or the Malthusian parameter. *m* measures the rate of survival and reproduction in a population for individuals possessing some genotype.[4] Fisher explains that 'any group of individuals selected as bearers of a particular gene, and consequently the genes themselves, will have relative rates of increase different

[4] Or, more precisely, *m* will solve the integral from 0 to infinity of $e^{-mx}l_xb_xdx=1$, where l_x = probability of survival from birth to age x, and b_x = rate of reproduction at age x of individuals with that genotype.

from the average' (Fisher, 1958, p. 37). m measures this rate of increase in numbers of individuals different from the average. The Malthusian parameter is thus 'the relative rate of increase [of the frequency of some gene] in a population' (Fisher, 1930, p. 26). In other words, it represents whether and how fast a population of bearers of some gene is increasing in numbers; if m is negative, the population size is decreasing, if positive, the population size is increasing. Fisher continues:

> Since m measures fitness by the objective fact of representation in future generations, the quantity
>
> $$\sum{}'(2pa\alpha)$$

will represent the contribution of each factor to the genetic variance in fitness. The total genetic variance in fitness being the sum of these contributions, which is necessarily positive, or, in the limiting case, zero. (Fisher, 1930, p. 37)

What is the 'total genetic variance in fitness'? Recall that α is a measure of the average excess, or the measure of a given gene substitution in some population (inclusive of effects due to association of genes or assortative mating), and a is a measure of average effect, where this is the precise effect on fitness due to substitution of a single allele. The total genetic variance is thus a summation over all changes in the genetic composition of a population that yield additive differences in some trait, or what is today called the additive genetic variance in fitness.[5] To continue, Fisher writes:

> Moreover, any increase dp in the frequency of the chosen gene will be accompanied by an increase $2a\,dp$ in the average fitness of the species, where a may, of course, be negative. But the definition of a requires that
>
> $d/dt \log \mathrm{p} = \mathrm{a}$
>
> or $\quad dp = (pa)dt$
>
> hence $\quad (2\alpha)dp = (2\mathrm{pa}\alpha)dt$

which must represent the rate of increase of the average fitness due to the change in progress in frequency of this one gene. Summing for all allelomorphic genes, we have

$$dt\sum{}'(2pa\alpha)$$

and taking all factors into consideration, the total increase in fitness is

$$\sum \alpha dp = dt \sum\sum{}'(2pa\alpha) = \mathrm{W}\ dt.$$

If therefore the time element dt is positive, the total change of fitness Wdt is also positive, and indeed the rate of increase in fitness due to all changes in gene ratio is exactly equal to the genetic variance of fitness W which the population exhibits. We may consequently state the fundamental theorem of Natural Selection in the form: *The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.* (Fisher, 1930, p. 37)

[5] Fisher didn't like to call this additive variance. He was, however, happy to call it the genic variance, to distinguish it from the genotypic variance. (Thanks to Jim Crow for this clarification.)

The rate of increase in frequency of a particular gene over time, dp/dt, is, by definition, proportional to the increase in fitness of a population due exactly to the substitution of the particular gene in question. $(2pa\alpha)dt$ thus necessarily represents the rate of increase in fitness due to the substitution of this gene. Fisher simply sums over all such cases, such that the total increase in fitness is proportional exactly to the effects of gene substitution.

Here is where the controversy over Fisher's theorem usually begins. Surely the mean fitness of any population can always decrease? In other words, variance is always positive, and change in mean fitness can be negative. So, surely Fisher's theorem is false? With Price (1972) and Ewens (1989), I take it that by 'rate of increase in fitness' Fisher did not intend the mean fitness of the whole population. Rather, Fisher was referring in the fundamental theorem to what Ewens calls a 'partial' change in mean fitness, or that part of fitness due to changes in the 'gene ratio' alone. Ewens (1989) defines the 'partial change in mean fitness' as:

$$\sum\sum(P'_{ij} - P_{ij})(w + \alpha_i + \alpha_j)$$

where P_{ij} is the frequency of some genotype A_iA_j, w is the mean fitness of a population, and α_i and α_j are the average effects due to substitution of one or another allele. Ewens explains that:

> since changes in gene frequency are the substance of evolution as viewed by Fisher, and, since the average effect of A_i is a_i, the fitness of genotype A_iA_j in the sense of contributions to future generations is thought of as being not w_{ij} but rather $(w + a_i + a_j)$. (Ewens, 1989, p. 170)

In other words, according to Ewens, the change in fitness in some population that Fisher is concerned with is that proportion of the change that is directly due to change in gene ratio, not the total average fitness. Fisher is not concerned with average fitness changes, but changes in fitness that are proportional directly to changes due to the average effect of a gene substitution. And so, change in fitness will necessarily be proportional to just that part of the genetic variance which is additive and thus 'visible' to selection.

Further evidence for this interpretation comes in the pages immediately following the derivation of the fundamental theorem (Fisher, 1930, pp. 40–46). Fisher points out that *total* fitness of a population can decrease over time due to a number of factors both separately and in interaction: dominance and epistatic effects between genes, the 'deterioration' of the environment, and overpopulation. He writes:

> In addition to genetic variance . . . a second element comprised in the total genotypic variance . . . this component, ascribable to dominance, is also in a sense capable of exerting evolutionary effects, not through any direct effect on the gene ratio, but through its possible influence on the breeding system. (Fisher, 1930, p. 40).

Further, Fisher comments on rapid changes in the environment, both genetic (mutation), or geological and climatological, that may cause a decrease in the overall fitness of a population: 'Against the action of Natural Selection in constantly increasing the fitness of every organism . . . is to be set off the very considerable item of the deterioration of its inorganic and organic environment' (ibid., p. 45). And, he notes 'an increase in numbers of any organisms will impair its environment' (ibid.). That is, overcrowding may cause a

deterioration of overall population fitness. He even goes so far as to quantify the effects of the deterioration of the environment on the rate of increase of a population.[6]

This seems ample evidence that Fisher's derivation was restricted to the partial change in mean fitness, rather than the fitness of the population as a whole. Further, Fisher acknowledges that the theorem is exact 'only for idealized populations, in which fortuitous fluctuations in genetic composition have been excluded' (ibid., p. 38). Such 'fortuitous fluctuations' are due to reduction in population size. In other words, he acknowledges that the theorem does not hold unless one rules out the effects of drift. And, he follows the above with an estimate of the standard of error for his theorem (under random mating), given the effects of drift. Moreover, Fisher makes clear that it also must be the case that there are fixed fitness values, no mutation, no fertility differences, and no geographic structure to the population. Of course, we know none of these conditions hold in real populations.[7]

The fundamental theorem is not a statement about the unending or necessary adaptation of the species to its environment, but expresses a fundamental relationship between the reservoir of genetic variation available for selection, and the rate of increase in fitness in a population, *ceteris paribus*. Fisher was well aware that genetic interactions, rapid changes or 'deterioration' in the environment, or overpopulation, could affect whether or not a population of organisms would increase in number or continue to adapt over time. The fundamental theorem is thus not a statement of the necessary improvement of the species, but about the relation between genetic variance in some trait and increase in numbers of individuals possessing such a trait.

4. Why so fundamental?

This conclusion seems, however, rather mundane. Why did Fisher regard his theorem as so very 'fundamental'? The answer is that the fundamental theorem was a culmination of Fisher's lifelong project to vindicate Darwinism and unify the biometrical gradualist model of evolution and Mendelism in a rigorous mathematical theorem analogous to the physical sciences. After his derivation of the theorem, Fisher notes the following:

> The statement of the principle of natural selection in the form of a theorem determining the rate of progress of a species in its fitness to survive (this term being used for a well-defined statistical attribute of the population), together with the relation

[6] Fisher uses the following differential equation to represent the relationship of 'm', or the Malthusian parameter, to the other factors at work in effecting the gene ratio (deterioration of environment, etc.):

$$dM/dt + M/C = W - D$$

where 'M is the mean Malthusian parameter, C is a constant expressing the relation between fitness and population increase . . . W is the rate of actual increase in fitness determined by natural selection, and D is the rate of loss due to deterioration of the environment' (Fisher, 1930, p. 46). In other words, the rate of increase in M, or the rate of growth of a population, will be roughly equal to the rate of increase in fitness in a population due to natural selection, minus that effect due to the deterioration of the environment, where that deterioration is, as we have seen, due to both the genomic and external environment.

[7] The theorem does not, however, require random mating (as some have suggested), and, as should be clear from the above discussion of the distinctions between average excess and average effect. Fisher introduces this distinction exactly because he is aware of the differences in phenotype that may be due to factors other than allele substitution alone. If mating were random, then he might have used average effect throughout and not considered average excess, since they are equal with perfectly random mating.

> between this rate of progress and its standard error, puts us in a position to judge of the validity of the objection which has been made, that the principle of selection depends on a succession of favourable chances . . . It is easy without any very profound logical analysis to perceive the difference between a succession of favorable deviations from the laws of chance, and on the other hand, the continuous and cumulative action of these laws. It is on the latter that the principle of Natural Selection relies. (Fisher, 1930, p. 40)

As we can see, the theorem was for Fisher not simply a quantitative relationship between abstract properties of populations of organisms, but rather an answer to objections to the possibility that natural selection genuinely lead to adaptive evolution. Recall that in 1930, laypersons and scientists alike were still, to some extent, skeptical of the power of natural selection. Evolution was simply due to a 'succession of favorable chances'. (This same objection is still made today). Fisher thought he had shown that this was not so, but rather that Darwinian evolution had been vindicated, insofar as he had shown that all that was necessary for evolution to go forward was that there was a reservoir of genetic variation for selection to act upon.

Thus, according to Fisher the key to evolution by natural selection was genetic variance. Fisher saw the theorem as one of his central contributions to evolutionary theory, and it shares the characteristics of both the biometry and the physical sciences which Fisher wished to emulate. He was continuing in the tradition of Weldon and Pearson, for whom the problem of animal evolution was essentially statistical. Fisher's project was to establish necessary relations between fundamental properties of the natural world. He aimed for universal generalizations in the style of Newton, or better, Boltzmann, as his theorem treated aggregates whose interaction involved elements of chance. And, Fisher was working also very much in the tradition of Pearson and Galton before him: the statistical, biometrical method of beginning with the phenomena, abstracting away from these, and deriving statistical relationships between these abstract characterizations.

Finally, the theorem is the culmination of Fisher's long-standing aim of synthesizing Mendelism and biometry—the particulate theory's mechanism of inheritance and the principle of natural selection. It was a demonstration of their consistency and cohesion for Fisher in the form of a necessary mathematical relationship between the variation preserved on a Mendelian, or particulate, theory of heredity, and the increase in fitness due to natural selection. Where the 1918 paper was Fisher's attempt to show that the two views may be compatible, the theorem goes one step further. According to Fisher, not only is the Mendelian system compatible with Darwinism, but the particulate nature and additive contribution of genes is essential for selection to operate at all. His theorem was thus a rigorous mathematical solution to a problem that he first addressed as early as 1916.

At the core of the divide between the Mendelians and biometricians was the question of the nature of heredity—how did it operate, how were its effects to be measured, and how, given the complicated nature of heredity, did selection operate? Fisher stepped into this divide and did what he did best—simplify and idealize. Instead of including all the complexity of heredity in his models, he showed that on a number of assumptions, one could partition the contributions to observed variation in traits, and determine exactly that part of the variation that was proportional to the rate of change in a population due to selection.

5. Epilogue

> In a growing research discipline, inquiry is directed not to rearranging old facts and explanations into more elegant formal patterns, but rather to the discovery of new patterns of explanation. (Hanson, 1958, p. 2)

> It is the great charm of these essays that they show the *reasons* which led Darwin to his conclusions, whereas the later works often give the *evidence* upon which the reader is to judge their truth. (Fisher, comparing Darwin's 1842 essays to his later works, 1930, p. 3)

Standard models of scientific progress tend to emphasize the empirical over the formal. Progress in science, on most philosophical accounts, is concerned primarily with the generation and testing of novel hypotheses (Popper, 1963). Fisher was not a good Popperian in the sense that his significant contributions were not experimental. Rather, his genius was in developing mathematical models and following through with the consequences of these models. My view is that this lack of fit of Fisher's contributions with the standard picture of science is a failure of the standard view of science, rather than a failure of Fisher's contributions.

Progress in science is not exclusively empirical. Developing a novel conception of the object of explanation, and developing new patterns of explanation and styles of reasoning, is a key and indeed necessary part of the development of the sciences. Moreover, deriving relationships between general properties of all objects in one's domain, however idealized, is a project that is worthy not only of the physical sciences. Much of what Fisher did was reconceive the object of explanation such that a mathematical theory of evolution was made possible. Fisher (1918, 1922) proposed a new way of picturing populations of organisms, as akin to statistical aggregates of particles in a gas, whose 'motion', or increase and decrease in frequency, was determined by the 'vectors' of selection, mutation, migration, and random extinction by drift. The purpose of the analogy was not to assimilate evolution and statistical mechanics, so much as to emphasize the advantages of a particulate theory of heredity, and set out the most significant conditions on the operation of selection. On the biometrical view, heredity was still viewed in the old model: as a force or tendency of offspring to more or less resemble their ancestors (Gayon, 2000). Fisher's analogy helped lead to the adoption of a Mendelian conception of heredity as a structure, such that the constitution of each individual is directly determined by their parental genotypes. Starting with this novel conception, Fisher, Haldane, and Wright developed models of the genetics of populations.

Moreover, this analogy allowed Fisher to vindicate Darwin's theory of natural selection, not by empirical demonstration, but by a mathematical argument to the effect that evolution was not only possible, but also necessary, given certain assumptions about the character of heredity and the existence of variation. Fisher's fundamental theorem, while its conclusion was more circumscribed than perhaps Fisher thought, still holds true under the idealized conditions he describes. Fisher's derivation requires that we abstract away from the details of actual populations. His theorem assumes no mutation, fixed fitness values, no fertility differences, and no geographic structure to populations. Of course, no population meets these conditions. Nevertheless, Fisher's abstraction enables one to understand the fundamental relationship between additive variance in fitness and rate of increase in fitness.

Fisher compared his fundamental theorem with the second law of thermodynamics. He also commented on the differences between the two:

(1) The systems considered in thermodynamics are permanent; species on the contrary are liable to extinction, although biological improvement must be expected to occur up to the end of their existence.
(2) Fitness, although measured by a uniform method, is qualitatively different for every different organism, whereas entropy, like temperature, is taken to have the same meaning for all physical systems.
(3) Fitness may be increased or decreased by changes in the environment, without reacting quantitatively upon that environment.
(4) Entropy changes are exceptional in the physical world in being irreversible, while irreversible evolutionary changes form no exception among biological phenomena. Finally,
(5) entropy changes lead to progressive disorganization of the physical world, at least from the human standpoint of the utilization of energy, while evolutionary changes are generally recognized as producing progressively higher organization in the organic world. (Fisher, 1930, pp. 39–40)

It is clear from this how many conditions Fisher needed to place on what he considered his fundamental contribution to evolutionary theory. His aim, to bring the kind of generality and rigor he found in statistical thermodynamics to the biological sciences, was thus importantly constrained by the contingency of the biological world. Fisher recognized that the terms in his models described quantitative properties of groups of organisms—such as fitness—which were qualitatively different. And, he saw that his models were idealizations. He noted that evolving populations were transient things: species go extinct, fitness may decrease due to changes in the environment, etc. Nevertheless, there was an important sense in which Fisher thought his fundamental theorem captured the essential character of evolution, vindicating Darwinian selection and demonstrating its compatibility with a particulate scheme of inheritance. His faith in the power of Darwinian selection was, ironically, what enabled him to measure selection's limitations, relative to the genetic variation in any population.

Fisher's own description of his project is illuminating: 'my primary job is to try to give an account of what Natural Selection *must be doing*, even if it had never done anything of much account until now' (Fisher, 1930. Letter to J.S. Huxlen, 6 May 1930, in Bennett, 1983, p. 222). Some might object that the theorem is an artifact of seeing populations of organisms as clouds of points, assemblages of particles whose effects may be partitioned into additive and non-additive components. Of course, organisms are not mere clouds of points, and we know now that most genes have multiple effects and different effects in different genetic environments. Nonetheless, the abstraction was for Fisher a way of answering those skeptical of natural selection. The theorem is an exact statement, even if only of what must be the case for evolution to go forward—a condition of adequacy, if not a universal law.

6. Technical appendix: Proof of the fundamental theorem of natural selection

The following is Fisher's derivation (1930, p. 37), followed by a modern derivation, due to Ewens, 1989).

> Any group of individuals selected as bearers of a particular gene, and consequently the genes themselves, will have rates of increase which may differ from the average. The excess over the average of any such selected group will be represented by a, and similarly the average effect upon *m* of introducing the gene in question will be represented by a. Since *m* measures fitness by the objective fact of representation in future generations, the quantity,

$$\sum{}'(2pa\alpha)$$

will represent the contribution of each factor to the genetic variance in fitness. The total genetic variance in fitness being the sum of these contributions, which is necessarily positive, or, in the limiting case, zero. Moreover, any increase *dp* in the frequency of the chosen gene will be accompanied by an increase $2\alpha dp$ in the average fitness of the species, where α may, of course, be negative. But the definition of a requires that

$$d/dt \log p = a$$
$$\text{or, } dp = (pa)dt$$
$$\text{hence } (2\alpha)dp = (2pa\alpha)dt$$

which must represent the rate of increase of the average fitness due to the change in progress in frequency of this one gene. Summing for all allelomorphic genes, we have

$$dt\sum{}'(2pa\alpha)$$

and taking all factors into consideration, the total increase in fitness is,

$$\sum \alpha dp = dt \sum\sum{}'(2pa\alpha) = W\ dt.$$

If therefore the time element *dt* is positive, the total change of fitness W*dt* is also positive, and indeed the rate of increase in fitness due to all changes in gene ratio is exactly equal to the genetic variance of fitness W which the population exhibits. We may consequently state the fundamental theorem of Natural Selection in the form:

> *The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.*

The following derivation is drawn in full from Ewens, (1989). Ewens there demonstrates that the same theorem holds for both multilocus and overlapping generations, or continuous time models. However, for simplicity's sake, I will give here only the derivation of Fisher's theorem in the single locus, discrete generation case.

There may be two of several alleles at a locus, A1, A2. . .Am. At the time of conception, the frequency of bearers of the AiAi allele will be P_{ii} and of AiAj will be $2P_{ij}$ (since there are two alternative combinations). The frequency pi of the Ai allele will be:

$$p_i = \sum_j P_{ij}$$

Under random mating, $P_{ij} = p_i p_j$. Further, the viability fitness (or fitness to survive to reproductive age) of individuals possessing the genotype AiAj $= w_{ij}$. The population frequency of Ai at generation $t + 1$ is:

$$p_i' = \sum_j P_{ij} w_{ij}/w$$

In other words, the frequency of individuals bearing the i allele will be the summation of all individuals bearing the allele in the prior generation times the average fitness of such individuals over the average fitness of the population as a whole. w, or the mean fitness of the population, is defined as:

$$W = \sum_i \sum_j P_{ij} w_{ij}$$

The average excess a_i of the Ai allele is defined as:

$$a_i = \sum_j P_{ij}(w_{ij} - w)/p_i$$

It follows from the above definitions that the change in frequency of the i allele between parent and daughter generations will be:

$$\Delta p_i = p_i a_i / w$$

Now, it is important to distinguish average excess and average effect. Average effect measures the effect on fitness due to substitution of a single gene; thus, it is a genic, as opposed to genotypic fitness. On Fisher's view, the total mean fitness of a population would be measured by:

$$\sum ijPij(w + \alpha_i + \alpha_j)$$

Or, we may break down the fitness of a population into several components, w, the mean fitness, α_i, the effect on fitness produced by substitution of the i allele, and α_j, the effect on fitness due to substitution of the j allele. Fisher assumes that these values are constant. In other words, in order to calculate the change in mean fitness in a population from one generation to the next, he only takes into account changes in frequencies, and not changes in fitness from one generation to the next. Or, the total fitness of the population in a subsequent generation will be given by:

$$\sum ijPij'(w + \alpha_i + \alpha_j)$$

This is key to his derivation of the Fundamental Theorem. In other words, for the purposes of his derivation, Fisher will assume that the effect on fitness with gene substitution is constant. All other effects on fitness Fisher attributes to what he calls the 'deterioration of the environment', which will include dominance and epistatic effects as well as changes in the environment external to the organism.

We will thus let $w_{ij} = w + \alpha_i + \alpha_j$ for some constants, α_i, ... α_m which satisfy the equation:

$$\sum p_i \alpha_i = 0$$

In other words, the effect on fitness of any individual whose genotype contains an Ai gene of replacing it by Aj is $\alpha_i - \alpha_j$. Further, α_i, . . .α_m are chosen so as to minimize the equation:

$$\sum_i \sum_j P_{ij}(w_{ij} - w - \alpha_i - \alpha_j)^2$$

We then take the sum of squares for this equation to get the additive genetic variance α_A^2. It is, from least-squares theory:

$$2\sum_i p_i a_i \alpha_i = 2w\sum_i (\Delta p_i) a_i$$

Thus, the partial change in mean fitness will be:

$$\begin{aligned}\sum_i \sum_j (P'_{ij} - P_{ij})(w - \alpha_i + \alpha_j) &= \sum_i \sum_j (P'_{ij} - P_{ij})(\alpha_i + \alpha_j) \\ &= 2\sum_i \alpha_i \sum_j (P'_{ij} - P_{ij}) \\ &= 2\sum_i \alpha_i \Delta p_i = \alpha_A^2 / w\end{aligned}$$

References

Bennett, J. H. (Ed.). (1983). *Natural selection, heredity, and eugenics. Including selected correspondence of R.A. Fisher with Leonard Darwin and others*. Oxford: Oxford University Press.

Crow, J. F., & Kimura, M. (1970). *An introduction to population genetics theory*. New York: Harper and Row.

Depew, D. J., & Weber, B. H. (1995). *Darwinism evolving: Systems dynamics and the genealogy of natural selection*. Cambridge, MA: MIT Press.

Edwards, A. W. F. (1994). The fundamental theorem of natural selection. *Biological Reviews of the Cambridge Philosophical Society*, *69*(4), 443–474.

Ewens, W. J. (1989). An interpretation and proof of the fundamental theorem of natural selection. *Theoretical Population Biology*, *36*(2), 167–180.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, *52*, 399–433.

Fisher, R. A. (1922). On the dominance ratio. *Proceedings of the Royal Society of Edinburgh*, *42*, 321–341.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Clarendon Press, New York: Dover.

Fisher, R. A. (1941). Average excess and average effect of a gene substitution. *Annals of Eugenics*, *11*, 53–63.

Fisher, R. A. (1958). *The Genetical Theory of Natural Selection* (Revised Edition). New York: Dover.

Galton, F. (1894). *Natural inheritance* (5th ed.). New York: Macmillan and Company.

Haldane, J. B. S. (1964). A Defense of Beanbag Genetics. *Perspectives on Biology and Medicine*, *19*, 343–359. (Reprinted in Garber, ed. 1985. Genetic Perspectives in Biology and Medicine. Chicago: University of Chicago Press.).

Hanson, N. R. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.

Hull, D. L. (1973). *Darwin and his critics*. Cambridge: Cambridge University Press.

Jenkin, F. (1867). Review of *The origin of species*. *North British Review*, *44*, 277–318 (Reprinted in Hull, 1973).

Kimura, M. (1958). On the change in population fitness by natural selection. *Heredity*, *12*, 145–167.

Moran, P. A. P., & Smith, C. A. P. (1966). *Commentary on R. A. Fisher's paper, 'On the correlation between relatives on the supposition of Mendelian inheritance'*. Cambridge: Cambridge University Press.

Morrison, M. (2002). Modeling populations: Pearson and Fisher on Mendelism and biometry. *British Journal for the Philosophy of Science*, *53*(1), 39–68.

Norton, B. (1975). Metaphysics and population genetics: Karl Pearson and the background to Fisher's multifactorial theory of inheritance. *Annals of Science*, *32*, 537–553, p. 543–4.

Norton, B., & Pearson, E. S. (1976). A note on the background to, and refereeing of, R. A. Fisher's 1918 paper 'On the Correlation Between Relatives on the Supposition of Mendelian Inheritance'. *Notes and Records of the Royal Society of London*, *31*, 151–162.

Norton, B. (1983). Fisher's entrance into evolutionary science: The role of eugenics. In M. Grene (Ed.), *Dimensions of Darwinism* (pp. 19–29). Cambridge: Cambridge University Press.

Pearson, K. (1894). Mathematical contributions to a theory of evolution. *Proceedings of the Royal Society of London*, *185*, 71–110.

Pearson, K. (1897). *The chances of death and other studies in evolution*. London & New York: E. Arnold.

Pearson, K. (1957). *The grammar of science* (3rd ed.). New York: Meridian Books (First published 1911).

Popper, K. (1963). *Conjectures and refutations*. London: Routledge & Kegan Paul.

Price, G. R. (1972). Fisher's fundamental theorem made clear. *Annals of Human Genetics*, *36*, 129–140.

Punnett, R. C. (1976). Report on Fisher's paper, Whittingham Lodge, Cambridge. In: B. Norton, & E. S. Pearson, A note on the background to, and refereeing of, R. A. Fisher's 1918 paper 'On the correlation

between relatives on the supposition of Mendelian inheritance' (pp. 154–155). *Notes and Records of the Royal Society of London*, *31*, 151–162. (First published 1916)

Ruse, M. (1997). *Monad to man: The concept of progress in evolutionary biology*. Cambridge, MA: Harvard University Press.

Sarkar, S. (1998). *Genetics and reductionism*. Cambridge: Cambridge University Press.

Sober, E. (1984). *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge: MIT Press.

Turner, J. R. G. (1983). Fisher's evolutionary faith and the challenge of mimicry. *Oxf. Surv. Evol. Biol.*, *2*, 159–196.

Turner, J. R. G. (1985a). Fisher's fundamental theorem. In R. Dawkins, & M. Ridley (Eds.). *Oxford surveys in evolutionary biology* (Vol. 2). Oxford & New York: Oxford University Press.

Turner, J.R.G. (1985b). Random genetic drift, R.A. Fisher, and the Oxford School of ecological genetics. chap. 12 in The Probabilistic Revolution. vol. 2. eds. Gigerenzer, Krüger & Moran, MIT Press.

Vicedo, M. (1992). "The work of W.E. Castle and E.M. East. A first step toward a reassessment of the relationship between Mendelism and Darwinism in American Genetics, 1900–1915. Paper given at the 1992 History of Science Society Meetings: Washington, D.C.

Weldon, W. F. R. (1893). On certain correlated variations in *Carcinus moenas*. *Proceedings of the Royal Society London*, *54*, 318–329.

Weldon, W.F.R. (1895a). Report of the Committee, consisting of Mr. Galton, (Chairman), Mr. F. Darwin, Professor Macalister, Professor Mendola, Professor Poulton, and Professor weldon, for conducting statistical inquiries into the Measurable characteristics of Plants and Animals. Part 1.

Weldon, W. F. R. (1895b). Remarks on variation and plants. To accompany the first report into the measurable characteristics of plants and animals. *Proceedings of the Royal Society of London*, *57*, 379–382.

Wright, S. (1931). Review of 'The genetical theory of natural selection' by R. A. Fisher. *Journal of Heredity*, *21*, 349–356 (Reprinted in W. B. Provine (Ed.), *Evolution: Selected papers by Sewall Wright*. Chicago: University of Chicago Press, 1986).

Yule, G. U. (1902). Mendel's laws and their probable relations to intra-racial heredity. *The New Phytologist*, *1*, 193–207, 222–238.

Yule, G.U. (1906). On the theory of inheritance of quantitative compound characters on the basis of Mendel's laws—A preliminary note. *Report of the Conference on Genetics*, 140–142.