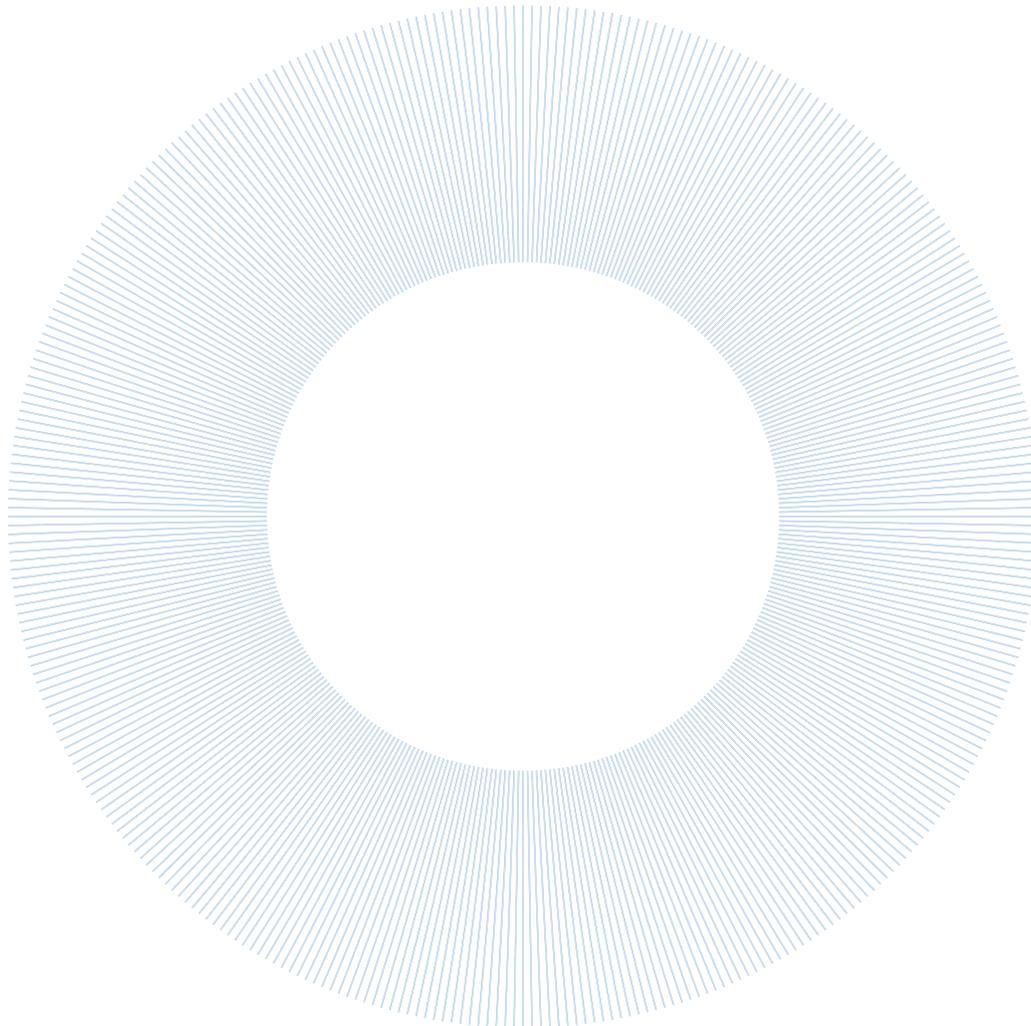


# R. A. Fisher and the Origins of Random Drift



Robert A. Skipper Jr

## *About Insights*

Insights captures the ideas and work-in-progress of the Fellows of the Institute of Advanced Study at Durham University. Up to twenty distinguished and 'fast-track' Fellows reside at the IAS in any academic year. They are world-class scholars who come to Durham to participate in a variety of events around a core inter-disciplinary theme, which changes from year to year. Each theme inspires a new series of Insights, and these are listed in the inside back cover of each issue. These short papers take the form of thought experiments, summaries of research findings, theoretical statements, original reviews, and occasionally more fully worked treatises. Every fellow who visits the IAS is asked to write for this series. The Directors of the IAS – Ash Amin, Michael O'Neill, Susan J. Smith and Colin Bain – also invite submissions from others involved in the themes, events and activities of the IAS.

## *About the Institute of Advanced Study*

The Institute of Advanced Study, launched in October 2006 to commemorate Durham University's 175th Anniversary, is a flagship project reaffirming the value of ideas and the public role of universities. The Institute aims to cultivate new thinking on ideas that might change the world, through unconstrained dialogue between the disciplines as well as interaction between scholars, intellectuals and public figures of world standing from a variety of backgrounds and countries. The Durham IAS is one of only a handful of comparable institutions in the world that incorporates the Sciences, Social Sciences, the Arts and the Humanities.

The focal point of the IAS is a programme of work associated with, but not exclusive to, an annual research theme. At the core of this work lies a prestigious Fellowship programme. This programme gathers together scholars, intellectuals and public figures of world standing or world-promise to address topics of major academic or public interest. Their mission is to anticipate the new and re-interpret the old, communicating across and working between disciplinary boundaries.

Every year, the Institute invites as many as twenty highly creative individuals to spend up to three months in Durham. They are located in Cosin's Hall, a magnificent and spacious 18th century mansion which, together with Durham Cathedral and Durham Castle, forms part of Palace Green, dominating the World Heritage Site of Durham Peninsula. During their stay, Fellows engage with departments and colleges, deliver public lectures and seminars, and, above all, join an international community of researchers to address the theme selected for that year. Further details of the IAS and its Fellowship programme can be found at [www.durham.ac.uk/ias/fellows](http://www.durham.ac.uk/ias/fellows)

## *Copyright*

The design and contents of Insights are subject to copyright. Copyright and Reproduction Rights in all submitted contributions remain with the authors, as described in the Author's Copyright Agreement. Copyright and Reproduction Rights of all other material remain with Insights.

Except under the terms of Fair Dealing (UK Copyright, Designs and Patents Act 1988), the user may not modify, copy, reproduce, retransmit or otherwise distribute the site and its contents (whether text, graphics or original research concepts), without express permission in writing from the Institute. Where the above content is directly or indirectly reproduced in an academic context under the terms of Fair Dealing, this must be acknowledged with the appropriate bibliographical citation.

The opinions stated in the Insights papers are those of their respective authors and do not necessarily reflect the opinions of the Institute of Advanced Study, Durham University, or the staff and students thereof.

## *R. A. FISHER AND THE ORIGINS OF RANDOM DRIFT*

*In his 1922 paper, 'On the Dominance Ratio,' R. A. Fisher introduced the mathematical treatment of (what would ultimately be called) random genetic drift to the emerging field of theoretical population genetics. The received historical narrative of the origins of theoretical population genetics, largely due to William Provine, focuses on Sewall Wright's emphasis on drift's role in evolution and his disagreement with Fisher, who believed drift to be evolutionarily insignificant. This paper examines Fisher's argument for the evolutionary insignificance of drift. The standard view portrays Fisher as coming to the conclusion that drift has little to no evolutionary role because of an assumption about large population size. As I will argue, Fisher's argument is subtler than that. Ultimately, the aim of this paper is to establish what are, arguably, Fisher's landmark contributions to the modeling of drift in evolution.*

### *Introduction*

In finite populations, the variation in the number of offspring between individuals may result in random fluctuations in allele frequencies. This is the phenomenon of 'random genetic drift,' or, more economically, 'drift.' Drift affects the chances of survival of a mutant allele in a population. It also reduces a population's genetic variation.

Sewall Wright, along with R. A. Fisher and J. B. S. Haldane, one of the architects of theoretical population genetics and the evolutionary synthesis more broadly, is credited with the most thorough development of drift as an evolutionary factor. Indeed, the phrase 'random drift' comes from Wright's 1931 landmark paper, 'Evolution in Mendelian Populations.' Notwithstanding Wright's obvious contributions to the development of the concept and mathematical modeling of drift, it was R. A. Fisher who, in 1922, was the first to explore mathematically the evolutionary consequences of drift in a Mendelian population.

Famously, Fisher and Wright became mired in controversy over, among other things, the relative evolutionary significance of drift. Fisher's view was that drift is evolutionarily insignificant; Wright's view was that drift is one of several key evolutionary factors. But Fisher's and Wright's disagreements over drift are not the topic of this paper. Here, I will explore Fisher's early examinations of drift, or as he variously puts it, random survival, steady decay, or the 'Hagedoorn effect,' and his theoretical argument that it is an insignificant evolutionary force. I focus on the period between 1922 and 1930, in which Fisher sets out his synthesis of Darwinism and Mendelism.

I argue, mainly, that Fisher's argument for the evolutionary insignificance of drift is not the argument that has been attributed to him, sometimes subtly, sometimes not so subtly, by historians and philosophers of biology (e.g. Gayon, 1998; Provine, 1971, 1986; Turner, 1987). The standard line is that Fisher thought that because populations are very large, drift is evolutionarily insignificant. I think the standard view gets Fisher's argument backwards. Indeed, my view is that he thought that treating populations as if they were very large was reasonable because the balance of evolutionary factors maintaining genetic variation in a population, including drift, is such that the effects of population size (and structure) are negligible.

Getting Fisher right is important because commentators, including historians, philosophers, and biologists, have thought that his assumption that populations are very large is not sensible and, in fact, is at odds with his own recognition that populations are not actually on the order of 'the total population of the planet,' the population size he uses for theoretical purposes (Fisher correspondence to Wright in 1929; in Provine, 1986).

### 1918-1922

The usual historical discussion of Fisher's views on drift, and more generally on Darwinian evolution in Mendelian populations, starts with his 1922 paper, 'On the Dominance Ratio.' Yet, that paper is not Fisher's first published 'synthetic' consideration of drift. His first such consideration is in his paper, 'Darwinian Evolution of Mutations,' (Fisher, 1922a) published earlier in 1922 than 'On the Dominance Ratio' (Fisher, 1922b) and in fact written and received by early 1921. Interestingly, one does not find Fisher's familiar line that effects of 'random survival' on the genetic variation of a population are evolutionarily insignificant.<sup>1</sup>

In 'Darwinian Evolution of Mutations,' Fisher divides the history of the survival of a new mutant allele into two periods, the first of which concerns the survival of the mutant due to chance; the second concerns the mutant's survival due to the advantages or disadvantages the allele confers upon its carrier in the struggle for existence. In this context, Fisher goes on to say that '[i]f those individuals which contain the new gene are found on the average of the chances of life, and on the average of the genetic natures with which it is combined, to be at a disadvantage in the struggle for existence, then the number of the mutant form will gradually diminish, *with large fluctuations due to chance*' (Fisher, 1922a, pp. 33-4, emphasis added).

I can identify no argument, quantitative or otherwise, to justify the claim. And interestingly, what little mathematics there is in the paper corresponds directly to that of the later paper, 'On the Dominance Ratio.' In 'Darwinian Evolution of Mutations,' Fisher sets up a framework of functional equations that he claims enables him to compute the probability that the new mutant will survive. And he claims that this framework is specifically for use in the first period of the history of the survival of a new mutant, that is, the period concerning the survival of the allele due to chance. This is exactly the same mathematical framework that Fisher uses in 'On the Dominance Ratio.' In the earlier paper he does not carry out the mathematics and in the later paper he does. And he comes to a different conclusion, namely that fluctuations in allele frequency due to chance are negligible in their effect.

In 'On the Dominance Ratio,' Fisher discusses, as he says, 'the distribution of the frequency ratio of the allelomorphs of dimorphic factors, and the conditions under which the variance of the population may be maintained' (1922b, p. 322). He sees this paper as following on the heels of his 1918 paper, 'The Correlation between Relatives on the Supposition of Mendelian Inheritance.' In broad brush strokes, what I take this to mean is that where the 1918 paper defended the principles of Mendelian heredity against the criticisms of the biometricians (and in fact showed the two schemes to be compatible), the (later) 1922 paper continues by defending Darwinism using the principles of Mendelian heredity. Specific to 'On the Dominance Ratio,' Fisher's aim was to respond to a set of criticisms aimed at demonstrating that Darwinian natural selection cannot be the correct explanation of the maintenance of genetic variation in populations because the genetics of populations are such that there is not enough variation available for selection to act on. Indeed, Fisher addresses three key issues concerning factors affecting the genetic variation of populations: (1) the steady state without

mutation and natural selection (drift), (2) equilibrium with mutations but without selection, and (3) the equilibrium distribution for mutations having very small selective effects. We are concerned here with the steady state, or Fisher's response to A. L. and A. C. Hagedoorn's (1921) claims that random survival is an important cause of the reduction of genetic variation in natural populations; what Fisher (1922b, p. 323) calls the 'Hagedoorn effect.'

As I remarked previously, Fisher carries his framework of functional equations from 'Darwinian Evolution of Mutations' to 'On the Dominance Ratio.' But now his mathematics is much richer. In particular, Fisher determines two key quantities for the situation in which a population is under the influence only of the steady decay of heterozygosity, i.e., the Hagedoorn effect. The first quantity describes the time course in generations of the Hagedoorn effect; the second describes the 'half-life' in generations of the effect. Fisher determined the time course to be the  $4N$ , where  $N$  is population size, and the 'half-life' to be  $2.8N$  (Fisher, 1922b, p. 330). This means that the Hagedoorn effect requires, in generations, the product of four and the population size to reduce the genetic variation in the population to the point that all alleles are identical by descent. The 'half-way' point is reached in  $2.8N$  generations.

Fisher used these quantities to 'weight' the significance of the effect of steady decay; the longer the time course, the weaker the effect. Given that the time course of the Hagedoorn effect depends on the population size, the larger the population, the weaker, or less significant, the effect. Simple inspection shows that as population size increases over  $10^4$ , the time course becomes considerable. Indeed, Fisher says, '[a]s few groups contain less than 10,000 individuals between whom interbreeding takes place, the period required for the action of the Hagedoorn effect, in the entire absence of mutation, is immense' (1922b, p. 330).

It is a mistake to think that Fisher's specific argument for the evolutionary insignificance of the Hagedoorn effect is made from a commitment to large populations. Indeed, you cannot get large populations out of the critique, unless you think populations on the order of  $10^4$  are very large. To be sure, Fisher's claim to the rarity of populations with fewer than 10,000 individuals is somewhat tantalizing, but it is not absurd on its face and, moreover, does not commit him to large populations. Thus, if he does commit himself to large populations in 'On the Dominance Ratio,' he must do so elsewhere in the paper. But I do not think that he does.

Recall that Fisher considers himself to be responding to a set of criticisms aimed at demonstrating that Darwinian natural selection cannot be the correct explanation of the maintenance of genetic variation in populations because the genetics of populations are such that there is not enough variation available for selection to act on. We have considered Fisher's rejection of one of those critiques – the Hagedoorn effect. In the surrounding arguments that Fisher sets out, he makes it clear that he is aware that as population size decreases, alleles may become extinct due to chance faster than mutation can replenish them. He also makes it clear that he is aware that as population size increases, mutation adds more variation than random survival can remove. These insights are what is behind Fisher's reflection of Darwin when he (Fisher) says, 'a numerous species, with the same frequency of mutation, will maintain a higher variability than will a less numerous species: in connection with this fact we cannot fail to remember the dictum of Charles Darwin, that "wide-ranging, much diffused and common species vary most"' (Fisher, 1922b, p. 324).

Ultimately, Fisher claims that 'very infrequent mutation will serve to counterbalance the effect of random survival' (Fisher, 1922b, p. 340). But that is in the context of his consideration of mutation and random survival when there is no selection. When there is selection, he thinks

'a much higher level' of mutation is needed. But none of this adds up to a commitment that natural populations must be very large. To be sure, there is reason to think that Fisher's arguments here add up to the view it is well-known that he had, which is that larger populations have more evolutionary potential (because they carry more variability). But again, that does not commit him to the view that natural populations are very large. Overall, there is no evidence in 'On the Dominance Ratio' that should convince us that Fisher thought that random survival is evolutionarily insignificant because he thought that populations are very large. Rather, his conclusions regarding population size are driven by his consideration of the balance of factors maintaining genetic variation.

I want now to leave Fisher's 1922 work on drift, or as he called it, the Hagedoorn effect, steady decay, or random survival. I think there is nothing in his 'On the Dominance Ratio' that commits him to large populations. But maybe that is not a strong claim even if some (e.g. John Turner, 1987) seem to think that he does have that commitment in 1922. As it happens, it is later, in 1929 and 1930, that Fisher makes claims that may more plausibly lead us to the conclusion that he thinks random survival is evolutionarily insignificant because he thinks populations are very large. But before we move on, I want to go back and address the dissonance between Fisher's arguments regarding random survival in 'On the Dominance Ratio' and those in his earlier paper, 'Darwinian Evolution of Mutations.' Unfortunately (for me), I cannot say much of substance at the moment about why in the earlier paper he seems to suggest a significant role for random survival and then in the later paper pushes hard on the claim that in fact the role is insignificant. But I can make the following speculation: In 'On the Dominance Ratio,' Fisher was responding to the 1921 Hagedoorn book, which gets no mention in 'Darwinian Evolution of Mutations.' Perhaps that is because he had finished and sent away the latter in early 1921, prior to seeing the book. If that is correct, then perhaps his reading of the Hagedoorn book spurred him to think more critically about the evolutionary significance of random survival in 'On the Dominance Ratio.' But that is just speculation. So let us move on.

### 1922-1930

It is well-known that in 1929 Wright, in correspondence, corrected Fisher's computation of the two time course quantities I discussed above. Fisher's quantities were twice too high, due to an error in his mathematics. So, the correct time course for steady decay to reduce the genetic variation in a population by half, or the 'half-life,' is  $1.4N$  measured in generations, and the full time course is  $2N$ . With the same correspondence, Wright included a draft of what would be his 1931 paper, 'Evolution in Mendelian Populations,' which articulated his own views of the importance of 'random drift' in evolution. For Wright, the global population is subdivided into 'demes' and semi-isolated so that random drift may create genetic variation for selection, in various modes, to act on.

Fisher's response to Wright's letter makes no mention of Wright's request for clarification about the discrepancy between his own and Fisher's time course quantities. Neither does it mention random drift. But it does mention the population size. Fisher says,

I am not sure that I agree with you [in the 'Evolution' paper] as to the magnitude of the population number  $N$ . To reduce it to the number in a district requires that there shall be no diffusions even over the number of generations considered. For the relevant purpose I believe  $N$  must usually be the total population on the planet, enumerated at sexual maturity, and at the minimum of the annual or other periodic fluctuation.

Here is a clear statement made by Fisher about population size. He thinks the correct assumption to make is that population size is the 'total population on the planet.' And his reason, expressed here at least, concerns what he takes to be the implausibility of the degree of isolation between subpopulations that Wright requires. Indeed, Fisher thought that subpopulations isolated in the way Wright theorized they were would likely drive them to extinction. It was more reasonable to Fisher that there would be sufficient (yet minimal) migration between subpopulations that would, for all practical purposes, unify them into a single population. And if that is correct, then his theoretical assumption that population size is the 'total population on the planet' was, practically speaking, sound: a point that worried Wright considerably at the time.

What is important to keep in mind here concerns, again, the main claim of this paper. Fisher's assumption that populations are, for theoretical purposes, on the order of the 'total population on the planet' did not drive his view that random drift was evolutionarily insignificant. Rather, his population size assumption was motivated by considerations regarding the balance of factors that maintains genetic variation in natural populations, including mutation, drift, migration, and selection.

In Fisher's 1930 paper, 'The Distribution of Gene Ratios for Rare Mutations,' he makes a full accounting of the error in the time course quantities pointed out by Wright. He acknowledges Wright's corrections in a friendly manner, and re-explores with extraordinary precision the same three issues concerning factors affecting the genetic variation of populations that he did in 'On the Dominance Ratio' in 1922: (1) the steady state without mutation and natural selection (drift), (2) equilibrium with mutations but without selection, and (3) the equilibrium distribution for mutations having very small selective effects. Fisher's conclusions in the 1930 paper are no different from those in his (later) 1922 paper. Indeed, regarding the Hagedoorn effect, technically directly affected by his error in 1922, he says, 'Both periods [both sets of time course quantities in his 1922 and 1930 papers] are in most species so enormous that they lead to the same conclusion, namely, that random survival, while of great importance in conditioning the fate of an individual mutant gene, is a totally unimportant factor in the balance of forces by which the actual variability of species is determined' (Fisher, 1930a, p. 205).

Given that the arguments and conclusions of the 1930 paper are the same as those of the 1922 paper, only more precisely mathematically supported, it stands to reason that the same claims regarding population size that I am convinced are implied by the arguments in the 1922 paper are implied by the 1930 paper. In spite of the fact that Fisher's time course quantities were doubly wrong in 1922, he was not forced to change much. So, for instance, when he said in 1922 (p. 330) that '[a]s few groups contain less than 10,000 individuals between whom interbreeding takes place, the period required for the action of the Hagedoorn effect, in the entire absence of mutation, is immense,' he must only change '10,000' to '20,000.' This seems hardly a significant change. Ultimately, the 1930 paper is the 1922 paper redressed in response to Wright's corrections.

When we finally arrive at Fisher's 1930 book, *The Genetical Theory of Natural Selection* (Fisher, 1930b), it is apparent that the treatment of drift is taken directly from the 1930 paper. But more generally, the argument strategy Fisher used in 1918, 1922 and earlier in 1930 of defending the principles of Mendelian heredity and defending Darwinism under the rubric of Mendelian heredity, is carried through. His aim in *The Genetical Theory* is to establish particulate inheritance against the blending theory and then demonstrate how plausibly Darwinian natural selection may maintain genetic variation in Mendelian populations. The

argumentative theme remains the same and, so, my conclusions about what Fisher thought was the relationship between drift and population size are preserved.

### Conclusion

What I have attempted to do in this essay is argue for the claim that Fisher thought that treating populations as if they were very large was reasonable because the balance of evolutionary factors maintaining genetic variation in a population, including drift, is such that the effects of population size (and structure) are negligible. This claim is an alternative to the standard view, a view most clearly expressed by William Provine (1971, 1986), that Fisher thought drift is evolutionarily insignificant because he thought that populations were very large.

My argument has been, essentially, that Fisher argued from the bottom up, from evolutionary factors to the large scale properties of populations, rather than from the top down. Understanding his overall argument strategy that, among other things, the variation upon which Darwinian natural selection can act is not depleted by blending inheritance (as he does masterfully in *The Genetical Theory*), nor by random drift (as he does in 'On the Dominance Ratio'), makes manifest Fisher's bottom up approach. Moreover, it makes sense of a favorite and sometimes puzzling analogy Fisher first deploys in 1918 and again in 1922 and 1930. That is the comparison of his approach to evolutionary theory to the kinetic 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 general laws as to the behavior of gases, leaving but a few fundamental constants to be determined by experiment' (Fisher, 1922b, pp. 321-2). Just as Fisher did not start with gross assumptions about the structure of populations, he did not start with the gas theory analogy. He arrived at each piecemeal.



*Notes*

<sup>1</sup> Thanks to Roberta Millstein for pointing me to this paper.

*Reference List*

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

- - - . (1922a) Darwinian evolution of mutations. *Eugenics Review* 14: 31-4.

- - - . (1922b) On the dominance ratio. *Proceedings of the Royal Society of Edinburgh* 42: 321-41.

- - - . (1930a) The distribution of gene ratios for rare mutations. *Proceedings of the Royal Society of Edinburgh* 50: 205-20.

- - - . (1930b) *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Gayon, J. (1998) *Darwinism's Struggle for Survival: Heredity and the Hypothesis of Natural Selection*. Cambridge: Cambridge University Press.

Hagedoorn, A. L. and Hagedoorn, A. C. (1921) *The Relative Value of the Processes Causing Evolution*. The Netherlands: M Nijoff.

Provine, W. (1971) *The Origins of Theoretical Population Genetics*. Chicago: University of Chicago Press.

- - - . (1986) *Sewall Wright and Evolutionary Biology*. Chicago: University of Chicago Press.

Turner, J. (1987) Random Genetic Drift, R. A. Fisher, and the Oxford School of Ecological Genetics. In Krüger, L., Gigerenzer, G. and Morgan, M. (eds.) *The Probabilistic Revolution*. Vol. 2. Cambridge, MA: MIT Press, pp. 313-54.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* 16: 97-159.

*Backlist of Papers Published in Insights***2008 Volume 1**

<b>No.</b>	<b>Author</b>	<b>Title</b>	<b>Series</b>
1	Boris Wiseman	Lévi-Strauss, Caduveo Body Painting and the Readymade: Thinking Borderlines	General
2	John Hedley Brooke	Can Scientific Discovery be a Religious Experience?	Darwin's Legacy
3	Bryan R. Cullen	Rapid and Ongoing Darwinian Selection of the Human Genome	Darwin's Legacy
4	Penelope Deutscher	Women, Animality, Immunity – and the Slave of the Slave	Darwin's Legacy
5	Martin Harwit	The Growth of Astrophysical Understanding	Modelling
6	Donald MacKenzie	Making Things the Same: Gases, Emission Rights and the Politics of Carbon Markets	Modelling
7	Lorraine Code	Thinking Ecologically about Biology	Darwin's Legacy
8	Eric Winsberg	A Function for Fictions: Expanding the Scope of Science	Modelling
9	Willard Bohn	Visual Poetry in France after Apollinaire	Modelling

*Insights*

Insights is edited by Susan J. Smith, IAS Director and Professor of Geography. Correspondence should be directed to Audrey Bowron ([a.e.bowron@durham.ac.uk](mailto:a.e.bowron@durham.ac.uk)).