errors of random sampling may modify it by a considerable proportion. The estimate of the possible effect of non-heritable factors, which is small, is consequently smaller than the probable error of this determination. The factor representing the proportionate effect of innate causes is actually greater than unity, but exceeds it by an amount less than its standard error; the actual figure is 1.01, with a standard error .08; the value is therefore very unlikely to be less than .93, and we make take .95 as a low estimate. This conclusion is strengthened by the fact that the mean of three determinations of this constant for three different features is .982.

There is, then, in this analysis of variance no indication of any other than innate and heritable factors at work, and a strong probability that whatever non-heritable factors are at work, including errors of measurement, do not contribute more than 5 per cent. and perhaps much less, to the total variance. Any such residue could be determined by a somewhat more accurate numerical knowledge of the relationship correlations, and in particular of the fraternal correlation. Fortunately this is of all the correlation coefficients the one for which it is easiest to gather data, and it is, in my opinion, the one with the simplest and most important meaning.

The conclusions of my investigation may be summed up as follows:—

1. The facts of Biometry do not contradict, but in many ways positively support the theory of cumulative Mendelian factors.
2. If this theory is correct a sufficient knowledge of the correlation coefficients for any one feature, between different pairs of relatives, would enable us to analyse completely and estimate numerically the percentage of variance due to heritable factors.
3. A provisional examination of the existing data shows it to be unlikely that more than 5 per cent. of the variance of the physical measurements of man is due to non-heritable causes. Of other features for which the data is at present insufficient, it would be wisest to judge by comparison of the known facts with those of the physical measurements.

It should not be forgotten that 5 per cent. of the variance would allow considerable scope for the action of environment in individual cases. It amounts to .338 sq. ins., so that the standard deviation if all else were constant would be .58 ins. We may put this another way by saying that any factor producing 5 per cent. of the variance is correlated with stature with a coefficient as high as .224. I am not aware that any environment correlation has ever been established which is more than one-fifth of this amount. So that if there really is a residue of 2 per cent. or 5 per cent. of the variance due to non-heritable cause, there is no evidence of any single environmental cause of sufficient importance to account for the whole of this residue.

In conclusion it is right that I should express my deep sense of gratitude to the Eugenics Education Society, who have most generously assisted me throughout; and in particular to Major Leonard Darwin whose continual kindness and encouragement has enabled me to carry through the work.

Commentary: Fisher 1918: the foundation of the genetics and analysis of complex traits

International Journal of Epidemiology, 2019, 10–12
doi: 10.1093/ije/dyx129
Advance Access Publication Date: 17 July 2017

Peter M Visscher[1,2]* and J Bruce Walsh[3]

[1]Institute for Molecular Bioscience, [2]Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia and [3]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

*Corresponding author. E-mail: peter.visscher@uq.edu.au.

Accepted 20 June 2017

Written while a high school teacher in 1916, Fisher's article (1918)[1] and its easier to understand summary[2] are landmark papers that founded the field of quantitative genetics, or in modern parlance, complex trait genetics. Like many landmark papers, it was not recognised as such at the time. With a century-worth of hindsight, Fisher

1918 did not suffer from a rejection by the Proceedings of the Royal Society, although it may have hurt Fisher's pride. Fisher 1918 is notoriously difficult to read and understand. It also contains an awful lot of new material—new theory, concepts and analysis methods, most if not all of which have stood the test of time.

It is important to place these papers in their historical context. In the 1880s, Francis Galton pioneered the quantification of the resemblance between relatives for traits like height, using linear regression and correlation. Mendel's laws were rediscovered in 1900, and a question that occupied a number of scientists (including Karl Pearson, George Udny Yule, John Brownlee and others) for the next few decades was whether the hereditary and evolutionary properties for a trait like human height were the same as those for Mendel's peas. A particular question was whether inheritance of complex traits was by 'blending' of parental phenotypes, which was seen as different to the inheritance of discrete characters as in Mendel's peas. The incompatibility between blending inheritance and natural selection (and therefore indirectly between blending inheritance and complex traits) was pointed out as early as 1867, in a critique of Darwin's Origin of Species by Fleemin Jenkin.[3]

Pearson and Lee in 1903[4] quantified the correlation between first-degree relatives for height and related measures using a large sample size of 1000s of families, and incorrectly concluded 'Thus for most practical purposes we may assume parental heredity for all species and all characters to be approximately represented by a correlation of .5'. The 'pea versus height' debate was lively and involved some big egos—a thorough historical account of the discoveries in population and quantitative genetics in the early part of the 20th century is given in the book by Provine.[5]

Enter Fisher (1918). In 35 pages packed with derivations and results, Fisher sets out to derive the theory of the resemblance between relatives due to their genetic covariance at 1, 2 and many loci. He defined the term 'variance' and introduced the analysis of variance to partition observed variation into underlying 'causal' factors. Fisher partitioned genetic variation into additive genetic (the variance of what we now call 'breeding values') and the variance of dominance deviations. He also showed the effect of epistatic variance on the resemblance between relatives. Fisher showed how phenotypic variance could be partitioned into meaningful genetic and non-genetic sources of variance, without knowing anything about the underlying genes. Fisher defined quantities that are now called narrow sense and broad sense heritability (variance due to 'essential genotypes' and 'genotypes', respectively). It was a masterful reconciliation between Mendelian inheritance and biometric analysis of quantitative traits.

In Fisher (1919),[2] he gives a mostly non-technical summary of his discoveries. The importance of the analysis of variance is first discussed and its meaning in terms of causes of variability. Fisher then discusses the application of the analysis of variance to quantify the total effect of 'ancestry' (genetic factors) on quantitative trait variation, with numerical examples of human height. Fisher points out that in addition to the parental contribution of variation in offspring (measured as the squared correlation of offspring-parent correlation, i.e. the variance explained in offspring by parental phenotypes), segregation variance ('segregation of hereditary qualities') should also be considered. This solved the problem of 'blending' inheritance mentioned previously. Fisher (1919) concludes that the 'facts' of biometry are consistent with the theory of cumulative Mendelian factors, and that the 'percentage of variance due to heritable factors' (i.e. what we now called heritability) could be estimated numerically from observations on the resemblance between different pairs of relatives. Fisher also makes an important observation that a large heritability would still allow 'considerable scope for the action of environment'.

Interestingly, Fisher (1919) states that his greatest analytical difficulty was the problem of how to allow for the observed phenotypic correlation in trait value between spouses (assortative mating). Indeed, it is easy to miss, but 24 out of the 35 pages in Fisher (1918) are devoted to assortative mating! In comparison, epistasis ('epistacy' by Fisher) only gets 1.5 pages, where two-locus epistasis was modelled. Notably and with great insight, Fisher wrote 'In addition it is very improbable that any statistical effect, of a nature other than that which we are considering, is actually produced by more complex somatic connections'. Indeed, even in the presence of multi-locus higher-order epistasis, Maki-Tanila and Hill (2014)[6] showed that most of this variation maps onto the additive variance, indicating very little contribution from non-additive genetic variances.

Fisher derived two important aspects of the theory of assortative mating, namely the effect of non-random mating on genetic variation in the population, and the effect on the resemblance between relatives. The theory of assortative mating was re-visited in the 1970s and 1980s in a number of important contributions by Lande, Nagylaki, Gimelfarb, Crow, Felsenstein and Bulmer, and empirically modelled mostly by twin researchers. However, with the advent of molecular genetic data it is now possible to revisit these old questions with new data, and test different hypotheses about the causes of trait similarity in spouses[7] and the effect of assortative mating on the genome.

Fisher appeared to adhere to the principle of 'I trust my own theory and somebody else's data', because he seemed somewhat uncritical of the inference he drew about the

contribution of the different sources of variation to human height—Fisher noted the sampling variance of the estimates but did not seem to question the models (assumptions) themselves. He concluded, effectively, that the broad sense heritability was nearly 100%. The conclusions from variance partitioning were that 62%, 21% and 17% of variance were due to additive genetic variance, dominance deviations and assortative mating, respectively. These parameters are not wildly different from estimates obtained subsequently over the past century, apart from the contribution of non-additive (dominance) variation, which seems large by today's standards.[8]

Conflict of interest: None declared.

References

1. Fisher RA. The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans R Soc Edin* 1918;**53:**399–433.
2. Fisher RA. The Causes of Human Variability. *Eugen Rev* 1919;**10:**213–20.
3. Jenkin F. Review of 'The origin of species'. *North British Review* 1867;**46:**277–318.
4. Pearson K, Lee A. On the laws of inheritance in man: I. inheritance of physical characters. *Biometrika* 1903;**2:**357–462.
5. Provine WB. *The Origins of Theoretical Population Genetics*. Chicago, IL: University of Chicago Press, 1971.
6. Maki-Tanila, A, Hill WG. Influence of gene interaction on complex trait variation with multilocus models. *Genetics* 2014;**198**:355–67.
7. Robinson MR, Kleinman A, Graff M *et al*. Genetic evidence of assortative mating in humans. *Nat Hum Behav* 2017;**1:**0016.
8. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* 2008;**4:**e1000008.

International Journal of Epidemiology, 2019, 12–13
doi: 10.1093/ije/dyx184
Advance Access Publication Date: 7 September 2017

Commentary: On R. A. Fisher's paper 'The causes of human variability', 1918

A W F Edwards

Gonville and Caius College, Cambridge, CB2 1TA, UK. E-mail: awfe@cam.ac.uk

Accepted 13 August 2017

R. A. Fisher published two papers in 1918. The first, universally known simply as 'The 1918 paper',[1] was 'The correlation between relatives on the supposition of Mendelian inheritance', the foundation paper of biometrical genetics. The second, 'The causes of human variability',[2] here reprinted, followed shortly after as an outline of the conclusions of its long, difficult, and highly-original predecessor. Published in the *Eugenics Review* it takes the form of a lecture introducing the concept of the variance of a character, in this case human stature, and how it may be partitioned so as to identify and quantify the separate causes influencing it. 'This point of view throws a flood of light upon the meaning to be attached to the results of biometrical research'.

The 1918 paper had been a long time in gestation. Undertaken at the suggestion of Major Leonard Darwin, Charles's fourth surviving son, it had been submitted for publication to the Royal Society of London but withdrawn after unfavourable reviews by Karl Pearson and R. C. Punnett. It was finally published in the *Transactions of the Royal Society of Edinburgh*, with Darwin's support. The attendant delay in publication made it all the more necessary to draw it to the attention of a wider audience. The *Eugenics Review*, to which Fisher was a frequent contributor, was an important outlet for papers on human genetics in the early days of the subject, reaching a wide audience.

In 'The correlation between relatives', Fisher had written:

> The contributions of imperfectly additive genetic factors divide themselves for statistical purposes into two parts: an additive part which reflects the genetic nature without distortion, and gives rise to the correlations which one obtains; and a residue which acts in much the same way as an arbitrary error introduced into the measurements.

