Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENCE, TECHNOLOGY, AND HUMANITIES

Ciclo XXV

Settore Concorsuale di afferenza: 11/C2 – Logica, Storia e Filosofia della scienza
Settore Scientifico disciplinare: M-STO/05 – Storia della scienza e delle tecniche

"MAKING SENSE OF FIGURES": STATISTICS, COMPUTING AND INFORMATION TECHNOLOGIES IN AGRICULTURE AND BIOLOGY IN BRITAIN, 1920s-1960s

Presentata da: Giuditta Parolini

Coordinatore Dottorato
Giuliano Pancaldi

Relatori
Staffan Müller-Wille

Giuliano Pancaldi

Esame finale anno 2013

ACKNOWLEDGEMENTS

My PhD research has been sponsored by a three-year scholarship in Science, Technology, and Humanities awarded to me by the University of Bologna in 2010. In Bologna my tutors have been Anna Guagnini and Giuliano Pancaldi, who has also co-supervised my thesis. I wish to thank them both for their support during my doctoral research.

I have a debt of gratitude to Staffan Müller-Wille, who has patiently read my writings – abstracts, presentations, research proposals, chapters of the thesis – and has greatly contributed to improve them with his advice. I want to thank him not only for his academic supervision and for putting me in contact with people who could help me further, but also for his human kindness and constant encouragement.

During my PhD I have spent several months abroad as a visiting researcher at the Centre for Medical History at Exeter University and at the Max Planck Institute for the History of Science in Berlin. The opportunity to share my ideas with people working in different institutions has been a precious chance to enrich my work and I would like to thank in particular Veronika Lipphardt and Susan Bauer that I met in Berlin for the interest they have taken in my work and their suggestions.

A consistent share of my PhD research has been represented by archival work and I have visited several research libraries and institutions and worked on many collections in order to gather the materials on which my thesis is based. This is the part of my research that I have most enjoyed and I remember with affection the hours that I have spent reading in the archives. I am therefore grateful to the people that under different capacities have made available for me documents, often many at a time of them, and have given me advice on further materials that might be of my interest. In particular, I wish to thank the staff at the Bodleian Libraries, University of Oxford; the John Innes Centre Archives, Norwich; the David Lubin Memorial Library in Rome; the National Library of Scotland, Edinburgh; the Royal Society Centre for the History of Science, London; the Royal Statistical Society Archive, London; Rothamsted Research Library and Archive, Harpenden; University College Special Collections, London; the Wellcome Library, London. I am also indebted to the staff of the Science Museum London for granting me access to archive materials and technical information and for the opportunity to view the collection of computing tools not already on display in the museum, and the mainframe Elliott 401, in possession of the institution since 1965, but never exhibited.

There are several archives that I have been unable to visit in person and from which I have received copies of documents. I am especially grateful to the staff at the Barr Smith Library, University of Adelaide, for providing several scans of the Fisher Papers and to the library of the department of philosophy at the University of Bologna, which acquired the documents on my behalf. Another consistent set of documents was reproduced from the Rockefeller Archives in New York and was related to the Galton Laboratory Serum Unit sponsored by the foundation. The collection of the National Research Development Corporation held at the National Archives for the History of Computing at the University of Manchester has been precious for exploring the computerization of statistics at Rothamsted Experimental Station. I am also grateful to the Special Collections Department, Iowa State University Library, the University of East Anglia and the Museum of English Rural Life, University of Reading for the reproduction of documents.

In archives, libraries, and foreign institutions I met several people that gave me freely their advice. Among them I wish to thank John Aldrich, Paul Brassley, Eileen Magnello, Simon Lavington and Anthony Edwards, who has shown me also several memorabilia of Ronald Fisher. Heartfelt thanks also to Roger Plumb, John Jenkins and the staff of the visual communication unit at Rothamsted Research for their help in going through archive materials and iconographic repositories.

When I did my archival research in the summer of 2010 and 2011, the records related to the Rothamsted statistics department were still divided between the main archive at Rothamsted Research and what was then the biomathematics and bioinformatics department, now the department of computational and systems biology. I wish to thank the head of the department Chris Rawlings for allowing me to spend time there as a visiting worker and to Gavin Ross for helping me to go through the collection. I am especially grateful to Gavin for the effort he has spent in conserving the records of the department in which he has worked for decades and for the kindness of him and his wife Rosemary during my visits to Harpenden.

My research travels in the UK have been an opportunity for keeping in touch with the people I met in London while attending my MSc in the History of Science, Medicine and Technology in 2008/2009. In particular I wish to thank Philip Stevens for his friendship, constant support, and the enjoyable time we spent in London, Exeter and Berlin. Among the people who supported me in Bologna with their friendship, Sara and Marco deserve a special mention.

To all the people who have helped me in one way or the other during my PhD and that it has not been possible to name here explicitly, I want to express my gratitude for their words of encouragement, advice and constructive criticism, and the time that they have dedicated to me in order to improve my work. My PhD research has been great part of my life for the past three years and I will never forget their contribution to it.

Quotations and images

For the permission to publish citations from archive materials I wish to thank the Barr Smith Library, University of Adelaide; the Biotechnology and Biological Sciences Research Council; the Bodleian Libraries, University of Oxford; the Medical Research Council (UK); the Museum of English Rural Life, University of Reading; The National Archives of the UK; the National Library of Scotland; Pearson Education Limited; the Rockefeller Archive Center; Rothamsted Research Library and Archive; the Royal Society Centre for the History of Science; the Science Museum London; the Special Collections Department, Iowa State University Library; University College London Special Collections; the University of Manchester Library; the Wellcome Library London. The images have been reproduced with permission of the British Red Cross Museum and Archives; Cambridge University Press; the Medical Research Council (UK); Rothamsted Research Ltd.; Wellcome Library, London.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Agricultural Improvement Council
ARC	Agricultural Research Council
BBSRC	Biotechnology and Biological Sciences Research Council
BL	Bodleian Libraries
BRC	British Red Cross
BSL	Barr Smith Library
IIA	Institut International D'Agriculture (International Institute of Agriculture)
MERL	Museum of English Rural Life
MRC	Medical Research Council
TNA	The National Archives of the UK
NAAS	National Agricultural Advisory Service
NAHC	UK National Archive for the History of Computing
NLS	National Library of Scotland
NRDC	National Research Development Corporation
RA	Rockefeller Archive Center
RES	Rothamsted Experimental Station
RR	Rothamsted Research
RSCHS	Royal Society Centre for the History of Science
RSS	Royal Statistical Society
SM	Science Museum
UCL	University College London
UEA	University of East Anglia
UML	The University of Manchester Library
WL	Wellcome Library

INTRODUCTION

I. *Preamble: Teatime at Rothamsted*

Fig. *i* Teatime at Rothamsted Experimental Station (around 1931).
Credits: *Records of the Rothamsted Staff Harpenden* (1931), p. 45. Copyright Rothamsted Research Ltd.

The snapshot here reproduced was first published in the staff journal of Rothamsted Experimental Station in 1931 to celebrate the cheerful atmosphere of afternoon tea in the agricultural institution. The scientific staff, men and women together, mingle on the lawn warmed by the sunshine, among the tables set out for tea. The picture, presumably taken by the official photographer of the institution, portrays them deep in conversation or pleasantly smiling, smoking pipes and drinking tea. The original caption reminded the reader that the image depicted "the staff at 4:15pm any summer afternoon".[1]

Rothamsted Experimental Station, now Rothamsted Research, has been a key institution of British agricultural science throughout its history. Its beginnings date back to the mid-nineteenth century, when John Bennet Lawes, a businessman engaged in the fertilizer industry, sponsored a

[1] RES (1931c), p. 45.

series of long-term experiments on crops and fertilizers in the fields of his private estate, Rothamsted, located in the village of Harpenden in Hertfordshire.[2]

Afternoon tea was a twentieth century addition to the routine of the station. It was instituted in 1906 as a courtesy to the botanist Winifred Brenchley, the first female member that entered in the Rothamsted scientific staff, and rapidly established itself as a daily social event for all the station research workers.[3] During one such teatime gathering in the 1920s the statistician and geneticist Ronald Aylmer Fisher, a founding father of modern statistics and a Rothamsted staff member since 1919, took inspiration for the statistical planning of experiments from the case of a lady tasting tea. Here the description of the tasting experiment in Fisher's own words.

> A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup. We will consider the problem of designing an experiment by means of which this assertion can be tested. [...] Our experiment consists in mixing eight cups of tea, four in one way and four in the other, and presenting them to the subject for judgement in a random order. The subject has been told in advance of what the test will consist [...]. Her task is to divide the 8 cups into two sets of 4, agreeing, if possible, with the treatments received.[4]

According to the statistician's biographer, his daughter Joan, the tasting experiment was performed by Ronald Fisher and two of his colleagues at Rothamsted, the algologist Muriel Bristol and the chemist William Roach. Bristol was the lady "who preferred a cup [of tea] into which the milk had been poured first", while Roach assisted Fisher, sceptical on Bristol's refined sense of taste, in setting up the experiment on the spot.[5] Bristol succeeded in her test and Fisher, who had gained food for thoughts as statistician from the occasional tasting experience, honoured her a decade later in *The Design of Experiments*, the book from which the previous citation has been extracted. Since then the lady tasting tea has achieved enduring fame in the statistical planning of experiments.[6]

"Fisher found the staff tea a particularly agreeable institution. He was a notable figure, with his shabby clothes and shaggy head", claims Fisher's biographer.[7] There is no reason to doubt such account, at least on the evidence offered by a series of photographs taken during teatime at the station. Fisher is portrayed quietly sit on the grass, talking with his colleagues or smoking his pipe.

---

[2] On the history of Rothamsted Experimental Station and its role in British agricultural science see E. J. Russell (1966).

[3] For the origins of the afternoon tea at Rothamsted see E. J. Russell (1966), p. 235.

[4] R. A. Fisher (1947), p. 11.

[5] J. Fisher Box (1978), p. 134. An alternative version of the story (D. Salsburg (2001), p. 1) sets the teatime experiment in Cambridge, but it has been criticized as inaccurate (J. Ludbrook (2005)).

[6] See for instance N. T. Gridgeman (1959) for a comparison of the several interpretations of the tea tasting experiment, and G. E. P. Box (1976).

[7] J. Fisher Box (1978), p. 132.

Fig. ***ii*** Teatime at Rothamsted Experimental Station (1920s) [1]. R. A. Fisher's head is on the left, half hidden.
Credits: Copyright Rothamsted Research Ltd. (RR Library and Archive, Ref. PHO 2.2.6)

Fig. ***iii*** Teatime at Rothamsted Experimental Station (1920s) [2]. R. A. Fisher (on the left) is portrayed with three other research workers at Rothamsted. From left to right: R. A. Fisher, J. Henderson Smith, assistant mycologist, A. D. Imms, head of the entomology department, D. Ward Cutler, head of the microbiology department.
Credits: Copyright Rothamsted Research Ltd. (RR Library and Archive, Ref. PHO 2.2.2)

Fig. *iv* Teatime at Rothamsted Experimental Station (1920s) [3]. R. A. Fisher, well identified by the thick beard, is clearly visible smoking his pipe in the top right of the photograph.
Credits: Copyright Rothamsted Research Ltd. (Visual Communication Unit Archive, uncatalogued)

Teatime at Rothamsted, however, was more than a social event. It represented an opportunity for an informal exchange of scientific ideas and the chemist Thomas Eden – colleague, co-worker and later even student of Ronald Fisher – several years after his resignation from the station still remembered "the grace and philosophical discussion of lab tea in Harpenden".[8] Fisher enjoyed scientific conversation at teatime and ventured also in the controversial topic of eugenics, in which he was engaged since his years as an undergraduate student in Cambridge.[9] But mostly he took advantage of the teatime conviviality to become acquainted with the problems encountered by the Rothamsted research workers and by the several scientific visitors of the station and he explained them how statistics could be helpful in the planning of field and laboratory trials and in the analysis of the results.[10]

In the turn of few years Fisher became a statistical consultant of the institution researchers and the teatime encounters certainly contributed to draw him closer to the experimentalists he met with. Analysis of variance and experimental design, the statistical methods Fisher developed during the 1920s, found immediate application in agriculture and biology at Rothamsted Experimental Station and in the following decades they spread in several other institutions and were applied in a wider range of experimental sciences.

The teatime preamble, I want to suggest, highlights the social dimension that was instrumental for introducing these statistical methods in the scientific life of the institution. Only the constant dialogue between mathematically minded statisticians – Fisher's appointment coincided with the institutionalisation of statistics and the opening of a dedicated department at the agricultural station – and experimentalists made possible the concurrence of mathematical tools and experimental needs that I am going to examine in my thesis.

I will investigate the development of analysis of variance and experimental design and their application to agriculture and biology in the period 1920s-1960s both at Rothamsted Experimental Station and at the Galton Laboratory. The latter institution was the main centre for research in eugenics and human genetics in Britain and Ronald Fisher moved there from Rothamsted in 1933. I am going to consider Fisher's mathematical tools an integral component of the history of the experimental sciences and I will discuss analysis of variance and

[8] Letter from T. Eden to E. J. Russell, 18th September 1941, E. J. Russell Correspondence, MERL, University of Reading, FR HERT 11/1/1.
[9] J. Fisher Box (1978), p. 132.
[10] G. E. P. Box (1976), p. 793.

experimental design from the perspective of the computing instruments and of the information technologies that promoted the application of these statistical methods to field and laboratory data.

The emphasis on the intertwining of statistical methods and experimental research and the prominence given to computing and information technologies are the novelties of my research and to point them out I will start with a review of the current literature on statistics and computing in Britain in the half century covered by my thesis. In this review I will consider as secondary literature only the historical accounts written with a critical aim, while I am going to count obituaries, institutional reports and eulogies among the primary sources. The review so drafted will show clearly the peculiarity of my approach and the aims of the thesis that I am going to discuss afterwards. An outline of the thesis chapters and of the primary sources I have used in my research will then complete the introduction.

II. *Review of existing literature*

A comprehensive history of British statistics during the half century covered by my thesis is not available. Internalist accounts – such as Anders Hald's *History of Mathematical Statistics* – and externalist ones – as *Statistics in Britain* by Donald Mackenzie – end in 1930 or 1935 with a discussion of Fisher's statistical methods.[11] Even Stephen Stigler's essay, 'The history of statistics in 1933', is very much concentrated on the 1930s and mainly concerned with the development of statistics in the United States.[12]

Regrettably, thus, there is no publication that covers step-by-step both the evolution of Fisher's ideas and the theory of hypothesis testing developed in the period 1926-1938 by the statisticians Jerzy Neyman and Egon Pearson. The Neyman-Pearson theory is the other main component of twentieth century applied statistics and found widespread use in the experimental sciences. It represents thus a natural comparison for Fisher's own methods, although it is not of immediate concern for my own research.[13]

The multiauthored volume *The Empire of Chance*, in which the use of probability and statistics in the sciences is investigated according to a long-term perspective, is partly an exception to the

[11] A. Hald (1998); D. A. Mackenzie (1981). Hald covers the development of Ronald Fisher's theory of estimation up until 1935, while Mackenzie's ends in 1930.
[12] S. M. Stigler (2002), pp. 157-172.
[13] For an account of the theory developed by Neyman and Pearson see E. L. Lehmann (1993, 2011) and G. Gigerenzer et al. (1989), pp. 98-106.

limited time-frame of the publications already mentioned.[14] But the unsystematic structure of the book – which deals with different case studies to explain how statistics transformed our ideas about nature, mind and society – does not offer to me a general framework suitable for my story. Moreover, like all the publications mentioned so far, its interest is limited to the development of statistical ideas and to the role acquired by statisticians as expert consultants in different fields, while computing instruments and information technologies, which are at the core of my narrative, are never substantially taken into account.

Therefore, for the development of statistics as a discipline after the 1930s and for the computing tools and information technologies popular among statisticians in the period 1930s-1960s, I have mainly referred to internalist accounts written by statisticians, commemorative papers with an historical gist, published debates or committee reports of associations, like the Royal Statistical Society or the American Statistical Association, and statisticians' accounts of the development of their own discipline.

If the historians' interest for statistics after the 1930s is scarce, more literature is instead available on the history of the discipline until that stage and in this literature also the computing equipment employed by statistical laboratories has been investigated, although never as a substantial element in the development of statistical theory. For this time period I will mainly refer throughout my thesis to Theodore Porter's work. Porter has extensively written on the development of probability and statistics during the nineteenth century and in the first decades of the twentieth century, and has also dedicated a biography to Karl Pearson, a key figure in British statistics at the beginning of the twentieth century and an essential reference point for understanding, by way of compare and contrast, the methods developed by Ronald Fisher.[15] Among Porter's contributions to statistics I will refer also to his *Trust in Numbers*, an investigation in the use of quantification in experimental research as well as in administration and actuarial statistics, and a reflection on how quantification has been entrenched with objectivity.[16] I will quote it at several stages in relation to the conventional thresholds accepted in statistics.

Statistics in Britain by Donald Mackenzie covers extensively the same period examined by Porter, as it traces the development of the discipline from 1865 to 1930. The book is not an account of applied statistics *per se*, but of the intertwining between statistics and eugenics, the theory of the betterment of mankind through the selection of favourable genetic features. Mackenzie examines three main British contributors to both statistics and eugenics – Francis Galton, the founding father of eugenics and the promoter of its statistical approach; Karl

[14] G. Gigerenzer et al. (1989).
[15] T. M. Porter (1986); T. M. Porter (2006).
[16] T. M. Porter (1996).

Pearson, Galton's scientific heir and the head of the first statistics department in Britain; the already mentioned Ronald Fisher –, but he leaves aside British statisticians who do not fit the eugenic pattern, such as George Udny Yule, a student of Karl Pearson and the first promoter of statistics at Cambridge University and at the Cambridge School of Agriculture.

Mackenzie's argument on the eugenic roots of British statistics has been influential during the 1980s. Both Porter and the authors of the above mentioned *Empire of Chance* have accepted Mackenzie's position on statistics and eugenics.[17] Since the 1990s, however, the strong association suggested between the two fields has been challenged. One example is Eileen Magnello's research on Karl Pearson, in which she claims the independence of the work done by Karl Pearson in statistics and eugenics, considering the activity of the laboratories – the Biometric Laboratory and the Galton Laboratory – Pearson managed at University College London.[18] In my thesis I do not disregard that eugenics mattered in terms of funding, training and opportunities for the development of statistical methods, but there will be comparatively little space for eugenics in my own account, because I am going to take, as I explain in section III of this introduction, a complementary approach to the one adopted by Mackenzie looking at other contexts, such as agricultural science, in which statistics developed.

Besides the critical literature already mentioned, other useful sources on the history of statistics are Steven Stigler's several publications on the topic. In his *History of Statistics*, Stigler focuses on the assessment of uncertainty and variability before 1900.[19] In particular, in the third section of the book Stigler investigates the use of statistics in the study of heredity and the English tradition established by Galton, Pearson and Yule. *Statistics on the Table* is instead a collection of essays, grouped under thematic areas and concerned with disparate aspects of probability and statistics, including the works of Ronald Fisher and Karl Pearson, and I have referred to it in relation to the controversy between Fisher and Pearson over the degrees of freedom in the chi-square distribution.[20]

An extensive literature is available on Ronald Fisher, who, beyond his involvement in statistics, was also a "fairly well-known" geneticist.[21] Fisher has been depicted on a case-by-case basis as a genius of mathematical statistics, a founding father of population genetics, an active

---

17 "More impressively, perhaps, statistical innovation remained closely tied to particular applications throughout what might be called its heroic period, from Pearson to Fisher. [...] Not only Pearson, but Fisher, too, moved to statistics from physics as a result of an infatuation with eugenics." (T. M. Porter (1986), p. 316); "As Donald Mackenzie has shown, the founders of modern statistics were deeply committed to eugenic control of human evolution." (G. Gigerenzer et al. (1989), p. 53)
18 M. E. Magnello (1993); M. E. Magnello (1999a, 1999b).
19 S. M. Stigler (1986).
20 S. M. Stigler (2002).
21 A. W. F. Edwards (2003), p. 311.

eugenicist, a contributor to the theory of statistical inference, while much less space has been devoted to his role as consultant of research workers. I present a detailed discussion of the literature available on Ronald Fisher in chapter one, in which such review helps to understand the novelty of my historiographical approach. It would be redundant to repeat it here and I refer the reader to the chapter introduction.

Little, instead, is available on Frank Yates, the other statistician that I will examine in my thesis. Biographical information is available only in the obituary notices written at the time of his death in the 1990s by his fellow statisticians.[22] As understandable, these obituaries are mainly collections of scientific and bibliographical records and of some personal memories, but they do not attempt any systematic evaluation of Yates' work within the broader scenario of statistics.

Moving from the people to the institutions in which they worked, again there is no secondary literature available on the Rothamsted statistics department, despite its long tradition and its relevance. Possibly the fragmentary nature of the archival collections and the difficulty of accessing the station's historical repositories have acted as deterrent. The agricultural institution, in fact, has not preserved systematically the materials related to his scientific activity and in the archive of Rothamsted Research can be found only the documents voluntarily donated by former researchers. Moreover the statistics department of the agricultural institution – and its successors, at first the biomathematics and bioinformatics department and now the department of computational and systems biology – have kept for decades an independent archive, but no comprehensive catalogue can be obtained for consultation. The only systematic accounts for the period 1920s-1960s nowadays available on the Rothamsted statistics department can be found, alongside the list of the personnel and a bibliography of the scientific publications, in the official reports of the agricultural institution.[23]

Instead, the Galton Laboratory as a centre for research on statistics and eugenics has received much more interest among historians. A comprehensive account of the laboratory originally founded by Francis Galton for the study of eugenics has been written by Daniel Kevles and covers the period from the 1900s to the 1960s.[24] The scientific activity of the Galton Laboratory under Pearson features also in the accounts of Theodore Porter, Eileen Magnello and David Grier.[25]

---

[22] See for instance D. J. Finney (1995), M. J. R. Healy (1995), G. Dyke (1995).

[23] Up to 1928 the institution published a report every two year, while afterwards the reports were prepared every year. Luckily there are not many gaps in the series. Only during WWII due to paper shortage the publication was suspended and a comprehensive report covering the war years was printed in 1946.

[24] D. J. Kevles (1995).

[25] T. M. Porter (2006); M. E. Magnello (1999a, 1999b); D. A. Grier (2007).

Kevles, Porter and Magnello have shown only a passing interest for computing in Pearson's laboratory, while the last contribution mentioned, Grier's, is exclusively concerned with this aspect. In particular, Grier's has been a crucial reference for the too often forgotten history of human computers in scientific computation and for the role that they had at the Galton Laboratory.[26] Unfortunately, there has not been a systematic attempt to follow the computing activity at the Galton Laboratory also in the decade in which Ronald Fisher ran the centre. As evident, this omission has been a relevant problem for my own research and I have filled the gap as much as possible with primary sources.

Turning now to the technological side of my research, I want to draw a general picture of the secondary literature available on computing instruments and information technologies in Britain during the period examined in my thesis. Among the historical accounts that I have used at several points there is Mary Croarken's *Early Scientific Computing in Britain*.[27] Croarken's book retraces the tools and organization of scientific computing in Britain in the first half of the twentieth century. I have referred to it in relation to the computing equipment available in the 1920s and 1930s, the adoption of punched-card equipment in scientific calculation and the development of digital computers in Britain. Besides Croarken's publication, Martin Campbell-Kelly's history of the computer company ICL has provided a framework for the development of the punched-card and computer industry in Britain and the adoption of these tools in scientific computation and office routine.[28]

Croarken and Campbell-Kelly have also worked jointly on the history of the British Association Mathematical Tables Committee, the main body engaged in table making in Britain during the nineteenth and twentieth centuries.[29] The efforts of the British Association in the preparation of mathematical tables are a necessary point of reference for understanding the general context in which the *Statistical Tables for Biological, Agricultural and Medical Research*, the collection co-authored by Ronald Fisher and Frank Yates and examined in my thesis, was published. Campbell-Kelly and Croarken are also among the editors of an extensive anthology of essays on mathematical tables.[30] The anthology sketches the history of these computing tools during the millennia and throughout different contexts of use (from tabulation in the ancient civilizations of the Middle East to spreadsheets), and it has been helpful as well in my discussion on the *Statistical Tables*.

---

[26] D. A. Grier (2007), pp. 108-113; pp. 130-133.
[27] M. Croarken (1990).
[28] M. Campbell-Kelly (1989).
[29] M. Croarken and M. Campbell-Kelly (2000).
[30] M. Campbell-Kelly et al. (2003).

In my analysis of computerization in the Rothamsted statistics department after WWII the author that I have more often quoted is Jon Agar. His comparative study of computerization in scientific research during the 1950s has provided food for thoughts in the acquisition of the Elliott 401, the first mainframe adopted by the Rothamsted statistics department.[31] More generally Agar's approach to the history of the digital computer – an approach which values a long-term perspective and the detailed examination of the human actors, in particular expert groups, in the mechanization of computing – has been influential in my analysis of computerization in statistics, as I regard a bias of large part of the history of computing the choice to focus only on the technology itself.[32]

Along with Agar's work, a relevant source on the Elliott 401 has been represented by Simon Lavington's book on Elliott Brothers, the company that developed the mainframe.[33] Lavington's book is rich of technical information on the company, its management and the technologies it realized, but the approach of business history that gives the general structure of the book is very far from my aims and no trace of it will be found in my thesis.

From the present review it is evident that the historical accounts up to now available on British statistics have privileged the theoretical outcomes over the systematic investigation of the relationships between statistical methods and experimental practices, nor they have been very interested in the material tools of statistics. On the other hand, the literature on British computing and information technologies have never thoroughly addressed statisticians engaged in experimental research as a users' group.

My research, thus, contributes to the current literature considering statistical methods a component of the history of experimental research and approaching them bottom-up, looking at the computing instruments and information technologies that contributed to the application and the dissemination of these mathematical tools. This approach to the history of statistics represents as well a contribution to the history of computing and information technologies. Statisticians, in fact, were keen users of these tools and thus my research is also a case study on the adoption of computing equipment and information technologies in science. On computing in agricultural science, for instance, very little has been written so far.[34] My research, thus, contributes also to the growing literature interested in going beyond the technology for technology's sake discourse that has been predominant until recently in the history of computing

[31] J. Agar (2006).
[32] Besides the paper mentioned above J. Agar (2006), a suggestive example of Agar's historiographic approach is J. Agar (2003), the book in which he describes the general purpose computer as an outcome of the British civil service.
[33] S. Lavington (2011).
[34] An exception is the work of D. A. Grier on the U.S. department of agriculture and the Iowa State College Computing Service more extensively mentioned in section IV of this introduction.

and has given us much information on engineering developments and business practices, but comparatively little on the users of these technologies.[35]

III. ***Statistics in agriculture and biology in Britain, 1920s-1960s***

Analysis of variance and experimental design, the statistical methods I am going to consider in my thesis, are two relevant examples of the mathematical tools developed for experimental research in the first decades of the twentieth century. By looking at the adoption of these methods, I want to account for the introduction of statistics in agriculture and biology – in Britain and worldwide – and discuss the new role gained by statisticians as consultants of field and laboratory scientists. Analysis of variance and experimental design, in fact, had as their primary user the 'research worker', a definition that broadly encompassed all the people engaged in experimental research, whether in academia or in industry, with a primary concern for soft sciences such as agriculture and biology, but with a general applicability in medicine, psychology, quality control, engineering and several other disciplines.

In so doing I am taking up the question left unanswered by Donald Mackenzie in the conclusion to his *Statistics in Britain*. "By the mid-1920s", Mackenzie's argues,

> there were [...] clear signs of the beginning of a new era in the development of statistical theory in Britain. The new role for the statistician in agricultural and industrial production, and in scientific research in general [... was] one of considerable importance. Its evolution, and the way in which the practical demands associated with it translated themselves into goals of statistical theory, are interesting problems.[36]

The diffusion of statistics in scientific research remains excluded from *Statistics in Britain* as the focus of the book is on the eugenic roots of British statistics, but there were other contexts in which statistics developed. As I will argue in my thesis, for instance, the adoption of statistics in agricultural science during the 1920s was the result of a long tradition connected to the making of field experiments on the growth of crops and was not immediately relevant to the eugenic tradition accounted by Mackenzie.

The aim of my doctoral thesis, therefore, is to begin where Mackenzie's book ends, that is investigating how statistics became part of experimental research during the half century from the early 1920s to the end of the 1960s. As pointed out by Mackenzie, the 1920s were the decade in which the new role of statistics in science, industry and agriculture revealed itself in Britain. Mathematicians like Yule and Fisher actively contributed to the application of statistics in British

[35] For an overview of the general trends in the historiography of computing see T. J. Misa (2007).
[36] D. A. Mackenzie (1981), p. 213.

agricultural institutions; the Guinness Brewery in Dublin officially opened its own statistics department under the chemist William Sealy Gosset, whose work was influential for Ronald Fisher; in 1925 appeared *Statistical Methods for Research Workers*, the book in which Fisher explained analysis of variance and experimental design to a wide public of experimentalists.[37] A few years later, in 1933, the Royal Statistical Society created an industrial and agricultural research section under the stimulus provided by the statistician Egon S. Pearson, the son of Karl Pearson.[38]

"Making sense of figures" – the quotation taken from Ronald Fisher's and used in the title of my thesis – has to be understood in this perspective. My interest, in fact, lies in the value that statistical methods acquired as tools of scientific research, side by side with the laboratory equipment and the field practices adopted by research workers. Statistical methods made quantification processes meaningful in disciplines, like agriculture and biology, which had traditionally relied on activities such as collecting and describing diversity rather than timing variation.

Fisher's paper from which the title quotation is taken is a manifesto of inferential statistics, but as evident from my approach, analysis of variance and experimental design needed more than "a logical process of the kind we call inductive" to establish themselves.[39] Statistical inference became a tool of experimental research, not only for the intrinsic prestige of mathematics and the claimed objectivity of numbers, but because a new alliance was borne between Ronald Fisher and the research workers who consulted him. Thus, I am going to seek an answer to the necessity of "making sense of figures" in the border area where different disciplinary traditions met and made possible the mathematization of agriculture and biology.

My account will begin with the development of Ronald Fisher's analysis of variance and experimental design, and the creation of the statistics department at Rothamsted Experimental Station. My interest for Fisher's work will not end in the early 1930s, when he moved from the agricultural institution to the Galton Laboratory at University College London. Chapter three of the thesis, in fact, will discuss the use of statistics, computing and information technologies in serology using as a case study the researches on the ABO blood groups – in particular the WWII

[37] On the contribution of G. Udny Yule to British agricultural science see B. Charnley (2011); on W. S. Gosset and the statistics department at Guinness see L. McMullen (1939). *Statistical Methods for Research Workers* was the first statistical textbook written by Ronald Fisher (see A. W. F. Edwards (2005a)).

[38] "Dr. E. S. Pearson, D.Sc., read a paper in December, 1932 (Journal, 1933, I), entitled 'A Survey of the Uses of Statistical Method in the Control and Standardization of the Quality of Manufactured Products'. In it he surveyed the work of Dr. Shewart of the Research Laboratories of the Bell Telephone Company New York, of the British Standards Association, and of others. The interest aroused by this paper and the discussion thereon was so great that the Council on 12 April 1933, authorized the formation of a Section of the Society under the title of 'The Industrial and Agricultural Research Section' and appointed a Committee under the Chairmanship of Dr. E. C. Snow (Honorary Secretary) to organize the Section. (*Royal Statistical Society Annals*, pp. 201-202)

[39] R. A. Fisher (1935), p. 272.

survey of the ABO blood group distribution in Britain – undertaken by Ronald Fisher and his co-workers at the Galton Laboratory.

Besides Ronald Fisher's career, I will follow also the work of Frank Yates, the statistician who assisted Ronald Fisher at Rothamsted during the early 1930s and took over the department in 1933. Yates was chief statistician at Rothamsted for over thirty years retiring only in 1968 and in the last chapter of my thesis I will describe how Yates promoted the computerization of his department during the 1950s and 1960s. As Fisher's arrival at Rothamsted and the opening of the station statistics department mark the start of my research, Yates' retirement in the late 1960s is the ending point. Fisher and Yates worked together at Rothamsted only for a couple of years, but their scientific collaboration lasted until Fisher's death in 1962 and Yates was a fierce promoter of Fisher's statistical methods throughout his life. The *Statistical Tables for Biological, Agricultural and Medical Research*, which I examine in chapter two, are a relevant example of the scientific collaboration between the two statisticians.[40]

My choice to examine statistics in agriculture and biology has been guided by the experimental contexts in which Fisher's statistical methods were developed and by the scientific mission of Rothamsted Experimental Station and the Galton Laboratory. That said, it will be evident at several points in my account that analysis of variance and experimental design became popular in a wider range of disciplines, not only so-called soft sciences, but also disciplines in which quantification processes had been in use for longer time. In fact, once Fisher's statistical methods had proved their value in agriculture and biology, many more users even in the hard sciences became interested in them and Fisher's methods found application beyond the areas for which they had been developed. For instance, if the tabular matter in Fisher and Yates' *Statistical Tables* was prepared having in mind biological, agricultural and medical research, even chemical industries and engineering associations showed their interest for reprinting the tables in Fisher and Yates' book.

IV. *Writing the history of statistics bottom-up*

Statistics is an information intensive and a computationally intensive discipline. Systems for storing and retrieving data as well as computing tools to ease the burden of the calculations required by the statistical examination of such data – calculations as simple as making averages or the more complex computations required by correlation and the analysis of variance – are thus

[40] R. A. Fisher and F. Yates (1963).

integral to statistics and its developments and I am going to approach the history of statistics starting from these technologies.

My choice to write the history of statistics bottom-up, looking at computing instruments and information technologies, rather than top-down, beginning with the theoretical achievements, will offer an historical account in which science and technology merge and in which statistics, computing and data management are interrelated activities in the mathematization of agriculture and biology.

Historical accounts, instead, have generally paid scarce attention to the technological and material constraints that influenced the development of statistics. A notable exception to this general trend is the analysis of the census of England and Wales made by Edward Higgs. The year 1911 was for the census examined by Higgs a statistical "Big Bang" with a consistent growth in output not only in terms of quantity, but also of quality: new areas connected with public health, such as marital fertility, were investigated and new socio-economic groupings adopted for classifications.[41]

According to Higgs the census transformation was prompted by its mechanization. In 1911 punched-card equipment was adopted for the first time in the analysis of the census returns and machine tabulation proved crucial in "providing new forms of information, and older forms in greater detail".[42] Higgs's argument, however, has little to share with technological determinism, because the historian does not attribute agency to the technology alone, but rather describes the transformation of the census as "the result of the complex interaction of technology, intellectual debate, and administrative necessity".[43]

A similar attitude towards the use of information and computing technologies in statistics can be found in JoAnne Yates' *Structuring the Information Age*.[44] In the book Yates examines the adoption at first of punched-card equipment and later of digital computers in life insurance, an industry guided by managers trained in actuarial statistics and able to assess risk "by determining who to insure, what products to offer, and what prices to charge for them".[45] Again, as in Higgs' case, Yates' argues that life insurance as a major user of computing and information technologies was not a technology-driven business, but rather that the industry and the technology mutually

[41] E. Higgs (1996), p. 409.
[42] E. Higgs (1996), p. 425.
[43] E. Higgs (1996), p. 425.
[44] J. Yates (2008).
[45] J. Yates (2008), p. 19.

shaped each other and that the computerization of life insurance was largely indebted to the previous adoption of punched-card equipment.[46]

The case studies mentioned so far are concerned with census and actuarial statistics, but computing equipment and information technologies are the tools of the trade in statistics, regardless of the subject matter. Indeed in my thesis I will argue on the one hand that data management and number crunching should be considered activities integral to statistics in agriculture and biology since the 1920s and on the other that the availability of data archives and suitable computing equipment have been crucial for the development of statistics throughout the half century here considered. Moreover, my in-depth analysis of computing tools for statistics offers the opportunity to see behind the apparent objectivity of number crunching how subjective ideas could be embodied in such instruments, as in the case of the five per cent threshold for statistical significance suggested by Ronald Fisher and faithfully promoted by the format he adopted in his tables of the chi-square and Student's distribution.

Fisher and Yates, the statisticians that I am going to examine closely in my research, were both actively engaged in data management and computing. Ronald Fisher, in particular, claimed to have learnt most of his statistics on a calculating machine and started at Rothamsted Experimental Station the tradition that equated computing tools to research tools in statistics.[47] This tradition constantly brought to the agricultural station up-to-date computing equipment, supported the making of the *Statistical Tables for Biological, Agricultural and Medical Research*, whose first edition was prepared by the human computers at Rothamsted Experimental Station and the Galton Laboratory, and eventually prompted the early computerization of the Rothamsted statistics department under Frank Yates during the 1950s.

The contributions of Rothamsted Experimental Station to the development of British computing are thus undeniable, but they have received so far scarce attention by historians and in general there is very little available in the secondary literature on the adoption of computing instruments in agriculture. On this latter aspect, a notable exception is David Grier's work on the U.S. Department of Agriculture in the 1920s and the Iowa State College Computing Service directed by the statistician George Snedecor. Grier describes the use of punched-card equipment and the work of human computers in the calculation of least squares and how the problems offered by agricultural research provided a context for a pioneering effort in computerization, such as the manufacturing of the ABC computer by John Atanasoff.[48] My research on the Rothamsted statistics department complements Grier's work in the British context offering the

[46] J. Yates (2008), pp. 1-8.
[47] F. Yates (1966), p. 235.
[48] D. A. Grier (2007), pp. 159-169; pp. 225-229; D. A. Grier (2000).

opportunity to follow the development of agricultural computing from the age of desk calculators to digitization.

Not only computing tools, but also information technologies were constantly at the core of Fisher's and Yates' work. The data in my case studies range from collections of experimental yields to returns of agricultural surveys, from family pedigrees to blood donor records. The instruments that were used to manage these data were standardised forms, tables, punched and index cards and the record systems adopted were in themselves a pre-requisite for the following statistical examination of the data. For instance, I will argue that the yields and meteorological records held at Rothamsted Experimental Station offered to Ronald Fisher and his co-workers suitable material for testing the analysis of variance. Information technologies were also crucial in the examination of the agricultural surveys that the Rothamsted statistics department undertook under Frank Yates and I will describe in chapter four how the computerization of survey analysis required an *ad hoc* adjustment in the input device of the Elliott 401 to allow the use of punched-cards. There is therefore a strong continuity between computing tools and information management in statistics.

Besides the study of the material tools adopted in data management, the interaction between statisticians and information technologies can be investigated at a deeper level. In particular, I will discuss how Ronald Fisher considered recordkeeping part of his mission as statistician and geneticist. Under Fisher, in fact, the Rothamsted statistics department became the repository of the data archive collected at the station since the nineteenth century, a move that positioned the statistics department at the centre of the experimental research conducted at the station. Again, Fisher's ABO blood group survey during WWII benefited from Fisher's attention for recordkeeping and his prompt decision to collect the medical data acquired by the Emergency Transfusion Services.

V. *Structure of the thesis*

I investigate the interdependence of statistics, computing and information technologies in agriculture and biology using four relevant case studies, each one developed in a separate chapter of the thesis. Even though the chapters can be examined independently, there is an overall unitary narrative. The same statistical methods, statisticians and institutions are taken into account and several cross-references from one chapter to the others build a common framework which spans the half century under consideration in my thesis and connects the different perspectives from which I examine the role of computing tools and information technologies in statistics for agriculture and biology.

In the choice of the case studies I have been guided by the aim to give on the one hand an overview of the main technologies – mechanical calculators, statistical tables, punched and index cards, standardised forms, digital computers – that were adopted in the period considered in my thesis, and on the other point out how these tools complemented each other and could be instrumental, as in the case of the *Statistical Tables*, for the dissemination of analysis of variance and experimental design.

In the first case study I examine the development of Ronald Fisher's statistical methods at Rothamsted Experimental Station and the impact of statistics on the research activity at the agricultural station. Fisher's career at Rothamsted is investigated as well as the organization of the department he founded there and the computing work he supervised in the institution. In the final section of the chapter I discuss the diffusion of analysis of variance and experimental design beyond Rothamsted and beyond agricultural science arguing for the generality of the statistical methods developed by Ronald Fisher.

Several are the research questions underpinning the case study. They are concerned with the contributions of statistics to agricultural science, the suitability of statistical methods for the analysis of agricultural experiments, the advisory services given by statisticians to the Rothamsted research workers and the role of computing equipment in supporting the adoption of statistics in agriculture. I will claim that Ronald Fisher's methods represented a concrete opportunity to estimate the experimental error and increase the precision of both field and laboratory experiments at Rothamsted Experimental Station and that the introduction of statistics reshaped experimental research in the institution. On the one hand, field and laboratory tools and practices had to receive the approval of the statisticians as to guarantee that experiments were planned and conducted according to sound statistical principles and on the other number crunching, either for the analysis of experimental results or for the preparation of computing tools for statistics, like mathematical tables, had to be accepted as an integral part of experimental research.

Although my case study refers to a specific institution and to a limited set of statistical methods, the conclusions are of general interest due to the leading role that Rothamsted Experimental Station had in British agricultural science and due to the success that analysis of variance and experimental design met well beyond the agricultural station. In appendix the list of visiting workers who came to Rothamsted to learn Fisher's methods offers an insight into the dissemination – both disciplinary and geographic – of analysis of variance and experimental design.

In the second chapter of the thesis my interest will shift from the development of analysis of variance and experimental design to the making of a computing instrument for their dissemination. I will discuss, in fact, the *Statistical Tables for Biological, Agricultural and Medical Research*, the collection that Ronald Fisher and Frank Yates co-authored during the 1930s to promote the application of Fisher's statistical methods. In the chapter I will provide a general background on the statisticians' contribution to the British tradition of table making before examining in detail the *Statistical Tables*. I am going to compare and contrast Fisher and Yates' book with its closer competitor, the *Tables for Statisticians and Biometricians* edited by Karl Pearson, and I will reconstruct the planning and development of the book throughout six editions and the liberal policy adopted by its authors for the reproduction of their tables in other statistical publications.

My aim is not only to investigate the making of the collection as a computing tool intended for statisticians and research workers. Rather, I want to discuss how the book contributed to spread Fisher's statistical methods in experimental research. The main question that the chapter addresses, in fact, is whether computing tools in statistics can be crafted as instruments for enforcing and promoting certain statistical methods over others, i.e. whether they can be political artefacts according to the definition given by Langdon Winner.[49] Overall, it will become evident that the answer to my research question is affirmative and that, despite the perceived objectivity of numbers and formulae, statistical tables are the result of human choices and in so far they can be imbued with specific values and ideas.

Besides this main research question, I discuss in the second chapter also Ronald Fisher and Frank Yates' as computers – a point that helps to clarify the relevance that they attributed to computing tools in statistics – and I use the requests for reprinting materials from the *Statistical Tables* as an instrument to track the diffusion of analysis of variance and experimental design. The second chapter, thus, complements the first one offering a further source of data on the diffusion over time of Fisher's statistical methods.

In the third chapter the focus will shift from the computing instruments to the information technologies adopted in statistics. Again Ronald Fisher will feature prominently in this case study, but the main setting will be provided now by the Galton Laboratory, the department that Fisher administered at University College London from 1933 to 1943 and where he founded in 1935, sponsored by the Rockefeller Foundation, a serum unit. In this case study I investigate how statistics, computing and information technologies contributed to the study of the human blood groups. I will examine in detail the survey of the ABO blood group distribution that Fisher

[49] L. Winner (1980).

undertook during World War II with the data provided by the Emergency Transfusion Services set up in Britain during the warfare. The research questions that underpin the chapter are related to the mutual relations between statistics and serology, the constraints and limitations that mathematization implied in blood group research, the problems related to data collection in genetics, the development of statistical methods for the analysis of blood group data, the reception of statistics among serologists.

As it will be evident from my account, information technologies for managing blood donor records had a key role in the making of the survey. I will trace the records employed in the warfare back to the information management adopted by the British Red Cross Blood Transfusion Service and I will discuss how the information practices employed for decades by the voluntary donor service were adapted to the necessities of the Emergency Transfusion Services. The information technologies, however, were a pre-requisite, but not the solution for the making of the ABO survey. I will discuss, in fact, how Fisher's statistical outlook was essential in order to turn data gathered for a medical aim into instruments for the study of British ethnography.

Beyond the analysis of the ABO survey, I will investigate serology as a further context for the use of statistical methods in experimental research. I will account why and how statistics was essential in the understanding of the ABO blood groups, which statistical methods were developed for extracting gene frequencies from phenotype ratios and how Fisher collaborated with the members of the serum unit he had founded and with the several physicians with whom he interacted during the wartime survey offering his statistical expertise in exchange for the transfusion records.

In the last chapter I am going back to where my quest started, that is the statistics department at Rothamsted Experimental Station, to account its computerization during the 1950s and 1960s. The statistician in charge of the transformation of the department was Frank Yates and the first mainframe used, the Elliott 401, remained at Rothamsted until the mid-1960s, when it was eventually acquired by the Science Museum in London as a valuable output of the British computer industry.

The computerization of the Rothamsted statistics department benefited from Frank Yates' involvement in operational research during WWII. During the conflict Yates established the network of alliances that were instrumental for the acquisition of the Elliott 401 in 1954. The mainframe, in fact, was leased to the agricultural station by the National Research Development Corporation, a government body instituted to promote the industrial exploitation of British inventions, and its acquisition was sponsored by the Agricultural Research Council, the main funding institution for agricultural research in Britain. In the decision of leasing the computer to

Rothamsted and in the choice to sponsor the acquisition, former acquaintances of Frank Yates in operational research were crucial. From this perspective the Rothamsted case can be added to the literature concerned with the role played by operational research for promoting computerization in the aftermath of WWII.

There is, however, a feature in the acquisition of the Elliott 401 that should not be missed. The computerization of the department was prompted by the idea, well established since Ronald Fisher's time, that statistics and computing are two interconnected activities and thus that computing instruments for the statistician are research tools and new developments in statistics depend on tinkering with these technologies.

Besides taking up the question 'Which difference did computers make to statistics?' – similarly to what already done by Jon Agar in his comparative study of computerization – I will also ask the opposite one, that is 'Which difference did statistics make to computers?', investigating in my case study the hardware and software development that were required at Rothamsted in order to adapt the Elliott 401 to the necessities of the department. In the final section of the chapter I will examine in detail some of the computer programs that were developed for dealing with the analysis of agricultural experiments and surveys and with genetic linkage. Overall, the account that will emerge from my narrative is far from the story of a computer revolution, but it will emphasize instead the role of the digital computer as a research instrument. Mainframes were accepted in statistics only after much tinkering and a preliminary exploration of their potentialities, and they did not replace outright the computing tools already in use in the field, such as desk calculators and punched-cards.

In the conclusion I will draw together the several threads interwoven in my account in relation to three main points: 1) the spread of statistical methods in agriculture and biology as a pattern for the mathematization of these disciplines; 2) the statisticians as a new expert group within agriculture and biology; 3) the role acquired by computing tools and information technologies in the process and their impact on research practices. Eventually, looking at the case studies examined in my thesis, I will be able to claim the relevance of computing tools and information technologies for the development and dissemination of statistical methods and to point out the mutual exchange of competences between statisticians and research workers as a necessary requisite for the mathematization of the soft sciences.

VI. ***Sources***

Great part of my research is based on archive materials and archives in three continents have contributed to it. A consistent set of documents is related to Ronald Fisher and an overview of the archival collections of interest on the statistician and geneticist is thus advantageous. Fisher spent the main part of his academic career in Great Britain working at first at Rothamsted Experimental Station in Harpenden, moving then to the Galton Laboratory at University College London and eventually to Cambridge University, from which he retired at the end of the 1950s. However, the archives of the institutions in which Fisher worked hold very little about him. The papers of Ronald Fisher are mainly conserved at the Barr Smith Library, University of Adelaide, as Fisher moved to Australia in 1959 and died in Adelaide in 1962.

If the documents held in Adelaide represent the more substantial collection of Fisher papers, there are other archival sources relevant for understanding the work of the statistician and geneticist. The value of the scientific correspondence between Fisher and the chemist William Gosset cannot be overestimated in understanding Fisher's research during the 1920s and 1930s. The original letters are conserved in the Archive of University College London, but there is also a published copy of these documents – the one I quote in my thesis – prepared by a former co-worker of Gosset in Dublin, Lance McMullen, and privately printed by the Guinness company.[50]

Another relevant collection of papers related to Fisher and, to my knowledge, never consulted before, is in the Oliver and Boyd collection held at the National Library of Scotland in Edinburgh.[51] The publisher Oliver and Boyd printed all Fisher's statistical textbooks – and almost all Fisher's books – and the journal *Heredity*, co-edited by Fisher and the geneticist Cyril Darlington. The extensive author-publisher correspondence of Oliver and Boyd has survived untouched and casts a new light on Fisher's scientific enterprise giving relevant information on the making and editorial success of *Statistical Methods for Research Workers*, *The Design of Experiments*, and the *Statistical Tables for Biological, Agricultural and Medical Research*.

At Rothamsted Research there are documents related to Fisher's early career (manuscripts of papers published in the *Journal of Agricultural Science* in the 1920s, sheets of calculations, graphs, annotations made in the preparation of papers) and to the work of his assistant statisticians during the 1920s and in the early 1930s. In Harpenden there have been two archives relevant for my research: the document repository of the former statistics department and the holdings of the main archive of the institution. Recently some of the historical materials held in the department

---

[50] W. S. Gosset (1962). Along with the letters there are also summaries of the correspondence prepared by Ronald Fisher, possibly while writing Gosset's obituary in the 1930s (R. A. Fisher (1939a)). The reference of the originals is UCL Special Collections, MS ADD 274.

[51] NLS, Oliver and Boyd Papers, Inventory Acc. 5000.

have been moved to the main archive to be merged with the documents there available on statistics. The catalogue of the whole collection in the main archive is now in preparation. In my thesis I will use, whenever possible, the new references attributed to the documents, otherwise I will conserve the references I found at the time of my research. When the cataloguing is completed, it will be possible to find correspondence between the old and the new references.

In relation to the development of analysis of variance and experimental design, I have used as main primary sources Fisher's published work – both his scientific papers and the textbooks in which he presented his methods to statisticians and research workers –, the correspondence with William Gosset, and the official reports of Rothamsted Experimental Station. To describe equipment and organization of Fisher's department during the 1920s and in the early 1930s, I have resorted to the accounting books of the institution for the acquisition of computing tools and to the correspondence between Fisher, Frank Yates, and the human computers of the department for determining tasks and main features of the assistant staff. In order to discuss how Ronald Fisher and his colleagues collaborated with the research workers in the station and the role that the newly founded statistics department gained in the institution, I have turned to the minutes of the Staff Council and the ones of the Field Plots Committee.[52] From them it is indisputable the key role that Fisher's department gained becoming essential in the set-up of field trials at Rothamsted, alongside the experimental farm and the departments in charge of the fieldwork.

In chapter two an essential piece of evidence has been provided by the correspondence between Ronald Fisher and his publisher, Oliver and Boyd. In the letters exchanged not only in relation to the *Statistical Tables*, but already for the appendix of tables that appeared in Fisher's first textbook, *Statistical Methods for Research Workers*, it is evident that Ronald Fisher attributed to these computing tools a strategic value in the promotion of his methods and that the liberal policy for granting permission to reproduce materials from the book was based on such assumption. The correspondence in the Oliver and Boyd papers, along with the letters exchanged by Fisher and Yates and a set of documents conserved in the archive of Rothamsted Research and related to the second and third edition of the *Statistical Tables* have provided the materials for compiling an extensive table, summarized in Appendix 2.3, with the requests for reprinting materials from the book.[53]

In chapter three I have resorted to the official publications related to statistics in the study of the ABO blood groups, in particular the papers of Felix Bernstein, the first statistician who

52 RR Library and Archive, Ref. STA 2.1; RR Library and Archive, Ref. FX.

53 NLS, Oliver and Boyd Papers, Inventory Acc. 5000; RR Library and Archive STATS 7.5; Fisher Papers, BSL, the University of Adelaide; RR Library and Archives, STF 31 (old catalogue reference).

contributed to the field, to the publications of the Galton Laboratory Serum Unit before WWII, and to the works that Fisher and his co-workers published on the ABO survey during the warfare and in its aftermath. Official papers, however, would have been unable to offer a comprehensive account of the wartime survey and of the lifetime collaboration between Ronald Fisher and serologists. For highlighting these aspects I had to resort to several archival collections. The correspondence between Ronald Fisher and George Taylor (the chief serologist of the Galton Laboratory Serum Unit), and the Blood Group Survey correspondence – both held in the Fisher Papers at the University of Adelaide – have provided an insight into the organization of the wartime survey.[54] As the Galton Laboratory Serum Unit started in 1935 with funding of the Rockefeller Foundation, also the foundation archives in New York have offered relevant material on the set up of the unit and the early war years.[55]

In the Wellcome Library London, I have had access to the papers of Percy Lane Oliver, the organizer of the voluntary blood donor service associated to the British Red Cross. At the Wellcome Library I have also examined the papers of Janet Vaughan, a physician engaged in the set up of the Emergency Transfusion Services; the papers of the blood transfusion service for the years 1943-1946; the collection of the Medical Research Council Blood Group Unit, in which there is correspondence both related to the early years of the Galton Laboratory Serum Unit and the wartime collaboration between Fisher and George Taylor.[56] Also The National Archives of the UK have offered interesting materials related to the Emergency Transfusion Services and to the Galton Laboratory Serum Unit during the war years. At The National Archives I have found samples of the donor cards adopted by both the emergency services and by the co-workers of Ronald Fisher in WWII.[57]

As mentioned in the review of the literature, I have resorted to primary sources for a comprehensive description of the computing equipment available at the Galton Laboratory during Fisher's time. A precious gift from the repositories of the Record Office of University College London has been represented by the ledger book of the Galton Laboratory in the years 1934-1939. The accounting book neatly records in ordered columns the expenses for feeding the laboratory animals, alongside with the income from the *Annals of Eugenics*, the journal published

[54] Fisher Papers, BSL, the University of Adelaide. Both the collections are partly digitized
- http://digital.library.adelaide.edu.au/dspace/handle/2440/67586;
- http://digital.library.adelaide.edu.au/dspace/handle/2440/68047.
[55] RA Center: Collection RF, Record Group 1.1, Series 401 England, Box 16, Galton Laboratory.
[56] WL London: Percy Lane Oliver Papers, uncatalogued (my attention has been drawn towards these materials by the archivist Amanda Engineer, whom I thank for the suggestion); Dame Janet Vaughan Papers, Ref. GC/186; Blood Transfusion Service (1943-1954), Ref. GC/107; MRC Blood Group Unit (1935-1995), Ref. SA/BGU.
[57] TNA: Medical Research Committee and MRC: Files, Ref. FD/1; MRC: Blood Group Unit: Reports and Papers, Ref. FD/8.

by the laboratory, and the expenditures for the publications acquired and the reprints bought. The final section of the ledger is reserved to the purchase of calculating machines and from it we can evaluate how much of the laboratory funds were expended into this kind of equipment. Unexpectedly, the accounting book has proved also a systematic use of punched-cards in Fisher's Galton Laboratory, an aspect not mentioned anywhere in secondary sources.

For the computerization of the Rothamsted statistics department in the fourth chapter I have mainly referred to the official reports of the experimental station and the papers held in the collection of the National Research Development Corporation at the National Archive for the History of Computing in Manchester.[58] Further information has been provided by the reports and minutes of meetings of the Agricultural Research Council held at The National Archives and by the Science Museum technical file that gathers the documentation with which the Elliott 401 was acquired by the institution in London.[59] Some documents on the accommodation of the statistics department are still conserved in the archive of Rothamsted Research and have been useful to reconstruct the whereabouts of the Rothamsted statisticians after WWII.[60] Regretfully, instead, none of the correspondence of the statistics department during Frank Yates' management has survived. Overall, only a small part of Yates' papers has been preserved after his death and among them there is surprisingly little correspondence, apart from the letters he exchanged with Ronald Fisher.[61] Over thirty years of scientific research have been lost in the several relocations that the Rothamsted statistics department has undergone after WWII, denying to the historian a precious source of information. As a partial compensation to this loss, I have had the opportunity to interview a couple of the former members of the staff department and their accounts have been reported in the appendices to the chapter.

Any historical research is partly a dialogue with the existing literature and partly an attempt to go beyond the established vision rummaging in archives and libraries and looking for sources able to offer novel perspectives. My work conforms to this rule, but I must say that the best ideas and suggestions for my research have come from the time spent in the archives. Both unexpected findings and inexplicable absences of some documents have offered me food for thoughts and valuable insights on the matter I was examining.

[58] UK NAHC, Manchester, Ref. NAHC/NRD.
[59] TNA, Ref. MAF 117; MAF 200; SM London T/1965-445.
[60] RR Library and Archive, Ref. SITE 1.5.2.
[61] Frank Yates Papers (RR Library and Archive, STF 31 (old catalogue reference)); Fisher-Yates correspondence (RR Library and Archive STATS 7.5 and Fisher Papers, BSL, the University of Adelaide).

Chapter 1

ANALYSIS OF VARIANCE, EXPERIMENTAL DESIGN AND THE RESHAPING OF RESEARCH AT ROTHAMSTED EXPERIMENTAL STATION, 1919-1933

1.1 *Introduction*

The experimental station of Rothamsted was set up in 1843 by John Bennet Lawes, English squire, amateur chemist and successful businessman in the fertilizer industry. In association with the professional chemist Joseph H. Gilbert, Lawes promoted and generously funded on the fields of his private estate, Rothamsted, agricultural experiments mainly concerned with the effects of fertilizers on the growth of crops. The scientific activity at Rothamsted did not end with the death of Lawes and Gilbert at the turn of the twentieth century, but went on in the following decades and the experimental station – the oldest British institution of its kind – became a landmark of agricultural science in Britain.[62]

In 1919 Rothamsted Experimental Station hired its first statistician, the mathematician Ronald Fisher. Fisher was engaged in the examination of the yields and weather records collected at the station since the mid of the nineteenth century and as a consultant of the local research workers in the examination of their experimental results and in the planning of the Rothamsted field trials. To deal with the problems arising in agricultural experiments Fisher developed during the 1920s the statistical methods of analysis of variance and experimental design, which represented his main contributions to applied statistics.

The emergence of the statistician as consultant for the research workers of the experimental station went hand in hand with the creation of an autonomous statistics department that Fisher set up from scratch during the fourteen years he spent at Rothamsted. The department distinguished itself also for its computing activity because gathering information from the results of agricultural experiments was a matter of computing labour, as well as mathematical skills, and calculating machines and human computers contributed to the examination of the station records and to the making of statistical tables useful for the application of analysis of variance and experimental design.

Ronald Fisher's acquaintance with the research activity in the experimental station and the challenges presented to him by the agricultural and biological experiments conducted there were

[62] On the history of Rothamsted Experimental Station and its founding fathers see E. J. Russell (1966). Biographies of Lawes and Gilbert are, respectively, F. M. L Thompson (2004) and E. Clarke [rev. A. E. Johnston] (2004).

crucial in the development of analysis of variance and experimental design. On the other hand, the adoption of Fisher's statistical methods at Rothamsted did not entail just a formal change in the examination of the station data, but analysis of variance and experimental design reshaped in depth the research activity of the institution requiring a different outline for the annual reports of the station and the adoption of new field practices and instruments, and imposing a redistribution of expertise among statisticians, research workers and the farm staff.

Thus, in my historical account Ronald Fisher is not presented as a mathematical statistician, but rather as a consultant of the Rothamsted research workers and his statistical methods are framed into the practices of experimental research adopted at the agricultural station. In the literature currently available on Ronald Fisher this is a neglected perspective, with the evident bias that Fisher's work at Rothamsted appears independent from the institution in which it was accomplished.[63]

For instance the preface to Fisher's collected papers introduces the statistician and geneticist as the man who probably had the greater impact on the methodology of scientific research during the twentieth century, but Fisher's published papers can do little to instance this statement, as they were rarely co-authored with experimenters and often adopted a sophisticated mathematical formalism, beyond the grasp of most research workers.[64] The dissemination of analysis of variance and experimental design was rather promoted by Fisher's textbooks, written without mathematical proofs, arranged as laboratory handbooks with solved examples and supplied with mathematical tables useful for computations.

Even in Fisher's selected correspondence, edited as the collected papers by the Australian geneticist Henry J. Bennett, there is no room for a systematic exploration of Fisher's engagement as statistical consultant of research workers. Most of the book is related to statistical inference and the correspondents are Fisher's fellow mathematicians interested in statistical theory and method.[65]

The same attitude is common in Fisher's obituaries. For example, the eulogy published in the *Biographical Memoirs of the Royal Society* mentions Fisher's involvement with the research workers at Rothamsted only as a brief introduction to his mathematical contributions to the design and analysis of experiments. However, the statistician Frank Yates, who co-authored the obituary,

---

[63] To some extent, an exception to this trend is the work of N. S. Hall (2002, 2007) on Fisher and randomisation. Besides discussing Fisher's reliance on *n*-dimensional geometry and his interest for the mathematical analysis of small samples, Hall provides also an account of experimental design before Fisher and of the research activity at Rothamsted.

[64] J. H. Bennett (1971), pp. 3-4.

[65] J. H. Bennett (1990).

worked at Rothamsted for over three decades, with Fisher and afterwards, and knew well how statistics was entrenched in the research practices of the institution.

Also several celebrations of Ronald Fisher's work, appeared later, presented him just as an outstanding mathematician. For instance, the essay on Fisher written in the 1970s by Leonard J. Salvage is a "re-reading" of Fisher conceived by a statistician for statisticians, and the same can be said of the collection of essays, *R. A. Fisher: An Appreciation*, published in the 1980s as a series of lecture notes on statistics.[66] In the *History of Mathematical Statistics from 1750 to 1930* written by the statistician Anders Hald in the 1990s, and in the later book from the same author, *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935*, Fisher is examined as well only for the role he had in mathematical statistics.[67] Even in the more recent *Fisher, Neyman and the Creation of Classical Statistics* by Eric Lehmann only Fisher's mathematical contributions to statistics are taken into account and discussed in relation to the theory of hypothesis testing developed by the statisticians Egon Pearson and Jerzy Neyman.[68]

Ronald Fisher's biography, written by the daughter Joan, offers some insights for reconstructing Fisher's engagement with the research workers at Rothamsted Experimental Station, but such interactions represents only the background for Fisher's portrait as a mathematical genius that Joan Fisher Box wants to set before the reader.[69]

If Fisher has been presented as an outstanding mathematician by his fellow statisticians and by his main biographer, on the other hand historians have mainly investigated Fisher as a founding father of population genetics and an active eugenicist.[70] For instance, William Provine has discussed Fisher, along with J. B. S. Haldane and Sewall Wright, in his classical study on the origins of theoretical population genetics. Donald Mackenzie in *Statistics in Britain* links Fisher's work in biology to his eugenics ideas; on Fisher and eugenics, and more broadly on Fisher's contributions as statistician and geneticist to the research on human blood groups has written extensively Pauline Mazumdar; James Moore has drafted a biographical sketch considering Fisher's family relations, his eugenic ideas and his Anglican faith.[71] Philosophers, like Ian Hacking

---

66 L. J. Salvage (1976); S. E. Fienberg and D. V. Hinckley (1980).
67 A. Hald (1998).
68 E. L. Lehmann (2011).
69 J. Fisher Box (1978), pp. 93-112; 140-166; 235-256. Fisher's daughter is the wife of the statistician George E. P. Box, whose advisory role in the making of Fisher's biography is openly recognised (J. Fisher Box (1978), p. *vii*). George Box's expertise might have influenced the biographical portrait of Fisher as statistician and mathematician rather than geneticist and eugenicist.
70 The following list does not aim to be a complete bibliography of Fisher's contributions to eugenics and population genetics, but is limited to the authors later mentioned in the chapter.
71 W. B. Provine (2001); D. A. Mackenzie (1981); P. M. H. Mazumdar (1995, 2011); J. Moore (2006).

and Teddy Seidenfeld, instead, have been interested in Fisher for his contributions to the theory of statistical inference.[72]

Overall, thus, the main interest for Ronald Fisher's work in applied statistics has been granted to his theoretical achievements leaving aside the "practical demands" associated with the development of statistical theory in the early decades of the twentieth century.[73] This is a serious limitation when referred to analysis of variance and experimental design, which were born in the context of an experimental station and had their usefulness and justification not in the realm of mathematical theory, but for their application at first to agriculture and than to a wider set of experimental disciplines.

Analysis of variance and experimental design, in fact, provided new opportunities for gathering information in agricultural research: they made more lab-like field experiments offering an estimation of their error and increasing the overall precision, but they proved also to be real field practices because they allowed the investigation of several factors at once, evidently a non lab-like condition. Moreover, analysis of variance and experimental design entered also in the practice of laboratory workers at Rothamsted – bacteriologists, entomologists, chemists – increasing their accuracy.

The application of Fisher's statistical methods was not limited to the experimental work at Rothamsted. They became popular worldwide due to the research workers that came to the station to learn Fisher's methods and introduced them in the scientific life of their own institutions. Fisher's lecture courses abroad and his textbooks, *Statistical Methods for Research Workers* (first edition 1925) and *The Design of Experiments* (first edition 1935), as well as the *Statistical Tables for Biological, Agricultural and Medical Research* (first edition 1938, co-authored with Frank Yates), contributed to spread further the application of his statistical methods.

The chapter is structured in five sections. I will start with an overview of agricultural science in Britain, before discussing the early applications of statistics to the analysis of agricultural experiments and the development of statistical methods *ad hoc* for agriculture. I will then present in detail analysis of variance and experimental design, and the contributions given by the ecologist Thomas Eden and the plant physiologist Ernest J. Maskell to their field implementation. My claim is that Fisher's statistical methods reshaped research at Rothamsted Experimental Station. To give evidence for my argument I will consider the role that the newly formed statistics department gained in the institution in relation to the format and management of the station records and to the activity of the Field Plots Committee, the body in charge of the set-up of the

[72] I. Hacking (1976); T. Seidenfeld (1979).
[73] D. A. Mackenzie (1981), p. 213. See also the thesis introduction, section III.

station experiments. I will also deal with the new outline of the station reports and the improved field practices required by analysis of variance and experimental design. I will claim that Fisher's statistical methods set new criteria of precision in agriculture providing a link between laboratory and field research at Rothamsted. I will then discuss the diffusion of Fisher's statistical methods beyond agricultural science and their dissemination worldwide.

1.2 *Agricultural science in Britain*

1.2.1 *The case of Rothamsted Experimental Station*

In his account of British agriculture before WWI the historian Paul Brassley qualifies as agricultural science only the enterprises "professional and disinterested" concerned with "*explaining* why the best practice [in agriculture] was so, or in finding out how things worked in order to produce *better* practice".[74]

According to Brassley's definition, Rothamsted Experimental Station was certainly an institution in which agricultural science had been practiced since the very beginnings because the Rothamsted founding fathers, Lawes and Gilbert, "attempted to explain how things worked by the application of skills or techniques not generally available to farmers in order to benefit the community in general".[75] The original goal set for Rothamsted at the mid of the nineteenth century remained unchanged for at least a century. The policy "to develop an agricultural science that experts and teachers could use in their daily work and that would stimulate good farmers to think and so to devise new and better methods of agricultural practice" was still recognised as "Rothamsted's chief claim to distinction" during the centenary celebrations of the station in 1943.[76]

Even though the policy of the station did not change, its implementation developed consistently during the first decades of the twentieth century. Agricultural science as practiced by Lawes and Gilbert was "mainly a branch of chemistry".[77] Gilbert was a pupil of the German chemist Justus von Liebig and Lawes and Gilbert began their collaboration disproving Liebig's report on *Organic Chemistry in its Application to Agriculture and Physiology*.[78]

Alfred D. Hall, the director who took over the management of the station in 1902, had instead a broader idea of agricultural science and he added to the Rothamsted staff a microbiologist, a

---

[74] P. Brassley (1995), p. 467.
[75] P. Brassley (1995), p. 467.
[76] RES (1943), p. 2.
[77] RES (1943), p. 2.
[78] J. von Liebig (1840). Gilbert was awarded a PhD in 1840 for its work in Liebig's laboratory in Giessen.

botanist, an organic chemist and a soil scientist.[79] The soil scientist was E. John Russell, who became the third director of Rothamsted in 1912. Under Russell, the scientific activity of the agricultural station expanded further.[80] In the aftermath of World War I Rothamsted had in total ten laboratories (bacteriological, botanical, chemical, fermentation work, antiseptics and insecticides, protozoological, statistical, entomological, mycological), a farm and experimental fields, and employed about fifty people.[81] The research activity of the station at that stage was "in the main restricted to the soil and the growing crop, leaving such questions as animal nutrition, feeding values, plant breeding, dairying, etc., to other Institutions in the country".[82]

The sustained growth of the scientific staff, from a handful of people until the early 1900s to about thirty in 1920, was prompted by the new funding opportunities offered to Rothamsted in 1911. Up until then Rothamsted had been a private enterprise supported by funding of the Lawes Trust and a few other sponsors, while from 1911 onwards the station received public grants of increasing amount provided by the Development Fund.[83]

The historian Robert Olby has strongly argued for the key role played by the Development Commission – that administered the Development Fund – in the growth of agricultural research in Britain during the first decade of the twentieth century.[84] Besides the provision of funding to agricultural institutions already in operation, the programme of the Development Commission consisted in the creation of a group of research institutes concerned with specific branches of agricultural science, such as the plant breeding institutes examined in detail by Paolo Palladino and Berrys Charnley.[85] The Commission set up also an advisory service to promote a closer collaboration between agricultural colleges, research institutions and farmers, and the training of new workers in agricultural science.[86]

Brassley supports Olby's conclusion on the relevance of the Development Commission, but suggests also that the change in agricultural science in Britain – in terms of people, funding and

---

79 For A. D. Hall's contributions to agricultural science see his obituary as fellow of the Royal Society (E. J. Russell (1942)) and the biographical entry in the *Oxford Dictionary of National Biography* (P. Brassley (2004)).
80 Russell's obituary as fellow of the Royal Society offers a biographical sketch of his work at Rothamsted (H. G. Thornton (1966)). A first hand account of Russell's life can be found in his autobiography (E. J. Russell (1956)).
81 RES (1921), pp. 4-5.
82 RES (undated), p. 1.
83 RES (1913), p. 3. The Rothamsted reports during the 1910s and 1920s record the grants received by the Development Commission for each year. In 1912 £2,500 were given to Rothamsted, while in 1919-1920 the grant had totalled almost £14,000. For the funding required by an experimental station see P. Brassley (1995), pp. 472-473.
84 R. Olby (1991).
85 P. Palladino gives an overview of the plant breeding institutes sponsored by the Development Commission in England, Wales and Scotland (P. Palladino, 1990), while Berris Charnley offers a closer examination of the Plant Breeding Institute and the National Institute of Agricultural Botany in Cambridge (B. Charnley, 2011).
86 Report of the Committee on Agricultural Research Organisation, Economic Advisory Council, 29th April 1930, p. 1, TNA, CAB/123/275.

results – was a much longer phenomenon than suggested by Olby.[87] University departments concerned with agricultural science were established since 1889 and in 1905 appeared the *Journal of Agricultural Science*, "devoted wholly to definitely scientific papers in agricultural subjects".[88] This was the right environment for the effective use of the support given to agricultural science by the Development Commission.

Both Olby and Brassley emphasize the role in the Development Commission of the second Rothamsted director, Hall, who had a distinguished career as agricultural scientist and civil servant. According to Olby, Hall gave direction to the commission in the coherent distribution of funding. The money was awarded in the form of 'block grants' to the successful institutions and Rothamsted largely benefited from this system.[89]

In 1931 the Agricultural Research Council (ARC) took over the management of agricultural science from the Development Commission on the recommendation that "a body mainly composed of men of high scientific standing in one or other of the basic sciences that serve agriculture" could meet at best the needs of agricultural research and promote the cooperation among the existent research organisations in Britain and the Empire.[90] Rothamsted became one of the grant-aided institutes of the ARC, referring now to the new body for funding. But the investigations at the research station went on as before, mainly concerned with crop production and use. They involved the study of "the growth of the plant in health and disease, its nutrition, its reaction to soil and climatic conditions, and its composition under various conditions".[91]

1.2.2 *Experiments in agriculture: assets and limitations*

A comprehensive history of agricultural experimentation has yet to be written.[92] This brief outline focuses only on the key aspects of field trials on crops and fertilizers as they represented the research activity associated with the development of analysis of variance and experimental design.

---

[87] R. Brassley (1995), pp. 476-477.

[88] P. Brassley (1995), p. 476. The quotation reported by Brassley is taken from the editorial published in the first issue of the *Journal of Agricultural Science*.

[89] R. Olby (1991), pp. 519-520. Rothamsted and Cambridge University, where a department of agriculture had been created since 1899, received the main grants offered by the commission.

[90] Report of the Committee on Agricultural Research Organisation, Economic Advisory Council, 29th April 1930, p. 6, TNA, CAB/123/275 (Quotation reproduced with permission from the BBSRC). On the history of the Agricultural Research Council see G. W. Cooke (ed.) (1981).

[91] RES (1931b), p. 3. An outline of the research activity at Rothamsted Experimental Station and the other British institutions for agricultural research during the 1930s is IIA (1933), pp. 78-97.

[92] As my aim is to discuss the introduction of statistical methods in the planning and analysis of field experiments, I am going to rely mainly on three (partial) reviews of experiments in agriculture written with the same attitude: W. G. Cochran (1976); G. Gigerenzer et al. (1989), pp. 70-72, pp. 84-86; and N. S. Hall (2002), pp. 34-49.

To simplify the discussion I am going to consider a specific experiment, the Broadbalk wheat trial, started at Rothamsted by Lawes and Gilbert in 1843 during their controversy with Liebig. According to Liebig, if a plant was supplied with the mineral constituents that remained as ashes when it was burnt, the plant did not require other manure, but extracted the carbon and nitrogen it needed from the atmosphere. The Broadbalk wheat experiment tested and disproved such hypothesis using several combinations of mineral manures (potash, soda...), with or without nitrogen, and comparing over a series of years the wheat yields of the experimental plots.

Besides testing Liebig's theory, the experiment – still in progress today – more generally assessed the effectiveness of different manures in the cultivation of wheat. In addition to the manured plots, since the beginning of the experiment, there were also two control plots – one left unmanured and the other dressed only with farmyard manure – to be used as comparison on the effectiveness of the combinations of fertilizers.

Comparison is a key point in agricultural experiments because uncontrollable factors such as soil fertility, climate, weeds, pests and drainage can heavily affect the results of a trial from season to season. Already in the eighteenth century Arthur Young, author of an extensive treatise on experimental agriculture, stressed this point when comparing in his book the practice of sowing wheat by the drill and the previous habit of broadcasting the seed. Young was also aware that the results of a one-off trial were not to be trusted. Thus he repeated his comparison in several experiments over five years.[93]

By the middle of the nineteenth century, when the Broadbalk wheat experiment began, the practice of comparative and repeated experiments was well established in agriculture. At that stage, it was also explicitly recommended – as in James Johnston's book *Experimental Agriculture*, published in 1849 – that repetitions of the same treatment were not distributed in adjoining plots to avoid biases due to soil fertility or exposure to sunshine.[94]

In the Broadbalk wheat experiment each treatment was applied to just one plot, but the same manurial scheme was repeated year after year and the continuation of the experiment over a long series of years represented a resource against the high variability of the elements under investigation.[95] The necessity to collect the results of long-term field experiments prompted the

[93] A. Young (1771). Young's experiments are mentioned in W. G. Cochran (1976), pp. 4-5.
[94] J. F. W. Johnston (1849). Johnston's work is examined in both W. G. Cochran (1976) and G. Gigerenzer et al. (1989).
[95] A. D. Hall (1905), pp. 33-34. The plot is the experimental unit of a field trial. Its size and shape can usually be adjusted by the experimenter according to convenience, except for constraints related to the natural conformation of the field. The plots in the Broadbalk field are long narrow strips that stretch across the length of the field, separated by uncropped paths. For the history of the Broadbalk wheat experiment and its rearrangements over time see the *Guide to the Classical and Other Long-term Experiments, Datasets and Sample Archive* prepared by Rothamsted Research (RR, 2006).

creation of a data archive at Rothamsted. Besides the crops data, the experimental station began to collect also meteorological records. The rainfall at Rothamsted was recorded since 1853, barometric and temperature records started in 1873 and since 1891 also observations of the sunshine were taken daily.[96]

According to Paul Brassley, the collection of field data on the efficacy of fertilizers was probably Lawes and Gilbert's major contribution to agricultural science.[97] Certainly the record archive was a key element in the experimental life of the station and it contributed to the adoption of statistical methods during the 1920s, as I will discuss in relation to Ronald Fisher's appointment.

The results gained in the Broadbalk wheat experiment were not of general validity, but could be trusted only for the trials conducted on soils similar to the "cold sticky clay" of the Rothamsted field.[98] Thus, agricultural experiments were usually replicated on several farms in order to gain an overview of the soils of the region. At Rothamsted this practice was routinely adopted after the acquisition in 1921 of the experimental farm of Woburn, because it became possible to compare the efficacy of fertilisers on crops, in particular for the long-term wheat and barley trials, in the sandy soil of Woburn and in the clay soil of Rothamsted.[99]

With the rediscovery of Mendel's work at the beginning of the twentieth century genetics became a further discipline able to contribute to agricultural science for plant and animal breeding.[100] Up until the 1930s the role of genetics at Rothamsted Experimental Station was ambiguous and it had not been decided yet "whether work of a genetical nature was in order or *ultra vires* at Rothamsted".[101] This disregard of genetical questions had consequences also on the field experiments, as it was not immediately perceived that it was necessary to use the same crop variety throughout time. For example in the Broadbalk wheat experiment, from 1852 to 1918, eight different wheat varieties had been used, some for just one year, others for decades.[102] In 1917 the variety Red Standard was adopted with the aim to settle on it for the future, but already in 1929 the Rothamsted director John Russell enquired with the plant breeder Rowland Biffen about a possible substitute, as the Rothamsted farm manager complained that Red Standard was

[96] A. D. Hall (1905), p. 15.
[97] P. Brassley (1995), p. 468.
[98] Letter from E. J. Russell to R. Biffen, 3rd October 1929, E. J. Russell Papers, RR Library and Archive, RUS 2.7.
[99] RES (1923), p. 26.
[100] For an overview of Mendelism in plant breeding in Britain and about Rowland Biffen, mentioned below, see B. Charnley (2011).
[101] Report of the Meeting of the Sub-committee on animal husbandry, 2nd February 1932, RR Library and Archive, STATS 6.3.
[102] R. A. Fisher (1921), p. 108.

"liable to become badly lodged, particularly on the high nitrogen plots".[103] A new change of variety was an intrinsic problem for the comparison of the experimental results and Biffen advised against such a change. "One does not know how much one is monkeying up the results by using different wheats", wrote Biffen to Russell, and this was a weakness for an experiment whose value relied in the repetition, year after year, of the same scheme.[104]

1.3 ***Statistics in agricultural science***

As described for the Broadbalk wheat experiment, several factors are involved in a field trial on crops and fertilizers and it is difficult to ascertain their respective influence on the final result. The replication of experiments and accumulation of data is not enough to decide whether the results gained in a specific case can be generalized, how the experimental error has to be assessed, when results are comparable and which is their overall precision.

At the turn of the twentieth century these inner limitations were a stumbling block for the development of agricultural science because more precision was needed for giving advice to farmers, as a variation of five per cent in the gross yield could make all the difference between profit and loss.[105] Statistical methods were called into cause to overcome this limitation and to attempt a better quantification of agricultural experiments.

1.3.1 ***Before Ronald Fisher: a review with speculations***

The first systematic interest for the statistical treatment of the experimental results arising in "some chemical, many biological and most agricultural and large scale experiments" in which "it is sometimes necessary to judge of the certainty of the results from a very small sample, which itself affords the only indication of the variability" can be traced back to the chemist William Gosset.[106]

Born in Canterbury in 1876, Gosset was educated at New College, Oxford, where he obtained a first in Mathematical Moderations in 1897 and two years later gained a first class degree in chemistry. In October of the same year he began to work for Arthur Guinness, Son & Co. at the

[103] Letter from E. J. Russell to R. Biffen, 3rd October 1929, E. J. Russell Papers, RR Library and Archive, RUS 2.7.
[104] Letter from R. Biffen to E. J Russell, 4th October 1929, E. J. Russell Papers, RR Library and Archive, RUS 2.7.
[105] RES (1925), p. 38.
[106] Student (1908), p. 2.

St. James's Gate Brewery in Dublin. There he remained until the end of 1935, when he was appointed head brewer at the new Guinness plant of Park Royal in London.[107]

The research work in which Gosset was engaged concerned agricultural experiments and mainly trials in an experimental brewery, and he suggested to the Guinness's board that the application of statistical methods could be an asset in the analysis of these experiments. The company, interested in a scientific approach to brewing, granted the chemist a leave of absence in 1906-1907 to spend two terms in the department of applied statistics headed by Karl Pearson at University College London.[108] Gosset's experimental problems were not of direct concern for Pearson's biometric school, as the so-called biometricians were only interested in large sets of experimental data, but Pearson with his mathematical expertise was able to help Gosset in his attempt to find a statistical approach to tackle the variable and sparse material offered by the experimental brewery.[109]

Gosset introduced the z distribution – nowadays more popular as the t distribution – and computed the related tables of the probability integral (**chapter 2**) with which it was possible to examine the accuracy of a small sample of experimental results. Under the pen name of Student, adopted to comply with the publication policy of Guinness, Gosset presented in 1908 his distribution and related tables in *Biometrika*, the journal edited by Karl Pearson.[110] *Biometrika* was the leading journal for statistical work in Britain, but it was not popular among research workers engaged in agricultural science and thus Student's distribution did not enjoy an immediate success in the analysis of agricultural experiments and was mainly employed by Gosset and his co-workers at Guinness.[111]

Nonetheless, Gosset was influential in the application of statistics to the planning and analysis of agricultural experiments in Britain during the first three decades of the twentieth century. Due to Guinness's experimental work on barley breeding both in Ireland and Britain, the chemist had a chance to gain a first-hand experience of agricultural experiments.[112] The barley trials of Guinness put also Gosset in contact with the maltster Edwin S. Beaven, who introduced him to the two main centres for agricultural science in the early 1910s, Rothamsted Experimental Station

[107] For a scientific biography of Gosset see E. S. Pearson (1939, 1990). Gosset's obituary written by Ronald Fisher is R. A. Fisher (1939a). A collection of Gosset's scientific papers is E. S. Pearson and J. Wishart (1942).
[108] E. S. Pearson (1990), pp. 16-17.
[109] For Karl Pearson's involvement in statistics and his biometric school see T. M. Porter (1986), pp. 296-314 and T. M. Porter (2006).
[110] Student (1908).
[111] W. G. Cochran (1976), p. 13.
[112] L. McMullen (1939), p. 207.

and Cambridge University. Since 1910 Gosset was a correspondent of the Rothamsted director, Hall, in the use of statistical methods for the examination of field trials.[113]

In 1918 Gosset mentioned to Ronald Fisher the vacancy of a statistician at Rothamsted and during the 1920s he was the principal discussant of Fisher's statistical ideas on the planning and analysis of field experiments.[114] Over the years the chemist contributed to the popularisation in the press of the mathematical analysis of agricultural experiments. He also revised two seminal contributions on the planning and statistical analysis of the results of agricultural experiments published respectively in 1910 and 1911 in the *Journal of Agricultural Science*.[115]

The first contribution was co-authored by Thomas B. Wood, Cambridge professor of agriculture, and by his friend and colleague, the astronomer Frederick J. M. Stratton.[116] The two authors resorted in their discussion on the accuracy of agricultural experiments to the mathematical techniques adopted by astronomers. They explained in which cases it was possible to consider the average of a set of experimental data and how the probable error could be calculated using least squares.[117] If the mathematical techniques employed were markedly astronomical – and Gosset's suggestions "were quite rightly turned down as being too refined for the purpose" – the examples examined were standards of agricultural research, from livestock experiments to field trials.[118]

The second paper to which Gosset contributed was co-authored in 1911 by the already mentioned Rothamsted director, Hall, and by the agriculturist W. B. Mercer. The data that prompted the publication were gathered from two uniformity trials on wheat and mangolds conducted at Rothamsted Experimental Station in 1910.[119] Hall and Mercer made suggestions for the planning of such experiments connecting the set up of the trials to the final accuracy of the

---

113 Gosset's acquaintance with Beaven is mentioned in E. S. Pearson (1939), p. 230. For Gosset's scientific collaboration with A. D. Hall see letter from W. S. Gosset to A. D. Hall, 8th December 1910, RR Library and Archive, STATS 12.

114 For W. S. Gosset and R. A. Fisher scientific correspondence see W. S. Gosset (1962). On Fisher's appointment at Rothamsted see in particular W. S. Gosset (1962), Letter No. 3, 30th December 1918. Fisher's wife, Ruth E. Grattan-Guinness, belonged to a collateral branch of the Guinness family who employed Gosset (J. Moore (2006), p. 124), but the correspondence between Fisher and Gosset does not suggest any previous acquaintance of the chemist with Fisher's wife.

115 In "Mathematics and Agronomy" (E. S. Pearson and J. Wishart (1942), pp. 121-134), originally published in the *Journal of the American Society of Agronomy* in 1926, Gosset presented an overview of the basic concepts of statistics useful for agricultural experiments. The papers published by the *Journal of Agricultural Science* are, respectively, T. B. Wood and F. J. M. Stratton (1910) and W. B. Mercer and A. D. Hall (1911).

116 For a scientific biography of Wood and Stratton see their obituary notices as fellows of the Royal Society, F. G. H. (1931) and J. Chadwick (1961). For the agriculturist Wood statistics was not a whim. In the 1910s he published other contributions on the *Journal of Agricultural Science* in which he made use of statistical methods (T. B. Wood (1910), T. B. Wood and G. Udny Yule (1914)).

117 On the relation between the probable error and the standard deviation see N. S. Hall (2002), pp. 41-44.

118 E. S. Pearson (1939), p. 230.

119 Uniformity trials are field experiments conducted on a large number of small plots, all treated similarly. These experiments "provided data from which the standard errors obtained from different sizes and shapes of plots, numbers of replications, and different experimental plans could be estimated." (W. G. Cochran (1976), p. 15)

results. The paper by Hall and Mercer was concluded by an appendix written by Gosset, whose "assistance and criticism" was acknowledged by the authors.[120] In the appendix Gosset gave reference to the Student's distribution, when the "samples are too small to give more than a rough indication of the S. D. [standard deviation]".[121]

Gosset's work and the papers mentioned above are usually considered the only forerunners of Fisher's work in the planning and analysis of field experiments. However, a more cautious attitude should be taken in outlining the introduction of statistics into agricultural experiments. In her discussion on experimental design in agriculture Nancy Hall offers a richer overview that includes, but it is not limited to these sources.[122] It should also be mentioned that in 1912 the Cambridge School of Agriculture had hired the statistician George Udny Yule, a former assistant of Karl Pearson at University College London, to advise the local agronomists and breeders. Yule's appointment marked the start of the formal teaching of statistics in Cambridge, as Yule was both statistician to the school and university lecturer.[123]

Moreover, as pointed out by Gerd Gigerenzer and colleagues, the British work in applied statistics has certainly been the most successful, but "[e]specially in the German-speaking countries, there existed in the nineteenth century a rich tradition of agricultural experimental research and a slowly emerging body of statistical techniques that, towards the close of the century, even began to obtain a probabilistic interpretation."[124] In the 1920s also Jerzy Neyman, another relevant voice in twentieth century statistics, began his career, like Fisher, with the analysis of agricultural experiments in Poland at the Agricultural Research Institute in Bydgoszcz, where he worked from 1921 to 1923. Neyman, like Fisher at Rothamsted, was the founder of the statistics unit in the institute.[125]

There is reason to believe, thus, that the interest for integrating statistics into agricultural experimentation at the beginning of the twentieth century was a general trend and not a British specificity and that this trend began during the late nineteenth century, reached some concrete results in the first years of the twentieth century and culminated in the direct appointment of statisticians, as Yule, Fisher and Neyman, into agricultural research institutes during the 1910s and 1920s. What caused such trend is more difficult to account, as local and global factors certainly interacted and thus the most I can attempt is to suggest a few speculations.

---

120 "Note on a method of arranging plots as to utilise a given area of land to the best advantage in testing two varieties" by Student. In: W. B. Mercer and A. D. Hall (1911), pp. 128-132. Quotation in the text p. 127.

121 W. B. Mercer and A. D. Hall (1911), pp. 131-132.

122 N. S. Hall (2002), pp. 34-49.

123 F. Yates (1952), pp. 310-311.

124 G. Gigerenzer et al. (1989), p. 85.

125 C. Reid (1998), pp. 43-44.

In relation to British agricultural science Paul Brassley remarks that one of the main achievements of the Development Commission was to bring "the scientifically literate into the bureaucracy".[126] The impact of scientifically trained people in agricultural science in Britain increased with the creation of the Agricultural Research Council. The 'scientifically literate' promoted a more quantitative approach to experimentation and in this context statistics began to be employed in research institutions concerned with agricultural science, such as Rothamsted. Paradigmatic is the case of Hall, who, as a scientist, supported the use of statistical methods in the analysis of field experiments and, as a civil servant, became a member of the Development Commission and later of the Agricultural Research Council and advisor of the Ministry of Agriculture.

On a global perspective, it is worth mentioning that since 1905 agriculture had an international discussion forum, the International Institute of Agriculture, besides the national societies devoted to the subject. The institute, based in Rome, gathered since its foundation representatives from forty nations with the aim to prepare statistics of production and commerce of agricultural commodities, make enquiries into the practical organisation of farming in each nation, but also touched upon typical problems of agricultural science such as plant pathology and fertilizers efficacy.[127] In the original charter of the institute the same office was in charge of general statistics and practical information concerning farming and plant pathology.[128] Thus, there was a long-standing association between agriculture and statistics – although in the actuarial meaning and not directly concerned with field experiments – and this association might have prompted the following adoption of statistical tools in agricultural science. Certainly it strengthened an idea of quantification in agriculture that was well suited for promoting an increased precision in the field experiments. The global scale of the International Institute of Agriculture might also suggest why the British case was certainly not isolated.

One last point I ought to touch upon is the impact of biometry in the development of statistical methods for agricultural science in Britain.[129] Plant and animal breeding represented a component of agricultural science and the statistical analysis of breeding experiments that began at the beginning of the twentieth century might have been a further element behind the adoption of statistical methods in agricultural experiments. However, I would be cautious to assess to what extent this element should be taken into account. Both the main schools of thought called into cause in Britain for the examination of breeding experiments – Mendelians and biometricians –

[126] P. Brassley (1995), p. 476.
[127] On the history of this body see IIA (1927).
[128] IIA (1908), art. 19.
[129] For biometry I understand here all the applications of statistical methods to biology and in particular genetics, not only Karl Pearson's school at University College.

recognised a role to Francis Galton, who had advocated the use of statistical methods in the study of heredity. However, Mendelians were rarely keen on mathematics and were "'satisfied' to judge the fit of actual to expected ratios 'by inspection merely'".[130] On the other hand the biometric school inaugurated by Karl Pearson at University College certainly did not lack mathematical sophistication, but as mentioned in relation to Gosset's work, was virtually uninterested in the problems of agricultural experimentation because concerned only with large samples of experimental data. It is thus difficult to establish at that stage a direct connection between the use of statistics in agricultural experiments and the application of mathematical tools in the analysis of biological phenomena.

I have presented my overview of statistics in agricultural science before Ronald Fisher's work at Rothamsted only as a review with speculation. Nevertheless, it suggests that in the outline of field experiments and in the analysis of their results statistics had a long tradition, predating Fisher, formed from empirical solutions and attempts to theorize best practices in making experiments and minimize their errors. Such tradition had grown throughout the nineteenth century and at the beginning of the twentieth century statistics was certainly on the agenda of agricultural science.

1.3.2 *Ronald Fisher's appointment at Rothamsted*

The first statistician hired by Rothamsted Experimental Station, Ronald Fisher, was a mathematician trained in Cambridge.[131] Born in London in 1890, Fisher entered Gonville and Caius College in 1909 and there studied mathematics with Godfrey H. Hardy. After his graduation in 1912 Fisher spent a further year in Cambridge at the Cavendish Laboratory where he studied the theory of error with the astronomer Stratton – the same who had co-authored with Wood the paper on the interpretation of experimental results in agriculture – and statistical mechanics with the physicist James H. Jeans. It was Stratton who introduced Fisher to statistics and it is likely that Fisher attended as an undergraduate Stratton's course on 'Combination of observations'. Stratton encouraged Fisher's interest for statistical methods and since 1912 put him in contact with Gosset.[132]

It is not surprising that Fisher learnt the fundamentals of statistics from an astronomer, as the normal distribution, crucial for statistics, had largely been a legacy of astronomy.[133] Instead,

[130] On Galton and statistics see T. M. Porter (1986), pp. 297-298. For Mendelians and Biometricians and their controversy see W. B. Provine (2001). The quotation is taken from G. Gigerenzer et al. (1989), p. 150.
[131] A comprehensive biography of R. A. Fisher is J. Fisher Box (1978).
[132] For Fisher's mathematical education see A. W. F. Edwards (2005a), pp. 857-860.
[133] G. Gigerenzer et al. (1989), pp. 53-54.

statistics itself was not a popular subject in British universities during the 1910s and Yule was appointed to the lectureship in statistics into the School of Agriculture only in the year of Fisher's graduation. Apparently Fisher did not benefit from Yule's teaching, but Yule's textbook, *An Introduction to the Theory of Statistics*, along with the papers published by Karl Pearson on statistical theory "with their emphasis on the normal distribution, the method of least squares, correlation and (in Yule's case) contingency, may be assumed fairly to reflect Fisher's undergraduate knowledge".[134]

During his university years, Fisher cultivated also his interest for biology and he is numbered among the founding fathers of population genetics.[135] Since the 1930s genetics became for Fisher a second career with academic appointments in 1933 at University College London, as Galton professor of eugenics, and ten years later at Cambridge University, as Balfour professor of genetics. In 1911 Fisher began also his association with the British eugenics movement that sponsored his early research career in genetics through Major Leonard Darwin, president of the Eugenics Education Society of London.[136]

Despite the intense activity as an undergraduate, when he left Cambridge in 1913, not many career opportunities were available for Fisher. He worked briefly as a statistician in the City of London and since 1915 as a mathematical schoolmaster in several institutions. Eventually he arrived at Rothamsted with a temporary position in October 1919, called by the then Rothamsted director, John Russell.

No appointment correspondence between Fisher and Russell has survived to offer an insight into the employment of the statistician at the experimental station. However, Russell mentions the event in several accounts.[137] The Rothamsted director justifies Fisher's appointment with the necessity to extract more information from the results of the field experiments and from the records of the meteorological observations held at the station. In his autobiography Russell goes as far as to explain that the statistical methods adopted by the census authorities for examining their data represented the model he had in mind for the examination of the files collected since the mid nineteenth century at Rothamsted Experimental Station.[138]

---

[134] D. A. MacKenzie (1981), p. 186, A. W. F. Edwards (2005a), p. 860. Yule's textbook (G. Udny Yule (1922)) was originally published in 1911.

[135] For R. A. Fisher's contributions to population genetics see W. B. Provine (2001), pp. 140-154.

[136] On the British eugenic movement and the association between Fisher and the Eugenics Society see P. M. H. Mazumdar (2011).

[137] E. J. Russell (1935b), E. J. Russell (1956) and E. J. Russell (1966). All Russell's accounts are reconstructions *ex post* of Fisher's career at Rothamsted. The first one was written at Fisher's resignation in the 1930s, while Russell's autobiography and his history of British agriculture were written decades later, respectively in the 1950s and at the mid of the 1960s. Despite the consistent time gap between the first account and the other two, they are all suspiciously alike, as if the version of Fisher's appointment given in the 1930s had become canonical.

[138] E. J. Russell (1956), p. 131.

The decennial census of England and Wales that Russell took as a potential model dated back to the beginning of the nineteenth century, but since 1911 there had been a drastic increase in the production of census statistics, due to the adoption of punched-card equipment in the tabulation of the data.[139] It is this spread of statistical analysis that likely impressed the Rothamsted director. However, the development of statistics at Rothamsted could not and did not take the routes that the census model offered. The statistical methods that Fisher developed for the analysis of agricultural data were different from the ones adopted by the census authorities, let alone the mechanisation of the statistical analysis that started at Rothamsted only in the late 1930s.[140]

Moreover, by 1919 the director of Rothamsted Experimental Station must have been aware that a statistician, besides dealing with the past data, could offer an immediate contribution to the experimental research of the station. Yule at the Cambridge School of Agriculture was certainly an authoritative example in this sense and already in 1923 Gosset mentioned a possible involvement of Fisher in the planning of the field trials done at Rothamsted.[141] It would be difficult, otherwise, to explain why Fisher, appointed only in the fall of 1919 to study the past records of the station, had already completed in the spring of 1923 the examination of a current Rothamsted experiment.[142]

[139] For an history of the census of England and Wales in the nineteenth and early twentieth century see E. Higgs (1996) and E. Higgs (2004), chapters 4-5. A detailed account of the information technologies adopted by the census is given in M. Campbell-Kelly (1996).

[140] As clarified below the statistical analysis of agricultural experiments required methods for dealing with small experimental samples, while the census statistics is, by definition, a statistics of large numbers. On the adoption of punched-card equipment at Rothamsted Experimental Station see the letter from L. J. Comrie to F. Yates, 18th October 1938. Comrie writes to Yates offering to train a Rothamsted operator in the use of punched-card equipment (D.A. Boyd Papers, RR Library and Archive, STATS 8).

[141] W. S. Gosset (1962), Letter No. 29, 25th July 1923.

[142] The *Journal of Agricultural Science* received the manuscript of R. A. Fisher and W. A. Mackenzie on the analysis of the Rothamsted potato experiment in March 1923 (R. A. Fisher and W. A. Mackenzie (1923)).

Fig. 1.1 R. A. Fisher (on the right) having tea at Rothamsted. The photograph offers an unconventional portrait of Fisher. The image is undated, but it can be reasonably assigned to the early 1920s. In this informal occasion the statistician is not surrounded by the tools of the trade, but is concentrated on his teacup. The man sat next to him is likely to be a Rothamsted research worker, as the picture belongs to a series of photographs taken during a tea party in the station fields (Rothamsted Research Library and Archive, Ref. PHO 2.2.6).

Credits: Copyright Rothamsted Research Ltd.

1.3.3 *The statistics of small samples*

At the beginning of every modern textbook of statistics the reader finds the distinction between the parameters (mean, standard deviation etc.) that qualify the infinite population on which inferences must be drawn and the similar quantities that characterize the sample, that is the limited subset of the population upon which there is first-hand knowledge. This distinction, taken for granted in modern statistics, was established only in the early years of the twentieth century by the brewer William Gosset. Gosset was concerned with the application of statistical methods to experimental contexts, like agriculture, in which often only a limited amount of experimental data, that is a small sample, was available. In this case the difference between sample and population could not be held irrelevant, as explained by Gosset while presenting Student's distribution:

> any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong. […] If the number of experiments be very large, we may have precise information as to the value of the mean, but if our sample be small, we have two sources of uncertainty: (1) owing to the 'error of random sampling' the mean of our series of experiments deviates more or less widely from the mean of the population, and (2) the sample is not sufficiently large to determine what is the law of distribution of individuals.[143]

The biometric school of Karl Pearson at University College London offered methods reliable only with large amounts of experimental data because "biometric statistical theory typically relied […] on the assumption that sample statistics could safely be substituted for population parameters".[144] The same could not be said of the experimental work done by Gosset for the Guinness brewery and the chemist was the first to introduce "different letters to denote sample (s) and population (σ) standard deviations, and sample (r) and population (R) correlations".[145] Student's distribution and its related tables of the probability integral gave the opportunity to deal with a small sample of experimental data.

The scientific friendship of Gosset and Fisher originated from their common interest for the statistical treatment of small samples. Using n-dimensional geometry, Fisher had derived Student's distribution and, advised by his Cambridge tutor Stratton, sent the mathematical proof

[143] Student (1908), p. 1.
[144] D. A. Mackenzie (1981), p. 114.
[145] D. A. Mackenzie (1981), p. 116.

to the chemist since 1912. This correspondence "petered out owing to my [Gosset] lack both of courtesy and mathematics" in a few months, but it was renewed in 1915 and went on until the death of Gosset in 1937.[146] *Statistical Methods for Research Workers*, Fisher's textbook on analysis of variance and (briefly) experimental design, aimed precisely at "tackling small samples problems on their merits" in order "to apply accurate tests to practical data".[147]

Analysis of variance and experimental design were developed for the analysis of the small samples used in agricultural and biological research, the two disciplines closer to Fisher at Rothamsted. However, as pointed out by Frank Yates in a eulogy of *Statistical Methods for Research Workers*, the application of Fisher's methods was not strictly limited to small samples, because "even when the data consist of observations on a large number of separate individuals these often require to be grouped according to a relatively small number of classes. Tests of significance involving these classes frequently involve small sample theory".[148] That accounts for the increasing popularity of Fisher's statistical methods that, as I will discuss in the final section of the chapter, despite being born in a very specific setting became of general use in a wide array of laboratory and field disciplines.

1.4 ***Ronald Fisher's statistical methods for field trials***

1.4.1 ***Analysis of variance***

The rationale behind the statistical method known as analysis of variance is to split the global variation of a phenomenon, i.e. the variance, in additive components, each one linked to an independent cause of variability.[149] In an experiment on the efficacy of fertilizers the analysis of variance allows to examine the variation of the yield both within plots that receive the same fertilizer or combination of fertilizers and between sets of plots that receive a different treatment. The global variation of the yield in the plots is subdivided in several components and it is possible to measure the effects of distinct causes – soil, fertilizers, pests etc. – on the final result. In so doing factors, such as the unequal fertility of the soil, can be set aside from the efficacy of fertilizers, that is the real point of interest for agricultural science.

---

146 W. S. Gosset (1962), No. 1, 15th September 1915.

147 R. A. Fisher (1925a), pp. *vii-viii*. In the preface to the first edition of *Statistical Methods for Research Workers* Fisher acknowledged the contribution to the book of Gosset and his assistant at Guinness, Edward Somerfield, in the proof-reading of the volume.

148 F. Yates (1951), pp. 30-31.

149 The variance is the square of the standard deviation. The latter is the quantity that measures the dispersion of the individual values in a statistic distribution. In Fisher's idea of statistics variation is a key concept and "the study of the causes of variation of any variable phenomenon, from the yield of wheat to the intellectual man, should be begun by the examination and measurement of the variation which presents itself." (R. A. Fisher (1946), p. 3)

Ronald Fisher used to describe the analysis of variance as "a simple method of arranging arithmetical facts so as to isolate and display the essential features of a body of data with the utmost simplicity".[150] However, Fisher's method was rather more than arithmetic for its potential applications to experimental research. It offered an alternative approach to the method of correlation that had dominated the statistics of Karl Pearson's biometric school and it made available to the research workers interested in the examination of their data tests of significance that could be easily applied and that were more flexible than the test derived from the Student's distribution, because the analysis of variance confronted not the mean of the sample, but its variance.[151]

Fisher developed the analysis of variance in the early 1920s, while he was re-examining the figures of the long-term experiments and acting as consultant of the Rothamsted research workers. However, the word variance and the idea to split the global variation in additive components predated his appointment at Rothamsted. Fisher used the term variance for the first time in 1918, in the seminal paper in which he proved that Mendelism and biometry were compatible. In the paper Fisher adopted the standard deviation as a measure of the variation of a character in a human population, but, instead of dealing with the standard deviation directly, he introduced the variance, i.e. the square of the standard deviation, that made possible to split the global variation in additive components, each one linked to an independent cause of variability.[152]

Donald Mackenzie and Theodore Porter have claimed that the analysis of variance was fully developed in Fisher's 1918 paper and that it derived from Fisher's involvement in eugenics.[153] However, I find in primary and secondary sources more convincing evidences of the development of analysis of variance during Fisher's work on agricultural experiments at Rothamsted. According to the Rothamsted report for the years 1925-26, "[t]he first example of an analysis of variance in its modern form was the examination of the results of T. Eden's experiment in 1922 on the response of different potato varieties to manures" (Fig. 1.2a and Fig.

---

150 Letter from R. A. Fisher to G. W. Snedecor, 6th January 1934, G. W. Snedecor Papers, Special Collections Department, Iowa State University Library, RS 13/24/51, Box 1, Folder 9.

151 F. Yates (1951), p. 24; 32-33. J. Hagen (2003), p. 368.

152 R. A. Fisher (1918). For a detailed analysis of this paper see W. B. Provine (2001), pp. 143-147.

153 "an interesting example being the use of the technique of the analysis of variance, originally developed in eugenics research (Fisher 1918a), as the basis for the design and analysis of agricultural experiments (Fisher and Mackenzie 1923)" (D. A. Mackenzie (1981) p. 211). "Analysis of variance, Fisher's most important addition to the techniques of statistical analysis, and one now widely used in a variety of fields, was invented as a method for studying heredity – indeed, one might almost say, as a theory of heredity." (T. M. Porter (1986), p. 316). On the other hand Porter claims that Fisher's methods of experimental design were linked to his engagement with the Rothamsted research workers (T. M. Porter (1986), pp. 317-318).

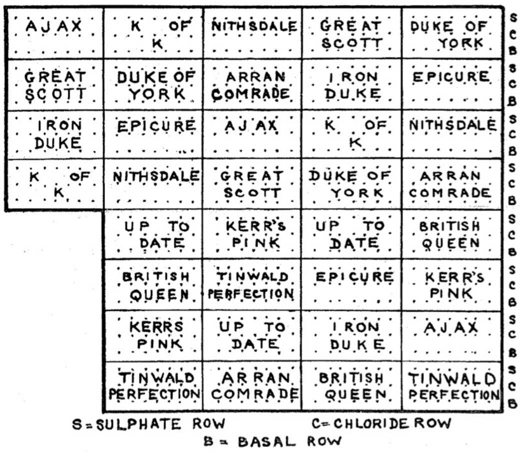
1.2b).[154] The historian Joel Hagen agrees in suggesting this paper as Fisher's first application of analysis of variance.[155]

Fisher and his first statistical assistant at Rothamsted, Winifred A. MacKenzie, examined the results of Eden's field trial interpreting the yield of each experimental plot as "the sum of two quantities, one depending on the variety and the other on the manure."[156] The statistical examination was presented in "a single compact table", an element that would become a hallmark of the analysis of variance, as the table could show "both the structure of the experiment and the relevant results, in such a way as to facilitate the necessary tests of their significance".[157]

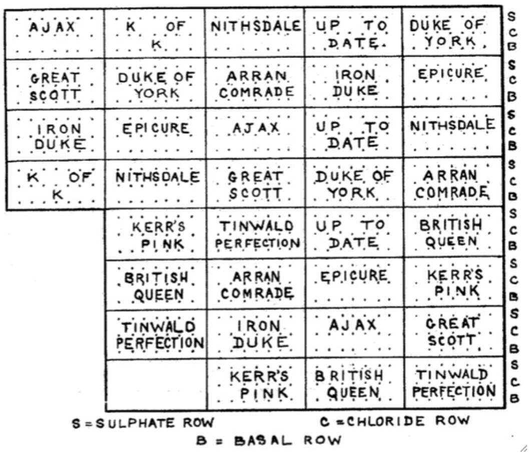
The analysis of the potato experiment was the second paper of the series "Studies in crop variation", published in the *Journal of Agricultural Science* from 1921 to 1930 and authored by Fisher and/or his co-workers in the Rothamsted statistics department. The data that prompted these studies were taken from the Rothamsted experiments, both the annual trials and the long-term experiments, and the statistical method of attack in the data analysis was the analysis of variance. John Aldrich mentions the first paper of the series on crop variation, concerned with the examination of the Broadbalk wheat experiment, as the first publication in which the technique known as analysis of variance was explicitly addressed.[158] In this paper, however, there is no table, the customary form in which Fisher arranged the analysis of variance. The first of these tables is the one published in the analysis of the potato experiment mentioned above.

The choice to present Fisher's analysis of variance in papers published in the *Journal of Agricultural Science*, a venue fit for agriculturists rather than statisticians, and the constant application of the method to the analysis of the Rothamsted data support further my claim that the agricultural research pursued in the experimental station prompted the development of Fisher's statistical methods.

154 RES (1927), p. 28. The paper mentioned is R. A. Fisher and W. A. Mackenzie (1923).
155 J. Hagen (2003), p. 368.
156 R. A. Fisher and W. A. Mackenzie (1923), p. 315.
157 R. A. Fisher (1947), p. 50.
158 R. A. Fisher (1921); J. Aldrich (2007).



2a. Farmyard Manure



2b. No Farmyard Manure

Fig. 1.2a, 1.2b Thomas Eden's potato experiment. The field was split in two sections, one with and the other without farmyard manure. Each section was then further divided into thirty-six small plots, where twelve potato varieties were planted each one three times in a chessboard arrangement (in the series without farmyard manure the variety K. of K. was only duplicated). In each plot there were three rows of seven plants each, and each row received a different combination of fertilizers as indicated by the letters S, C, B.

Credits: *Journal of Agricultural Science*, Vol. 13, pp. 312-313. Reproduced with permission of Cambridge University Press.

1.4.2 *Experimental design*

In the statistical analysis of the potato experiment set up by Thomas Eden, Fisher and Mackenzie argued that there was no significant difference in the response of the several potato varieties to the manurial treatment, but their conclusion was weakened by the mistakes in the arrangement of the trial (Fig. 1.2a and Fig. 1.2b).[159] Just after the publication of the paper, Gosset – by then a regular scientific correspondent of Fisher – pointed out that the experiment was badly planned and that the combination of fertilizers tested could not give a clear-cut answer on the response of the potato varieties. The brewer reminded Fisher that "[t]he experiment can't bias its errors so as to increase the difference up to significance for you!".[160]

The faults of the potato experiment proved that an efficient statistical technique for the analysis of experimental results was not enough to guarantee the quality of the information extracted from the field data. Planning was as vital a duty and the statistician should be consulted also in this initial stage. Money savings in the execution of the experiment and safeguard of the

[159] "[Eden's experiment] intended to investigate in particular the interaction between two variables: different types of potash manure and different varieties of potatoes. In that case although three variables were introduced, they were tested on areas of different size: basal manures were compared by half-fields, potato varieties on plots within the fields, and potash manures in rows within the plots. In consequence, the comparisons were not all equally precise, and three estimates of error would have been required to test the three effects. Moreover, the progressive fragmentation of the field required in order to introduce three factors in this way resulted in rows consisting of only seven plants each, rather too small a sample to be reliable." (J. Fisher Box (1978), p. 162) For the analysis of the potato experiment see J. Fisher Box (1978), pp. 109-112. The experiment was re-examined by Fisher in *Statistical Methods for Research Workers* (R. A. Fisher (1946), pp. 236-241).

[160] W. S. Gosset (1962), Letter No. 29, 25th July 1923.

results reliability – i.e. the soundness of the statistical method of analysis – were at stake because "the estimate of error is not created by the statistician out of nothing, but is inferred from the observations by a process of estimation analogous to that used in the estimation of any other quantity, and requiring the same care in experimental design if the estimate is to be a valid one".[161]

Ronald Fisher singled out four basic principles in arranging field trials at Rothamsted: 1) the replication of experiments on small plots of land; 2) randomisation, i.e. the chance allocation of treatments to plots; 3) the use of factorial experiments in which several questions are combined together; 4) 'confounding' that is the decision in relevant cases to sacrifice information on minor interactions.[162]

Randomisation is nowadays recognised as Fisher's breakthrough in experimental design in and beyond agriculture.[163] Before Fisher, the allocation of treatments to plots in field experiments was done according to a systematic arrangement at risk of introducing a bias, if in the field there was a gradient of fertility or an uneven distribution of sunlight. Randomisation, instead, was a way to limit the variability of the soil or, in an experiment where multiple factors were tested, a tool to prevent that the mutual influences of these factors were mistaken for the treatments efficacy. Moreover, it represented the necessary condition for examining the experimental results with the analysis of variance.

In agricultural experiments randomisation was achieved through two basic schemes, the Latin square and the randomised blocks (Fig. 1.3a, Fig. 1.3b). In the former the plots were arranged with as many rows and columns as the number of treatments to be tested, while in the latter the experimental area was divided into strips or blocks, each one containing one plot of each treatment. In both cases treatments were assigned to the plots in the scheme at random. According to Fisher, randomisation did "not mean that the experimenter writes down the names of the varieties, or letters standing for them, in any order that may occur to him, but that he carries out a physical experimental process of randomisation, using means which shall ensure that each variety has an equal chance of being tested on any particular plot of ground".[164] Such actions can be shuffling and drawing cards from a deck or going through a table of random numbers.

Nancy Hall has pointed out that "Fisher imported randomness from sampling into experimental design".[165] In the same way as randomisation in experimental design required a

161 R. A. Fisher (1934), p. 47.
162 R. A. Fisher (1934), pp. 46-49.
163 For randomisation in experimental design in relation to R. A. Fisher see N. S. Hall (2002) and N. S. Hall (2007).
164 R. A. Fisher (1947), pp. 49-50.
165 N. S. Hall (2007), p. 296.

physical action, also in sampling randomisation was literally a physical process, as described by the Rothamsted research workers, Thomas Eden and Ernest Maskell. During the spring of 1925, the estimate of the plants in a uniformity trial was obtained with a rather peculiar procedure for randomisation: "1 9ft. rod was thrown haphazardly on to the plots and aligned with that drill row which was nearest to a definite end of the rod; the number of plants was then counted on four such randomly distributed lengths".[166] The use of a physical action in randomisation, either in sampling or in the arrangement of an experiment, prevented the research worker from introducing a subjective bias in the scheme.

Randomisation is now widely applied for planning experiments in a wide range of disciplines, but it was greeted with scepticism in the 1920s. Even Gosset was highly critical about Fisher's suggestion: he could see the usefulness of randomisation from the theoretical point of view, but he considered impossible to apply randomisation in practice unless one would "want a large lunatic asylum for the operators who are apt to make mistakes enough even at present".[167] In fact, it was not the statistician, but the 'operators' – at Rothamsted the farm manager, the superintendent of the field experiments and the labourers that they directed – who set up in practice the trials and worked on them following the crops from sowing to harvest.

Despite Gosset's fears, by the early 1930s randomisation was widely mentioned in agricultural science, if not yet agreeable to everyone in the field.[168] During the summer lectures given by Fisher at Iowa State University in 1931, for instance, "a good share of the audience were biologists and, indeed, the agronomists" and they "were very eager to discuss randomisation".[169]

Rows	Columns I	II	III	IV
I	*12N*	*9N*	*6N*	*2N*
II	*9N*	*12N*	*2N*	*6N*
III	*6N*	*2N*	*12N*	*9N*
IV	*2N*	*6N*	*9N*	*12N*

3a. Latin square

I	III	II	IV
A	C	F	B
E	D	C	A
D	B	E	G
C	A	B	D
F	E	A	C
B	G	D	E

3b. Randomised blocks

Fig. 1.3a Nitrogenous top dressing on roots. Sugar beet experiment, Rothamsted 1926. System of replication: Latin square 4×4, plots 1/145 acre.

Credits: *Rothamsted Experimental Station Report 1925-1926* (1927), p. 142. Copyright Rothamsted Research Ltd.

Fig. 1.3b Top dressing on cereals. Oats (Grey Winter) experiment, Rothamsted 1925. System of replication: randomised blocks with additional plots F or G, plots 1/40 acre.

Credits: *Rothamsted Experimental Station Report 1925-1926* (1927), p. 145. Copyright Rothamsted Research Ltd.

166 T. Eden and E. J. Maskell (1928), p. 170.
167 W. S. Gosset (1962), Letter No. 50, 20th October 1924.
168 N. S. Hall (2007), p. 314.
169 Letter from R. A. Fisher to T. Eden, 1st December 1931, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

1.4.3 *Ronald Fisher's co-workers, Thomas Eden and Ernest Maskell*

A comprehensive account of the development of analysis of variance and experimental design at Rothamsted Experimental Station cannot disregard the scientific staff that closely assisted Ronald Fisher in the fieldwork necessary to implement his statistical methods. The annual reports of the station mention, in fact, that analysis of variance and experimental design "were the outcome of long previous investigations in which several workers, including the agriculturist, the ecologist, the plant physiologist and the statistician took part".[170]

The mathematical tools devised by Ronald Fisher required "a correspondingly rigorous field technique" and the workers of the field experiments department collaborated with him in the development of adequate practices. At Rothamsted plot techniques and methods were optimized "to give the maximum of accuracy with the minimum of labour".[171] A field laboratory, available to ecologists and plant physiologists, was established to measure and observe the growth of the crops in order to acquire information of statistical relevance.[172] The workers of the field experiments department contributed also to test the reliability of Latin squares and randomised blocks with uniformity trials and they took part in the development of sampling techniques for agricultural experiments.[173]

The ecologist Thomas Eden and the plant physiologist Ernest Maskell were the field workers who closely collaborated with Fisher in the practical implementation of analysis of variance and experimental design during the mid-1920s. Eden, trained in chemistry at the Victoria University of Manchester, started his work at Rothamsted as exhibitioner in 1921. From 1923 to 1927 he worked at the experimental station as ecologist in charge of the field experiments and during these years he collaborated with Fisher in the development of analysis of variance and experimental design. "Since then – Eden added on his résumé – statistical control has been a feature of all my work on soil and cropping problems".[174] Eden was the first secretary of the Field Plots Committee, since 1924 the body in charge of the field trials conducted at the station. Left Rothamsted in 1927, Eden moved to Ceylon and then East Africa as agricultural chemist, in particular interested in the cultivation of tea, but in 1932, during a leave from Ceylon, he came

---

170 RES (1927), p. 27.

171 Ministry of Agriculture and Fisheries, Agricultural Research Council, annual report 1924-1925. TNA, DSIR/36/4239 (Quotation reproduced with permission from the BBSRC).

172 RES (1927), p. 27; T. Eden and E. J. Maskell are explicitly mentioned as the people in charge of the field observations, "[t]hese results to be reported with a view to statistical investigation and to modification for rapid estimation on a routine scale." (Minutes of the Field Plots Committee, 31st January 1924, RR Library and Archives, FX1.1.1)

173 RES (1927), p. 28.

174 Thomas Eden's résumé, ca. 1946, E. J. Russell Correspondence, MERL, University of Reading, FR HERT 11/1/1. All the biographical information on Eden is gathered from the résumé.

back to Rothamsted as a worker in Fisher's department "picking up new ideas and clarifying old ones".[175]

Maskell instead was trained as a botanist in Cambridge and joined the staff of Rothamsted Experimental Station as plant physiologist in 1924. In the short period that he spent at Rothamsted Maskell "rapidly acquired a sound basis of statistical knowledge" and statistics contributed to his further research activity. For instance, Maskell used his knowledge of experimental design and adopted Student's distribution for the data analysis in the investigations that he undertook at the Cotton Research Station in Trinidad.[176]

Both Eden and Maskell became early promoters of Fisher's statistical methods, not only as users of analysis of variance and experimental design, but also for their contribution in the dissemination of these statistical methods in agricultural science.[177] Moreover, the experimental co-workers of Fisher did not limit themselves to the field practice avoiding any direct involvement with statistics. In his book *The Design of Experiments* Fisher gives credit to Maskell for the idea of fiducial probability.[178] The correspondence between Thomas Eden and Ronald Fisher shows as well a constant interest of Eden in refining his statistical knowledge to the extent that Eden became a fellow of the Royal Statistical Society.[179] Eden acted also as statistical consultant of the planters' Association of Ceylon and he was an advisor in statistical matters of the Ceylon Government during WWII, when he was "more concerned with the cost of living index than with the analysis of variance".[180]

1.5 *The reshaping of research at Rothamsted Experimental Station*

1.5.1 *The Rothamsted statistics department, 1919-1933*

Ronald Fisher's appointment at Rothamsted became permanent in 1920 and during the early 1920s Fisher set up from scratch the Rothamsted statistics department. The department was both a statistical and a computing unit, because the application of analysis of variance and experimental design to field trials could not be done 'on thumbs' but required calculating

[175] Letter from T. Eden to R. A. Fisher, 11th August 1930, R. A. Fisher Papers, BSL, The University of Adelaide.
[176] G. E. Briggs et al. (1961), p. 162.
[177] T. Eden (1935) and E. J. Maskell (1929).
[178] R. A. Fisher (1947), p. 187. Fisher introduced the concept of fiducial probability in 1930 to make probability statements on an unknown parameter without using the theorem of inverse probability. Fisher's fiducial probability is examined in A. W. F. Edwards (1976).
[179] See for example the letter in which Eden discusses with Fisher a possible problem to examine during his leave at Rothamsted in 1932. Eden is interested in the use of the analysis of variance for dealing with a set of widely different treatments giving markedly different responses (Letter from T. Eden to R. A. Fisher, 1st April 1931, R. A. Fisher Papers, BSL, The University of Adelaide).
[180] Letter from T. Eden to R. A. Fisher, 25th August 1942, R. A. Fisher Papers, BSL, The University of Adelaide.

machines and computing aids, the making of new statistical tables and the hiring of human computers.[181]

Throughout Fisher's time at Rothamsted the permanent staff of the statistics department was always scarce: Fisher could count on one or at most two assistant statisticians and a handful of human computers. On the other hand, over fifty temporary workers came to the department to learn analysis of variance and experimental design contributing both to the examination of the Rothamsted data and to the dissemination of Fisher's statistical methods. Appendix 1.1 gives an overview of all the people – statisticians, computers and visitors – who worked in the Rothamsted statistics department during Fisher's time.

a) ***Statisticians***

Fisher's first statistical assistant, Winifred A. Mackenzie, was appointed in June 1920.[182] Despite their different background – Fisher was a mathematician, while Mackenzie had a bachelor degree in economics – they set up a fruitful collaboration in the analysis of the station records and of the current experiments co-authoring several papers.[183] In 1924 Mackenzie was awarded an MSc from the University of London for her work at Rothamsted. In her master's dissertation, Mackenzie applied the analysis of variance to the barley data collected at the station and the University of London appointed as her examiners Ronald Fisher and William Gosset. [184] She left the department in 1927 after her marriage.[185]

With the departure of Mackenzie two new assistant statisticians were appointed, John Wishart (1927) and Joseph O. Irwin (1928).[186] Fisher's new co-workers had already received training in biometry and mathematical statistics under Karl Pearson at University College London, before joining Rothamsted.

In Fisher's department Wishart developed the more mathematical and computational aspects of statistics. He was also engaged in the design of field experiments, the study of meteorological factors in agriculture, the experiments with fertilizers, the analysis of agricultural experiments

[181] "You can't do ANOVA [i.e. analysis of variance] on thumbs", is a quotation taken from the amusing poem *Ronald and his line* written by the former Rothamsted statistician Gavin Ross. (Gavin Ross, undated)
[182] The first payment (£5.0.0 per week) to Winifred Mackenzie is registered on 30th June 1920 (Rothamsted Laboratory Cash Account October 1919-January 1921, RR Library and Archive, LAT 34).
[183] R. A. Fisher and W. A. Mackenzie (1922); R. A. Fisher et al. (1922); R. A. Fisher and W. A. Mackenzie (1923).
[184] See W. S. Gosset (1962), Letters: No. 40, 5th February 1924; Letter No. 42, 19th February 1924; Letter No. 43, 26th February 1924; Letter No. 45, 20th May 1924.
[185] Mackenzie married Rev. R. T. Tyrrell, Leopoldville Africa. She returned to England in November 1928 with her two children. (RES (1929b), p. 26)
[186] J. Fisher Box (1978), p. 139.

when results were incomplete. Wishart left Rothamsted in October 1931 to take up the appointment of reader in Statistics at the School of Agriculture in Cambridge, succeeding Yule.[187]

Joseph Irwin joined the Rothamsted statistics department a few months after Wishart to take part in the Ministry of Agriculture Crop-Weather scheme, an initiative to collect uniform observations on the state of crops, the incidence of insects and fungi and the weather conditions at several agricultural stations scattered throughout Britain.[188] In 1930 he moved to the London School of Hygiene and Tropical Medicine where he joined the staff of the Medical Research Council and started a life-long career in biostatistics.[189]

With the resignations of Wishart and Irwin, two new assistant statisticians entered in Fisher's department: in 1929 A. Margaret Webster and in 1931 Frank Yates. Webster was appointed with the support of the Royal Agricultural Society to examine with Fisher's statistical methods the data collected at the experimental station of Woburn, a research institution under Rothamsted's patronage.[190] There was a cold relationship between Webster and Fisher and Fisher left to his other statistical assistants the task to advise her.[191] She left her job in 1933 after marrying E. Walter Russell, the son of the Rothamsted director, John Russell.[192]

Frank Yates graduated in mathematics at Cambridge with first class honours in 1924 and, after an unsatisfactory experience as schoolteacher, joined the Gold Coast (Ghana) Geodetic Survey in 1927.[193] Yates' appointment in Fisher's department started in August 1931 with a salary of £360.[194] At Rothamsted he spent the first months learning the statistics he did not know, "reading up – following Fisher's advice – what one may call the biometrical side of statistics, and in particular the work on tests of significance and the analysis of variance, as this is the method which everyone who comes here wants to learn and on which the sooner you are an authority the better".[195]

Yates learnt quickly Fisher's statistical methods and established his scientific career contributing to experimental design and sampling techniques for censuses and surveys. Unlike his predecessors he remained at Rothamsted for all his career succeeding Fisher in 1933 and leading

[187] RES (1932), p. 10. A scientific biography of J. Wishart is E. S. Pearson (1957).
[188] For Irwin's work at Rothamsted see the RES (1929a), p. 39. A scientific biography of J. O. Irwin is B. G. Greenberg (1983), p. 527.
[189] RES (1931c), p. 46.
[190] E. J. Russell and J. A. Voelcker (1936), p. *xii*.
[191] F. Yates (1981), p. 138.
[192] RES (1935b), p. 62.
[193] A scientific biography of F. Yates and a comprehensive bibliography of his publications is D. J. Finney (1995). For Yates' appointment at Rothamsted see F. Yates (1981), pp. 135-137, and the appointment correspondence held at RR Library and Archives, STF 31 (old catalogue reference).
[194] Appointment of a member on the staff of the Rothamsted Experimental Station: Frank Yates, 14th September 1931, RR Library and Archive, STF 31 (old catalogue reference).
[195] Letter from R. A. Fisher to F. Yates, 9th May 1931, RR Library and Archive, STF 31 (old catalogue reference).

the Rothamsted statistics department as a pioneer institution in the computer age during the 1950s (**chapter 4**). With Fisher he kept a thirty-years friendship and scientific collaboration – they co-authored the *Statistical Tables for Biological, Agricultural and Medical Research* (**chapter 2**) – ended only by Fisher's death in 1962. Fisher, moreover, remained associated to Yates' department as honorary consultant throughout the 1930s and scientific visitors often moved from the Galton Laboratory to the statistics department at Rothamsted and vice versa.[196]

b) ***Human Computers***

As observed by David Grier, the stories of human computers "are often difficult to tell, as the vast majority of computers left no record of their lives beyond a single footnote to a scholarly article or an acknowledgment in the bottom margin of a mathematical table."[197] For the human computers who worked in Fisher's department sometimes not even the full name is known.[198] The first official mention of assistant staff in Fisher's department is represented by the station report for the years 1921-1922. At that stage the department had a laboratory assistant, A. D. Dunkley, and an honorary computer, W. D. Christmas. Both Dunkley and Christmas remained in the department for about ten years and in all the later reports Dunkley is associated to the more precise role of assistant computer, but probably doing calculations was never his main task.[199] Instead, he certainly cooperated with W. D. Christmas in the "compilation of the monthly weather records" at Rothamsted and prepared graphs and diagrams for the publications of the department.[200] Therefore, the qualification of computer in the Rothamsted statistics department should be understood in a very broad sense, encompassing all the activities that were of assistance to the statisticians' work and not just computing *per se*.

Under Fisher four other people, all females, joined at different times the department with the qualification of assistant computers: Kathleen Abbott, Florence Pennells, Alice Kingham, Kitty Rolt. In 1933, the year in which Fisher left the department, another assistant computer, J. M. West, was hired. It is not surprising that most of the computers in the statistics department were women. Since the mid of the nineteenth century women had entered into the computing rooms as "desk laborers, who were earning their way in this world with their skill at numbers",

---

[196] For instance Mildred M. Barnard worked both at the Galton Laboratory and the Rothamsted statistics department in 1935-1936.
[197] D. A. Grier (2007), p. 8.
[198] In the Rothamsted Experimental Station reports the male staff is mentioned only with the initials of the first name and the extended version of the surname, while for the female staff the full names and surnames are given. The convention helps in making at least a gender distinction among the computers.
[199] Dunkley's name is not associated to the request of a pay rise made by the other computer girls in 1931 despite being qualified as assistant computer.
[200] Anonymous (1931), p. 136.

subordinate to a professional staff, handling routine calculations and paid less than their male counterparts, thus preferred when the budget was tight.[201]

The Rothamsted computers of the 1920s and early 1930s fit at best this portrait. Certainly they did not receive any university education and no qualification is listed for them in the official reports of the station. Their role as computers should not have been seen as a real professional qualification as Alice Kingham moved from the assistant staff in the department of statistics right to the task of laboratory attendant in the department of chemistry.[202]

The only insight in the working life of the human computers in Fisher's department is offered by the unsuccessful request of wage increase filed in December 1931 by Florence Pennells and Kitty Rolt, after the departure of Alice Kingham. They appealed to Fisher for a better salary – they were earning respectively twenty-two shillings and six pence and twelve shillings per week – in relation to their heavier labour after Kingham's departure and complaining that their work as computers was a much harder chore than the one of the other assistants in the station.[203]

Fisher's answer to the letter suggests a sympathetic attitude towards his computing staff – "I shall certainly feel responsible for seeing that you are not driven to work harder than is good for your health, or indeed beyond what is necessary to attain full competence in computing practice" –, perhaps increased by the young age of the two assistants, indicated as "girls".[204] He thus committed himself to support their request if they could prove a "real increase" in their computing capacity.

Fisher invited the "girls" to show to Frank Yates, in charge of the computing staff, which machines they could use "skilfully and quickly for the different routine processes required" and, "what is very valuable when there is a shortage of machines", he invited the human computers to show the tasks for which they were "able to use the slide rule or logarithm tables".[205] The human computers in the Rothamsted department, thus, were invited not just to show dedication to the routine work, but also initiative in employing minimal resources for doing their job. Yates' report on the computing girls was positive and the request of pay rise forwarded by Fisher to the

---

201 D. A. Grier (2007), p. 81.

202 Letter from E. J. Russell to R. A. Fisher, 21st December 1931, RR Library and Archive, STF 31 (old catalogue reference).

203 Letter from F. Pennells and K. Rolt to R. A. Fisher, 3rd December 1931, R. A. Fisher Papers, BSL, The University of Adelaide. For Pennels and Rolt wages see Letter from F. Pennells and K. Rolt to R. A. Fisher, 12th December 1931, R. A. Fisher Papers, BSL, The University of Adelaide.

204 Letter from R. A. Fisher to F. Pennells and K. Rolt, 5th December 1931, RR Library and Archive, STF 31 (old catalogue reference). For the qualification of the computers as 'girls' see Letter from R. A. Fisher to F. Yates, 5th December 1931, RR Library and Archive, STF 31 (old catalogue reference).

205 Letter from R. A. Fisher to F. Pennells and K. Rolt, 5th December 1931, RR Library and Archive, STF 31 (old catalogue reference).

station's director, who eventually denied it, as "it is not possible for us to make exceptions [to the annual increments for the assistant staff], which would create dissatisfaction".[206]

As evident from the correspondence mentioned above, the human computers at Rothamsted were mainly engaged in routine computations. However, from time to time they were called to handle a bit of statistics on their own. For instance, in the absence of Fisher's assistant statisticians, Yates and his colleague, Webster, both ill at home, Pennels and Rolt had to examine a set of data related to a sheep experiment. The Rothamsted farm manager, Henry G. Miller, wanted the human computers "to work out some correlations" for him and Pennells and Rolt resorted to Yates for advice:

> There are two sets of figures one is the weights of the young ewes and the other is weights of older ewes do these have to be worked out separately? In some cases the ewe has died or been sold do we leave these out of account from the beginning? also for the twins do we add the two weights together?[207]

Not without some hesitation the human computers in Fisher's department turned themselves into statisticians.

c) ***Computing activity and equipment***

Since the mid-1920s Fisher's department was engaged in heavy computational labour for the analysis of the field trials, as the summary tables of the replicated experiments in the Rothamsted reports were now supplied with the standard error calculated using the analysis of variance.

Besides the work involved in the analysis of the local experiments and of the experiments conducted at outside centres under the supervision of Rothamsted, Fisher promoted in his department also the computation of mathematical and statistical tables and chose co-workers able to assist him in the task. The tables in the appendix of *Statistical Methods for Research Workers* were computed by Fisher and his assistant, Mackenzie, in the early 1920s, and represented – revised and extended – the first set of tables for the collection prepared by Fisher and Frank Yates during the 1930s.[208] During the years he spent at Rothamsted, Fisher was also actively engaged in the British Association Mathematical Tables Committee, in which were co-opted also his assistants, Wishart and Irwin. The Committee made available for the Rothamsted statisticians a Brunsviga calculator to be used for the table making activity of the British Association.[209]

---

206 Letter (copy) from E. J. Russell to R. A. Fisher, 21st December 1931, RR Library and Archive, STF 31 (old catalogue reference). Yates' report is the Letter from F. Yates to R. A. Fisher, 7th December 1931, R. A. Fisher Papers, BSL, The University of Adelaide.

207 Letter from F. Pennells and K. Rolt to F. Yates, 29th December 1931, RR Library and Archive, STF 31 (old catalogue reference).

208 R. A. Fisher (1925a), R. A. Fisher and F. Yates (1963, first edition 1938).

209 M. Croarken (2003), p. 244. The allocation of the machine to Rothamsted is explicitly mentioned in the Minutes

Three main computing tools supported the activity of Fisher's department during the 1920s and early 1930s: 1) mechanical and electro-mechanical calculators, 2) slide rules, 3) mathematical tables.[210] In the early decades of the twentieth century calculating machines were expensive and their cost was related to their ability in performing different operations. The main categories of computing machines were pinwheel and stepped drum calculators, so called in relation to their working mechanism. The basic operations were performed by the two kinds of machines almost in the same way: addition and subtraction directly, multiplication and division as repeated addition or subtraction. Besides these two basic categories there were a few machines, like the Millionaire of which Fisher was a keen user, able to perform products directly on the basis of a multiplication table, making computation quicker and less tiresome.[211]

Before coming to Rothamsted, Fisher asked advice to Gosset about the calculating machine that might be more suitable for a statistician, but the Rothamsted records do not give details on the machine that was rented for Fisher in December 1919.[212] In 1921 or 1922 Fisher acquired, instead, a Millionaire motor calculator, which he praised for the straightforward multiplication.[213] The machine, in fact, with "the astounding speed with which it operates, especially doing multiplication and division", was well suited for the calculation of the several sum of squares required in the application of analysis of variance.[214] Fisher's affection for this type of calculating machine has been constantly emphasised in secondary sources and the Millionaire calculator credited as being Fisher's own was still in the office of his successor at Rothamsted, Frank Yates, in the 1970s (Fig. 1.4).[215] Nowadays, not unlike the hand-cranked Brunsviga calculator for Karl Pearson, the Millionaire has been elected as a hallmark of Fisher and his statistical methods.[216]

---

of the British Association Mathematical Tables Committee, 19th November 1930 (Minute Book Ms. Eng. Min. d 1157, BL, University of Oxford).

[210] On the tools for scientific computation popular in the 1920s-1930s in Britain see E. Croarken (1990), pp. 4-21.

[211] For an overview of calculating machines available in the first decades of the twentieth century and for their technical features see E. Martin (1992). In particular on stepped drum and pinwheel machines see respectively pp. 3-11 and pp. 12-17. For the Millionaire calculator see E. Martin (1992), pp. 119-125. A 1970s catalogue of the calculating machines and instruments available at the Science Museum London (D. Baxandall, revised by J. Pugh (1975)) offers a photo gallery of several calculating machines popular in Britain in the early decades of the twentieth century with a brief description.

[212] W. S. Gosset (1962), Letter No. 4, 19th September 1919. "C. Beveridge: hire of Calculating Machine for R. A. Fisher £3.3.- (December 10th 1919)" (Rothamsted Laboratory Cash Account October 1919-January 1921, RR Library and Archive, LAT 34)

[213] "Can you let me have some information about your motor 'millionaire'? I understand it has been in use two or three years" (W. S. Gosset (1962), Letter No. 49, 15th October 1924).

[214] E. Martin (1992), p. 119.

[215] J. Fisher Box (1978), pp. 273-274. The presence of the Millionaire calculator in Yates' office can be gathered from the correspondence between Yates and J. Fisher Box in relation to the making of Fisher's biography (Letter from J. Fisher Box to F. Yates, 15th October 1974, RR Library and Archive STATS 7.5). For the book Yates agrees to be photographed sit at the Millionaire (J. Fisher Box (1978), Plate 23).

[216] For Karl Pearson and the Brunsviga calculator see D. A. Grier (2007), p. 106; p. 117; and T. M. Porter (2006), p. 241.

For a few years probably the Millionaire was the only calculating machine available to the Rothamsted statisticians, but in 1925 Fisher's department began to acquire on a regular basis more calculating machines. Certainly the human computers mentioned above, Rolt and Pennells, used electrical Monroe machines in the early 1930s.[217] Nevertheless, calculating machines were not a commodity in large supply in Fisher's department and finding a machine for the several research workers that visited Rothamsted to learn Fisher's statistical methods was not an easy matter. The first worker who joined Fisher's department for a few months at the end of 1922 was Edward Somerfield, the assistant of William Gosset at the Guinness Brewery. Somerfield came to Rothamsted with a calculating machine of his own, and again in 1930 his colleague, A. L. Murray, sent by Guinness to Fisher's department, had to bring along his own calculating machine, because Fisher wrote to Gosset "[w]e are rather choc-a-bloc at the moment, and are putting people about in the libraries etc. However the fact that you can send a machine too, makes a great difference as we only need to find him 'somewhere to put it' ".[218]

At Rothamsted the shortage of calculating machines did not hit only upon the visitors. Machines might be scarce also for the department staff and, as mentioned above, the local computers should learn to increase "their resources for emergencies when no machine is available" using slide rules and mathematical tables, which, to some extent, supplied to the scarcity of the more expensive computing tools.[219] The accounting books of the experimental station record the purchase of two cylindrical slide rules: in 1922 a second-hand Fuller's slide rule was bought for six pounds and two years later an Otis King's calculator model K was acquired for one pound and one shilling.[220] In the same years the library of the department began to purchase also journals and collections of mathematical tables useful for statistics. Since 1922 the department subscribed to *Biometrika*, the journal edited by Karl Pearson, and bought also a few booklets in the series *Tracts for Computers*, again edited by Pearson and considered a working tool for every human computer of the time.[221]

[217] For the acquisition of calculating machines in Fisher's Department see Rothamsted Laboratory Cash Account October 1919-January 1921, 27th April 1926, RR Library and Archive, LAT 34 and the Rothamsted Staff Council Minutes, 6th January 1925 and 6th December 1929, Volume 2, RR Library and Archive, STA 2.1. For the computing equipment of Rolt and Pennells see F. Yates (1981), p. 138.
[218] W. S. Gosset (1962), Letter No. 17, 20th November 1922; W. S. Gosset (1962), Letter No. 117, 9th October 1930.
[219] Letter from R. A. Fisher to F. Yates, 5th December 1931, RR Library and Archive, STF 31 (old catalogue reference).
[220] On the history of the slide rule see F. Cajori (1994). For an overview of the slide rules available in Britain in the early decades of the twentieth century see also E. M. Horsburgh (1914), pp. 155-180.
[221] "Prof. Karl Pearson, 1 copy Biometrika Vol. 14 Subs 40/. postage 4/. £2.4.0" (28th August 1922) (Rothamsted Laboratory Cash Account October 1920-June 1926, RR Library and Archive, LAT 34). For the purchase of the booklets in the series *Tracts for Computers* see minutes 4th November 1921, minutes 8th December 1921, minutes 7th May 1926 in Rothamsted Staff Council Minutes, RR Library and Archive, STA 2.1.

The equipment acquired by the statistics department represented a consistent upgrade of the computing facilities at Rothamsted. Before the opening of Fisher's department it is likely that the experimental station did not possess any calculating machine and in the 1930s some departments of the station still borrowed or rented computing machines from other departments.[222]

On the other hand, the equipment available to Fisher at Rothamsted was very limited if compared with other computing centres engaged in statistical research, as the mathematical and statistical laboratory run by George Snedecor at Iowa State College in the United States. Snedecor's laboratory can be compared to the Rothamsted statistics department for its mission, because research workers in agriculture mainly benefited from its services. The Iowa facility had already available punched-card equipment since the 1920s and Snedecor was a pioneer in the development of mechanical methods for the computation of least squares.[223] Instead, the statistics department at Rothamsted began to employ these tools only in the late 1930s, when Fisher had already left the station.[224] Fisher himself began to use punched-card equipment during his time at the Galton Laboratory, where he set up a new computing room and benefited from the services of the British Tabulating Machine Company (**chapter 3**).

Fig. 1.4 Millionaire Calculator, credited as being the one used by Ronald Fisher.
Credits: Copyright Rothamsted Research Ltd.

[222] Rothamsted Staff Council Minutes, Vol. 2, 11th January 1934, RR Library and Archive, STA 2.1. In the accounting book from 1913 to September 1919 no calculating machine is mentioned (Rothamsted Laboratory Cash Account 1913-30th September 1919, RR Library and Archive, LAT 34).

[223] On Snedecor's laboratory see D. A. Grier (2007), pp. 166-168 and T. A. Bancroft (1982). For Snedecor's contribution to the computation of least squares with punched-cards equipment see G. W. Snedecor (1928).

[224] In Britain the use of punched-card equipment in business and scientific computation spread during the 1930s, as described by M. Campbell-Kelly (1989), pp. 72-172.

1.5.2 ***The role of the statistics department in the station***

Since the mid-1920s the Rothamsted statistics department had already acquired a crucial position in the research activity done at the experimental station. Its relevance can be examined in relation to two main issues: on the one hand the format and management of the records of the field experiments and on the other the role played by the members of the statistics department in the Field Plots Committee, the body in charge of the planning and execution of the field trials.

a) ***Format and management of the Rothamsted records***

Since Fisher's appointment in 1919, the new discipline of statistics was involved with the analysis of the yields and meteorological records collected at the station and their public presentation. The consistent set of data available at Rothamsted represented the necessary input for the work of the statisticians, as hoped by the Rothamsted director John Russell, and the analysis of variance became in the 1920s the main instrument used in the examination of the past records and in the determination of the experimental error for the current field trials.

As Staffan Müller-Wille has argued in relation to the plant breeding station of Svalöf, record-keeping systems are not neutral tools, but they constrain and are constrained by the kind of scientific research pursued by the institution, which adopts them.[225] The record-keeping system used for the field trials at Rothamsted was formalised well before the opening of the statistics department in the institution and thus it is legitimate to ask how statisticians engaged with the format and management of the station records and to what extent they reshaped it.

A typical record sheet for the Broadbalk wheat experiment completed at the beginning of the twentieth century mentioned the number of the season, the time of the harvest and offered a table with detailed data on the quantity and quality of the yield (dressed grain, offal, straw). The same information was faithfully recorded in the station reports, with at most an indication of the yield averages calculated over a long series of years.[226] The same format was adopted also for the presentation of the results related to the annual experiments with replicated plots conducted at the station since the 1910s.

With the emergence of the statistics department field trials at Rothamsted began to be planned according to Fisher's principles of experimental design and the use of randomisation in the

---

[225] S. Müller-Wille (2005).

[226] Examples of the Broadbalk wheat data were held, uncatalogued, in the former biomathematics and bioinformatics department (now department of computational and systems biology) at the time of my research. They have now been moved to the RR main archive (STATS 6.1, STATS 6.2). For the presentation of the results related to the wheat experiment see, for instance, RES (1923), pp. 85-86 and RES (1925), pp. 108-109.

distribution of treatments to plots allowed the determination of the experimental error with the analysis of variance.[227] The standard error permitted

> to discriminate between cases in which a particular difference in yield can be reasonably set aside as accidental, and cases in which such an explanation requires that an improbable coincidence should be postulated, and in which therefore we have a sound basis for interpreting the difference as a real response to the treatments applied.[228]

A rule of thumb was given: "differences between treatments exceeding three times the standard error may be accepted as significant".[229] As pointed out by Theodore Porter, statistics in the form of error theory "was a strategy for eliminating inference by subjects", but still it needed conventional thresholds, like the rule here discussed or the rejection level adopted in the chi-square test that I mention in the following chapter.[230] Such conventional thresholds remind us that the increase in accuracy offered by the application of statistical methods requested to put aside the intrinsic variability of nature in favour of shared norms.

The adoption of the analysis of variance in the examination of the station experiments did not entail a radical change of the record-keeping system for field trials, but further information should be archived due to the more complex experimental settings. The detailed plans of the field trials should be set out on paper, the field practices adopted in the new trials were also recorded, as well as observations on the growth of crops that were deemed useful for the statistical analysis of the data.[231] If the record-keeping system was not completely transformed, the official presentation of the station experiments required instead a new format, which presented the detailed plan of the experiments and the statistical analysis of their results alongside the original data. Ronald Fisher was instrumental in reshaping such records. He outlined a 'produce sheet' to be used as standard for the presentation of the experimental results including (1) plan of plots, (2) actual plot produce, (3) statistical analysis, (4) summary tables in quantity per acre with statistical summary.[232] This was the format followed for the presentation of field trials in the station reports since the season 1925-1926.[233]

Besides reshaping the presentation of the station records, the Rothamsted statistics department actively contributed to their surveillance and safeguard. In the early 1920s the records

---

[227] RES (1927), p. 122.
[228] RES (1927), p. 122.
[229] RES (1927), p. 122.
[230] T. M. Porter (1996), p. 201.
[231] The layout and working details of the field experiments, as well as the observations on plant physiology, were held by the field experiments department. Some of these data were also duplicated in the farm records. (Interim report (1935) on the system of recording results at Rothamsted, Woburn and on the Farm, RR Library and Archives, E. J. Russell Papers, RUS 4.31). Further information on the data collected during the growth of crops is in the Minutes of the Field Plots Committee, 31st January 1924, RR Library and Archives, FX 1.1.1.
[232] Minutes of the Field Plots Committee, 2nd October 1925, RR Library and Archives, FX 1.1.1.
[233] RES (1927), final section on field experiments.

were kept in a filing cabinet in the administrative office of the station and the research workers could borrow them at their leisure. Ronald Fisher raised the issue of a better surveillance on the records "whereabouts" urging that an efficient system were established to avoid that they went missing. It was thus decided to prevent the fragmentation of the folders of data and to establish a book for recording the loans.[234]

In 1927 the station records were "transferred to the Statistical Dept. with a view to their being eventually incorporated with the records in that department". Fisher took them under his responsibility and made them "accessible to all members of the staff desiring to consult them."[235] By the mid-1930s the statistics department hosted records of the field experiments (yields, chemical analysis of crops, growth observations), further records deposited by other departments for the statistical analysis, and some historical records of the station.[236]

The transformation of the Rothamsted statisticians into recordkeepers was a choice of Ronald Fisher rather than a request of the experimental station. Fisher, in fact, conceived as a task of the statistician not only the analysis of experimental results, but also the collection and preservation of data suitable for statistical examination.[237] This is proved also by his attempt, in the mid-1930s, to put the Rothamsted archive in contact with the Natural History Museum Archives in London, as to guarantee preservation and circulation of the Rothamsted data well beyond the experimental station.[238]

On the other hand, the physical move of the station records into Fisher's department set it at the core of the scientific life of the institution consolidating the role that it had gained with its involvement in the planning and analysis of the field experiments and establishing a decade long tradition. In the archive of the biomathematics and bioinformatics department at Rothamsted Research – the successor of the Rothamsted statistics department, now renamed in turn department of computational and systems biology – most of the materials were folders of data, statistical computations and graphs related to the experiments that the department approached throughout its history, from Fisher's time onwards.[239]

---

[234] Rothamsted Staff Council Minutes, 8th April 1925, RR Library and Archive, STA 2.1.

[235] Rothamsted Staff Council Minutes, 11th July 1927, RR Library and Archive, STA 2.1.

[236] Interim report (1935) on the system of recording results at Rothamsted, Woburn and on the Farm, RR Library and Archives, E. J. Russell Papers, RUS 4.31.

[237] In chapter 3 I describe also the data collection of the ABO blood groups in Britain organised by Fisher during WWII.

[238] Interim report (1935) on the system of recording results at Rothamsted, Woburn and on the Farm, Rothamsted RR Library and Archives, E. J. Russell papers, RUS 4.31. Letter from R. A. Fisher to F. Yates, 29th May 1935, RR, RR Library and Archive, STATS 7.5.

[239] My comment arises from a first-hand experience of the department collection as I found it in 2010-2011, because a comprehensive catalogue of the materials in this archive was not available.

b) *Statisticians in the Field Plots Committee*

The roots of the Field Plots Committee – the Rothamsted body "set up to make sure that experiments are statistically and agriculturally sound, that they are sited on suitable land and that both farm staff and experimenters know their respective responsibilities at every stage" – date back to 1924.[240] The Committee was a forum that offered the opportunity to discuss research plans to experimentalists and statisticians. Both expertises were required as the statistical consultants could suggest solutions for an optimal layout of the field trials in the Rothamsted farm, but crucial to the sound realization of the experiments was the first-hand knowledge of the fields, which was a prerogative of the botanists and ecologists of the station. The principles of experimental design set out by Ronald Fisher had always to be confronted with the peculiarity of the Rothamsted fields and with unforeseeable events, such as weeds and pests, which could undermine even a well-planned agricultural experiment.

Among the members of the Field Plots Committee in the 1920s there were the Rothamsted director Russell, the staff concerned with the field experiments – at the beginning the crop ecologist Eden and the plat physiologist Maskell –, the farm manager, Fisher and his assistant statisticians. My aim is to outline briefly the relevance of Fisher's statistical methods in the planning activity of the Committee and to discuss the capacity in which statisticians took part in it.[241]

Since 1925 the Field Plots Committee considered the suggestion to re-examine the design of all the field experiments at the station "in the light of Mr Fisher's methods".[242] At that stage the field experiments department was already cooperating with Fisher and a year later it was decided that "proposals for field experimental designs might be made known to the Secretary for consultations with Dr Fisher before coming before the meeting".[243] In the same years Latin squares and randomised blocks became the experimental designs widely applied at the station and in the other research institutions under Rothamsted's influence. Since the second half of the 1920s, the overall plan of the field experiments from year to year was discussed using exclusively randomised blocks and Latin squares. Only the long-term experiments were maintained in their original format. [244]

---

240 H. V. Garner (1962), p. 180.

241 The activity of the Field Plots Committee, the problems it discussed and the influence of its members can be reconstructed from the minutes of its meetings. The Library and Archive at Rothamsted Research has a complete record of the Field Plots Committee Minutes (Ref. FX 1).

242 Minutes of the Field Plots Committee, 26th October 1925, RR Library and Archives, FX 1.1.1.

243 Minutes of the Field Plots Committee, 26th November 1926, RR Library and Archives, FX 1.1.1.

244 In each year the Field Plots Committee examined the experiments for the following season discussing plans and arrangements. A few examples of these discussions are in the Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 13th November 1928, 27th May 1931.

Alongside the planning of experiments, in the routine work of the committee Fisher and his assistants were constantly called to examine previous experimental results and suggest whether a trial was worth continuing or not.[245] Their advice was asked in order to combine several investigations into the same experiment or to give suggestions about harvesting plots on which the treatment had been applied without uniformity.[246] Their criticism of proposed experiments was sought and carefully considered by the research workers.[247] Fisher's assistants – at first Wishart and later Yates – gave also full reports of the results extracted from the statistical analysis of the annual experiments before the Committee.[248] Besides his advisory role, Fisher was an active proponent of experiments. He suggested field trials on the manuring of sugar beet and the top dressing of winter corn and to test the 'resistance formula' that Ernest Maskell had worked out to relate the crop yields with variations of the manure dressing.[249]

Fisher was eager to stress also the necessity of efficient sampling techniques in order to safeguard the whole trial against an inaccurate collection of the results.[250] In the late 1920s the Rothamsted statistics and plant physiology departments cooperated in the development of sampling techniques for gathering the yield results. Just before harvest, several samples were taken at random from measured lengths of the rows and weighed if roots or threshed in a miniature machine, if cereals. From the random sample acquired, the whole yield was then estimated without necessity "of separate harvesting, separate stacking, and separate threshing, with all the losses involved".[251] In the early 1930s the systematic development of sampling techniques – again the statistics and plant physiology departments worked in collaboration – involved also testing the efficiency of different forms of sampling units to be taken in the field for the determination of the final yield and for the study of the progress and growth of crops.[252]

---

[245] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 26th November 1926, 4th October 1929, 1st February 1932.

[246] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 22nd February 1927, 13th November 1928, 21st June 1929, 4th October 1929, 4th December 1931.

[247] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 2nd December 1930, 12th January 1931, 1st July 1932.

[248] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 22nd January 1929, 4th December 1931.

[249] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 7th March 1927 (sugar beet), 28th September 1927 (winter corn), 12th January 1931 (Maskell's resistance formula).

[250] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.1: 13th November 1928, 12th January 1931, 1st February 1932, 1st July 1932.

[251] The technique for the random sampling of the yield is described in RES (1929a), p. 39. In the development of the sampling technique the Rothamsted research workers and statisticians collaborated also with the staff of the plant physiology department of Imperial College London under the supervision of V. H. Blackman. (Ministry of Agriculture and Fisheries, Agricultural Research Council, annual report 1927-1928. TNA DSIR/36/4239)

[252] RES (1932). For a further description of the sampling techniques in agricultural experiments developed at Rothamsted see R. A. Fisher's contribution (p. 615) to the discussion of J. Neyman's paper read before the Royal Statistical Society in 1934 (J. Neyman, 1934).

The increasing involvement of Fisher's department in the Field Plots Committee and its collaboration with the research workers engaged in the set-up of the agricultural experiments can be proved also browsing the list of colloquia held in the institution during the 1920s and early 1930s. Since March 1924 Fisher, Eden and Maskell presented a joint seminar for the station workers on a 'Discussion of Field Plot Technique'. This tradition went on in the following years and the sampling of crops was discussed by A. R. Clapham, the plant physiologist who was the second secretary of the Committee from 1927 to 1930. Again in the early 1930s Fisher and his assistants spoke before the scientific staff of the institution on field experiments and their design. The last seminar that Fisher gave at Rothamsted in June 1933 was on 'The results of the field plots experiments'.[253]

In just ten years the statistics department had become a crossroad for all the field research conducted at the institution, from experimental design to the sampling of the crop at harvest, to the best methods of sowing or manuring. The Rothamsted statisticians were thus able to compete for attention and prestige with traditional elements of the scientific life in the station, such as the farm and the field experiments department. That is evident in the cooperative scheme for the set-up of field trials proposed by Fisher in 1930 and accepted by the Committee.

> The initiator shall submit in writing a detailed plan, including the object of the experiment, the number, size, method of arrangement, and treatment of the plots, the crop grown and its proposed position on the Farm to the Statistical Department, the Field Experiments Department and the Farm each of whom must endorse the plan with a statement that they can suggest no improvement, from their respective viewpoints.[254]

In Fisher's scheme statisticians, farm manager and field experiment departments had the same relevance in accepting or rejecting the proposal of any trial to be conducted at the station. To all intents and purposes, Fisher's department had thus become a crucial component of the experimental work at Rothamsted: it was influential in the committee that supervised the field trials and it was in charge of accounting the experimental results both in the official reports and before the scientific staff of the institution.

1.5.3 *Rethinking agricultural research in statistical terms*

So far I have described the structure of the statistics department and the role that it acquired during the 1920s in the field research conducted at Rothamsted Experimental Station. I want to outline now how analysis of variance and experimental design reshaped the approach to field trials at Rothamsted considering the official publications of the station and the practices and instruments adopted in the field experiments.

253 Rothamsted register of the colloquium meetings, RR Library and Archive, E. J. Russell Papers, RUS 4.34.
254 Minutes of the Field Plots Committee, 19th January 1930, RR Library and Archives, FX 1.1.1.

a) ***The Rothamsted Experimental Station reports***

The Rothamsted report for the years 1925-1926 gives the first official account of how field experiments were reshaped by Fisher's statistical methods. Beginning with this report the tables that summarised the results of the replicated field trials done at Rothamsted and Woburn systematically reported the standard error associated to the trial and computed with the analysis of variance, as a benchmark of the accuracy of the experiment.[255]

By 1925 analysis of variance and in part experimental design had found a codification in Fisher's book *Statistical Methods for Research Workers* and – as evident from the previous description of the activity in the Field Plots Committee – they had been integrated in the scientific life of the agricultural station. Thus, it is not surprising that the 1925-1926 station report ratified the new statistical outlook of the research conducted at Rothamsted. The choice to present for the first time Fisher's statistical methods at that stage, however, was also influenced by the termination of the Development Fund, which since then had provided grants to Rothamsted. A full account of the work done at the station could represent an asset for the provision of further money and thus it "was hoped to publish the Report, if possible, some time in January [1927], in order that it might be sent out well ahead of the time when the question of the provision of funds for the continuance of agricultural research work after March 1927, would have to be considered by the Government."[256] Analysis of variance and experimental design were deliberately proposed as a noteworthy contribution of Rothamsted to agricultural science and thus as an asset for further grants.

The 1925-1926 report was just one of the official publications prepared by the institution in which mention was given to the new statistical methods adopted in field trials. Throughout the 1920s several accounts were published in the station reports, insisting on the higher precision now possible in experimental research. The field trials realized according to Fisher's statistical methods were more costly than the traditional experiments because a greater number of replicated plots had to be employed, but they repaid in accuracy. With the new statistical methods the accuracy of the Rothamsted experiments was between two and four per cent in 1929, while traditional experiments rarely gave an accuracy superior to ten per cent.[257] The Rothamsted reports constantly emphasised this aspect as the improvement offered by Fisher's statistical

[255] RES (1927), pp. 26-29; pp. 122-155. The difference between the report for 1925-1926 and the one related to the previous season offers a clear-cut example: in the latter no experimental error is reported in the appendix tables and the only sign of statistical analysis are the averages for the long-term experiments that the statistics department had computed in the previous years from the historical series of data.
[256] Rothamsted Staff Council Minutes, 8th April 1925, RR Library and Archive, STA 2.1.
[257] RES (1930), pp. 45-46.

methods resonated with the Rothamsted agenda to contribute to the development of better agricultural practices. The management of the station and the governing bodies of the institution, in particular the Field Plots Committee, offered to Fisher's statistical methods the institutional backing necessary for their establishment at the experimental station.

In 1933 when Ronald Fisher left Rothamsted, the report for that year paid homage to him printing an account of the scientific activity of his department.[258] A table summarized the growing activity of the local statisticians during the 1920s and in the early 1930s: if in 1925 only eight experiments had been analysed with Fisher's methods, in 1933 this number had reached the amount of ninety-three considering both the experiments done at Rothamsted and Woburn, and the experiments for which the statistics department gave advice to outside centres.[259]

b) *Field practices and instruments*

To comply with Fisher's statistical methods new instruments and practices were employed at Rothamsted for the cultivation of the experimental plots. In the field new drills for sowing seeds and new strategies for manuring were tested, because these two passages were critical for the reliability of the trial.[260] For the harvest of cereal crops a small thresher was purchased as a quicker and more satisfactory solution than the use of the large machine available in the Rothamsted farm, ill-suited to thresh the small harvest given by each experimental plot. Moreover, in order to reduce the dispersion of the yield from the field where it was cut to the farm where it was threshed, the whole product from a plot was placed in a cloth and tied into a bundle in order to improve further the accuracy in the determination of the final yield.[261] For harvesting root crops, instead, a portable scale was provided in order to do the weighing already in the proximity of the experimental plots.[262]

If the planning of the field experiments was the task of the research workers and statisticians at the station, their practical implementation – measurement of the land, preparation of the soil, sowing, manuring, harvesting and weighing of the produce – was instead a duty of the farm staff

---

258 R. A. Fisher (1934).
259 RES (1934), p. 31.
260 "The fact that the size of the standard errors remains approximately the same from year to year, suggests that our present appliances and our methods have reached their limit of accuracy; new and more accurate ones are now being sought. None of the various devices so far tried has constituted any real improvement, and so far as we can see the limit is set by the implements. Both seeds and manure drills are admittedly defective; we have had to return to the old Coulter drill as the best we could find. Application of manures to the replicated plots is always by hand, but we urgently need better seed drills and better methods of distributing the fertiliser so that it shall act most effectively." (RES (1931a), p. 48) For a description of the equipment used on the experimental fields see also B. Weston (1962), pp. 17-18.
261 B. Weston (1962), pp. 36-37.
262 B. Weston (1962), p. 24.

that had to supervise the work of the labourers hired for the task. With the adoption of Fisher's statistical methods in the planning of field trials the overall complexity of the experiments increased and, in order to assist the farm staff, the secretary of the Field Plots Committee revised the experimental plans for the year with the superintendent of the field trials.

B. Weston, who worked as field superintendent at Rothamsted for over thirty years, remembers that the first secretary of the Committee, Eden, helped him in interpreting these new plot designs.[263] The same fruitful collaboration continued with the following secretaries of the Committee, at first A. R. Clapham, and since 1930 the plant physiologist D. J. Watson. In particular, Weston remembers that Watson prepared plans and manure tables and carefully examined them with him. Watson explained to the field superintendent each experiment, and advised him in case of doubts on the execution.[264]

The field superintendent was also in charge of preparing and storing the farm records which consisted of 'white books' – for recording copies of the experimental plans, instructions for the realization of the experiments, dates and details of the field work and crop observations – and an 'harvest book' with the weights of the crop yields, which was handed to the statistics department for the preparation of the station report.[265]

The collaboration between statisticians, research workers and the farm staff was not without frictions, because it was not always clear to the farm staff in charge of the daily working on the experimental plots why the fieldwork had a crucial influence on the statistical analysis of the results. One example is the Broadbalk harvest of 1933. Upon suggestion of Fisher's assistant, Frank Yates, the Field Plots Committee decided that "headlands should be cut out between each fallow strip of Broadbalk before harvest".[266] The farm manager Henry Miller disregarded the decision of the Committee on account that it "would have taken a long time and involved extra expense", and would have prevented the farm staff from taking full advantage of the rapid ripening of the harvest in the season.[267] On behalf of the station director, B. A. Keen, head of the physics department, firmly reproached the farm manager remarking that, regardless of the peculiarity of the season, "the real issue is that a considered decision of the Director and Plot[s] Committee on one of our classical fields, has been over-ruled by you [Miller] on the grounds of expense and impracticability".[268] Keen remarked how the insubordinate behaviour of the farm

---

263 B. Weston (1962), pp. 32.
264 B. Weston (1962), pp. 32-33.
265 Interim report (1935) on the system of recording results at Rothamsted, Woburn and on the Farm, RR Library and Archives, E. J. Russell Papers, RUS 4.31; B. Weston (1962), p. 48.
266 Letter from D. J. Watson to B. A. Keen, 28th July 1933, RR Library and Archive, E. J. Russell Papers, RUS 2.9; Minutes of the Field Plots Committee, 6th July 1933, RR Library and Archives, FX 1.1.2.
267 Letter from H. G. Miller to B. A. Keen, 2nd August 1933, RR Library and Archive, E. J. Russell Papers, RUS 2.9.
268 Letter from B. A. Keen to H. G. Miller, 3rd August 1933, RR Library and Archive, E. J. Russell Papers, RUS 2.9.

manager was a potential menace for all the fieldwork of the experimental station, "since there would be no guarantee that the details of any experiment were correctly carried out".[269]

To the accusation Miller replied that the decision of the Committee, taken in his absence, could not be a considered one, as it disregarded the increase of cost due to the new harvesting plan for the Broadbalk and the effects of this greater work on the harvest of the other field experiments at the station.[270] To avoid further problems for the Broadbalk field, in January 1934 the Field Plots Committee decided that "the headlands [...] should be marked out by paths so as to facilitate the scything" at harvest.[271] No official rebuke to Miller is listed in the minutes of the Field Plots Committee, which instead had to call an extraordinary meeting in September 1933 to make special arrangements for the autumn season, as Miller fell ill with tuberculosis. The Rothamsted farm manager never recovered and prematurely died, aged 31, in April of the following year.[272]

Despite the controversy described, all in all, the scheme for field trials adopted at Rothamsted worked for decades guaranteeing both the set up of hundred of experiments and the preservation of the experimental farm land.[273]

Fig. 1.5 Tying cloth on the harvest of each experimental plot.
Credits: B. Weston, *Fifty Years at Rothamsted 1910-1960* (1962).

[269] Letter from B. A. Keen to H. G. Miller, 3rd August 1933, RR Library and Archive, E. J. Russell Papers, RUS 2.9.
[270] Letter from H. G. Miller to B. A. Keen, 4th August 1933, RR Library and Archive, E. J. Russell Papers, RUS 2.9.
[271] Minutes of the Field Plots Committee, RR Library and Archives, FX 1.1.2: 6th July 1933, 24th January 1934.
[272] Minutes of the Field Plots Committee, 29th September 1933, RR Library and Archives, FX 1.1.2; Obituary of H. G. Miller published in RES (1935b), p. 63.
[273] H. V. Garner (1962), p. 183.

Fig. 1.6 Threshing the harvest of the experimental plots with a miniature machine (1934).
Credits: B. Weston, *Fifty Years at Rothamsted 1910-1960* (1962).

Fig. 1.7 Weighing sugar beet leaf and roots (1933). To be noted the field notebook in the hands of the farm staff for recording the results of the harvest.
Credits: B. Weston, *Fifty Years at Rothamsted 1910-1960* (1962).

1.5.4 *Statistical methods and the lab-field border in agriculture*

Since the foundation of Rothamsted in the nineteenth century the agricultural science practiced at the station included both field trials and laboratory investigations. Every year, besides the agricultural experiments, soil and crop samples were collected and analysed in the laboratories and became part of the sample archive instituted at the experimental station.[274]

In the 1920s the four main departments at Rothamsted – biological, chemical, physical and statistical – and the farm were both distinct and interrelated experimental contexts.[275] Crop and soil samples were brought into the biological, chemical and physical laboratories for examination, while a field laboratory was built in the experimental fields for a closer scrutiny of nature. In this

274 RR (2006), p. 46.
275 The method adopted was "to start from the farm and work to the laboratory, or vice-versa" (RES (1921), p.8).

two-way relation between laboratory and field research the statistics department offered tools that made more lab-like field experiments and that were suited, as well, for laboratory investigations.

Robert Kohler has named lab-field border the cultural space "where laboratory and field practices can meet and mingle".[276] According to Kohler

> [t]he domains of laboratory and field are cultural domains first and foremost, where different languages, customs, material and moral economies, and ways of life prevail. […] the boundary between lab and field cannot be demarcated by a line [...] rather it is a zone of mixed practices and ambiguous identities.[277]

In the examination of the agricultural experiments the analysis of variance allowed to isolate the influence of several factors (fertilisers, weeds, soil heterogeneity etc.), thus reducing the complex field phenomena to a series of single causes investigation "which can be dealt with by the methods of plant physiology".[278] That was in agreement with the philosophy of laboratory research in which "experimenters analyze and reveal causes and effects".[279] Furthermore, the use of a randomised experimental arrangement allowed to estimate the experimental error of the field trials and to increase their accuracy. Field practice was so aligned to laboratory standards striving for precision, but due respect was also paid to the complexity of nature in which the mutual relations of several factors could not be neglected.

This change was not without consequences, as I have described above. It required the acceptance of mathematical reasoning and number crunching as activities integral to the experimental research of the station. In fact, the Rothamsted director prized the statistician's ability "to impress upon the biologist and agriculturist the necessity for rigid mathematical tests of significance in place of the older and still much too common 'biological feeling'".[280] Mathematical and computing expertise belonged to the Rothamsted statisticians rather than agronomists and life scientists, thus statisticians closely collaborated with the research workers in the examination of the experimental results and in the planning of the field trials.

The role acquired by Fisher's statistical methods was precisely the desideratum that accompanied the opening of the statistics department in 1920.

> On the farm [...] many factors may operate and elimination results in conditions so artificial as to render the enquiry meaningless. In place, therefore, of the ordinary single factor method of the

[276] R. E. Kohler (2002), p. 51.
[277] R. E. Kohler (2002), p. 5.
[278] RES (1925), p. 15.
[279] R. E. Kohler (2002), p. 2. However, the observational component typical of fieldwork was still alive at Rothamsted as ecologists, agriculturists, plant physiologists constantly monitored the conditions of the experimental plots.
[280] Letter from E. J. Russell to R. H. Fowler, 29th January 1929, E. J. Russell Papers, RR Library and Archive, RUS 2.5. The letter belongs to the correspondence related to Fisher's election as fellow of the Royal Society.

> scientific laboratory, liberal use is made of statistical methods which allow the investigation of cases where several factors vary simultaneously. Thus in the crop investigations a large number of field observations are made; these are then treated statistically to ascertain the varying degrees to which they are related to other factors – such as rainfall, temperature, etc. – and to indicate the probable nature of the relationships. Thus the complex problem becomes reduced to a number of simpler ones susceptible of laboratory investigation.[281]

As I will discuss in the following section, analysis of variance and experimental design were introduced as well in the laboratory investigations conducted at Rothamsted and statisticians were also called as consultants at the lab-bench. Therefore, Fisher's statistical methods transformed the whole lab-field border at the station, as they set new standards of precision common to both laboratory and field research.

1.6 *The diffusion of analysis of variance and experimental design*

1.6.1 *Beyond agricultural experiments*

Since the early 1920s it was recognised that analysis of variance and experimental design could have a wider application than just field trials in agriculture. At Rothamsted they were employed in the bacteriology department for the study of the numbers of bacteria in soil, in the entomology department for studying bees and other insects, in the chemistry department for extracting information from the figures accumulated during the laboratory investigations.[282]

It was biology the discipline which, alongside with field experiments in agriculture, immediately benefited from Fisher's statistical methods.[283] Already in 1922 the Rothamsted biologists and Ronald Fisher published joint papers on the statistical analysis of laboratory data. In particular, I will mention here the collaboration between Fisher and the bacteriologist Henry G. Thornton and the cooperation of the statistician with the entomologist James Davidson.

Thornton was appointed head of the Rothamsted bacteriology department in 1919 to study the effects of the partial sterilisation of soil.[284] The first researches that he undertook in the experimental station concentrated on the counting of the number of bacteria in soil and some of the problems of enumeration were solved with the application of Fisher's statistical methods. The results were published in the *Annals of Applied Biology* in a joint paper by Fisher, Thornton and Fisher's assistant, Mackenzie.[285] The statisticians proved that "under ideal conditions the

[281] RES (1921), p. 8.
[282] RES (1931a), p. 53.
[283] RES (1931a), p. 53.
[284] A scientific biography of H. G. Thornton is offered by his obituary as fellow of the Royal Society (P. S. Nutman (1977)).
[285] R. A. Fisher et al. (1922).

bacterial counts on parallel plates will vary in the same manner as samples from a Poisson Series" and using the analysis of variance, they gave Thornton an estimate of the accuracy of his method of counting bacteria on agar medium.[286] The statistical tests proved that "the necessary perfection of technique was effectively realised".[287]

Again in the *Annals of Applied Biology* it was published a research to which Fisher contributed with his statistical expertise. The entomologist James Davidson presented a comparative study of different varieties of beans infested by the pest *Aphis rumicis*.[288] In Davidson's paper a method for counting the bean pest was presented and Fisher wrote a statistical appendix in which he discussed "the probable error to be attached to Dr Davidson's aphis infestation numbers".[289]

Thornton's and Davidson's researches were laboratory investigations: the former engaged the Rothamsted protozoologists and bacteriologists in the standardisation of the plating method of counting soil bacteria, the latter was instead conducted in a glasshouse using bean plants in flower pots. In both cases statistics was instrumental in assessing the accuracy of the experimental set-up adopted in the laboratory.

The design of experiments entered, as well, in the practice of the Rothamsted laboratory workers. Already in the early 1930s Frank Yates advised the botanist Winifred Brenchley to randomise her pot cultures for testing the effects of boron dressing on beans. Otherwise, the competition for light and air among the plants within each set of replicates would have concealed the real effects of the fertilizers.[290]

The prompt passage of analysis of variance and experimental design from agriculture to biology and from field trials to laboratory investigations is evident in *Statistical Methods for Research Workers*. The book, in fact, belonged to a series of biological monographs and manuals and the original title planned for the book was *Statistics for Biological Research Workers*.[291] Since the first edition, instead, the book was more generally called *Statistical Methods for Research Workers*, but the author presented it stressing the special acquaintance with biologists – "for several years the author has been working in somewhat intimate co-operation with a number of biological research departments; the present book is in every sense the product of this circumstance" – and giving credit to his "daily contact with the statistical problems which present themselves to the laboratory worker" as the motivation behind his mathematical investigations.[292] Nevertheless, the

[286] R. A. Fisher et al. (1922), pp. 357-358.
[287] RES (1923), p. 35.
[288] J. Davidson (1922). For a biography of Davidson see T. O. Browning (1981).
[289] R. A. Fisher (1922b), p. 142.
[290] F. Yates (1990), p. *xxii*.
[291] A. W. F. Edwards (2005a), p. 860.
[292] R. A. Fisher (1925a), p. *vii*.

practical examples in the book were freely chosen in agriculture and biology and concerned field trials, as well as laboratory investigations.

Besides biology, many more disciplines began to apply Fisher's statistical methods.[293] In **chapter 2** I trace the diffusion of Fisher's methods using the several requests for reproducing materials from the *Statistical Tables for Biological, Agricultural and Medical Research*, the collection of tables co-authored by Fisher and Yates and planned as a tool for the dissemination of analysis of variance and experimental design. By the end of the 1950s Fisher's methods had proved their usefulness and versatility in laboratory and field research and had come to stay in psychology, sociology, education, chemistry, medicine, engineering, economics, quality control, just to mention a few of the disciplines which adopted them.[294]

1.6.2 *Beyond Rothamsted Experimental Station*

During the time that Fisher spent at Rothamsted over fifty people came to his department as visiting workers to learn analysis of variance and experimental design.[295] They were postgraduate students, temporary workers or, as called in general by Joan Fisher Box, voluntary workers.[296] In Table c (Appendix 1.1) I have prepared a list of such workers using as sources the *Rothamsted Experimental Station Reports*, the *Records of the Rothamsted Staff Harpenden*, the minutes of the Rothamsted staff council and a list of voluntary workers already published by Nancy Hall.

The majority of the workers in Fisher' department came from abroad, especially the United States and the British Empire. They belonged to a wide array of disciplines: over half of them were agriculturists, botanists, plant breeders, but there were also statisticians, like the English Leonard Tippett and the American Harold Hotelling, one zoologist, C. H. N. Jackson, from the Tanganyika Territory, and a sociologist, the American Samuel A. Stouffer, who promptly applied Fisher's methods to his social research.[297] A few workers came also from private companies, like the Imperial Chemical Industries or the Guinness brewery in Dublin that sent two assistants of William Gosset to work under Ronald Fisher.

The visitors in Fisher's department were both researchers already established in their careers, as well as postgraduate workers, and they were all supported directly by their home institutions or

---

293 For the diffusion of Fisher's statistical methods see G. Gigerenzer et al. (1989), pp. 114-115; p. 118.
294 The diffusion of statistics in medicine after WWII, with specific reference to Fisher's statistical methods, is examined by H. M. Marks (2000), pp. 129-163.
295 The stream of visitors in Fisher's department became consistent from the late 1920s onwards.
296 J. Fisher Box (1978), pp. 241-243.
297 Stouffer is explicitly mentioned by N. S. Hall in her unpublished talk "Did Fisher's voluntary workers at Rothamsted make a difference in the spread of statistical techniques in agriculture?". Hall lists the publications written by the sociologist and influenced by the application of Fisher's statistical methods. These contributions appeared in both sociology journals and the *Journal of the American Statistical Association* (N. S. Hall, unpublished).

through scholarships offered by foundations and research councils. For example the American botanist Edgar Anderson benefited from a scholarship of the International Education Board of the Rockefeller Foundation for spending a couple of months in Fisher's department, while Frances Allen, the only woman in the list, was sponsored by a studentship from the Council for Scientific and Industrial Research, Melbourne.

Nancy Hall has examined a selection of the voluntary workers that came to Fisher's department arguing that they contributed to the dissemination of Fisher's statistical methods in their own disciplines and institutions.[298] To support her thesis Hall focuses on the publications of the voluntary workers after their Rothamsted experience and claims that the statistical books and papers that they wrote made Fisher's methods known to their fellow researchers. In **chapter 3** I examine the adoption of statistical methods in serology, and I argue, as well, that the serologists who closely collaborated with Fisher were instrumental in the dissemination of his statistical methods among their colleagues.

However, this general trend admits exceptions. Not all the research workers who learnt Fisher's statistical methods at Rothamsted immediately applied them in their own profession. A counterexample is discussed by the historian Joel Hagen and relates to the botanist Edgar Anderson, who came to Fisher's department in 1929. Anderson was interested in statistics as a subject, but for his work in systematics he felt that refined statistical tools of high mathematical complexity did not give significantly better results than more intuitive procedures, like the graphs he had devised. Anderson was grateful to Fisher for his teaching, but Fisher "did not in the end sell me [Anderson] on using his methods for my problems".[299]

Besides the stream of visitors in the Rothamsted statistics department, a notable contribution to the dissemination of analysis o variance and experimental design came from the books that Fisher published during the 1920s and 1930s. With the integration of Fisher's methods in new scientific disciplines, his books became a necessary companion for many research workers and statisticians. The editorial success of *Statistical Methods for Research Workers*, the *Design of Experiments* and the *Statistical Tables for Biological, Agricultural and Medical Research* is a good yardstick of the popularity acquired by Fisher's statistical methods in the next three decades. By 1963 thirty-six thousand copies of the English edition of *Statistical Methods for Research Workers* had been sold and the book had been translated into French, German, Italian, Spanish and Japanese. The success of *The Design of Experiments* was not much different, with thirty-two thousand and five hundred copies sold in the English edition and translations into Italian and Spanish. The *Statistical Tables*,

[298] N. S. Hall, unpublished talk.
[299] J. Hagen (2003), p. 361.

presented as a companion to Fisher's textbooks, were translated in Spanish and by 1963 their English edition had sold twenty-two thousand copies.[300]

A further opportunity for the dissemination of Fisher's statistical methods was offered by the lectures that Fisher gave abroad. In particular, the diffusion of Fisher's statistical methods in the U.S. was promoted by Fisher's personal and professional acquaintance with George Snedecor, director of the already mentioned statistics laboratory of Iowa State College. Snedecor called Fisher as a lecturer in the Iowa summer courses of 1931 and 1936 and Snedecor "in his teaching and in his well-known book *Statistical Methods*, made Fisher's methods available to a host of workers in agronomy and animal husbandry".[301]

1. 7 *Conclusion*

Analysis of variance and experimental design are almost a century old, but still very much alive in the experimental research conducted in field and laboratories, in and beyond agriculture. In the present chapter I have retraced their development and how they reshaped the research in the institution in which they were first applied. My aim has been to examine Ronald Fisher's statistical tools not as a mathematical achievement *per se*, but as a new opportunity for experimental research in order to increase the precision of both field trials and laboratory investigations.

I have presented analysis of variance and experimental design as instrumental in supporting the role of Ronald Fisher and his assistants as consultants of the Rothamsted research workers and I have described the development of Fisher's methods as an answer to the concrete problems of experimentation in the station.

I have approached the history of statistics not as a branch of the history of mathematics, but as a basic component of the history of experimental sciences during the twentieth century. Thus I have discussed the development of Fisher's statistical methods as a joint effort of statisticians and research workers and I have highlighted how the introduction of these new mathematical tools impacted on practical elements, such as the instruments used in the cultivation of the experimental plots, the organization of the field trials and the presentation of their results.

Writing my account I have used not only primary and secondary sources related to Fisher's statistical methods, but I have also resorted to the official reports of the station and the archive materials related to the organisation of the research work at Rothamsted. I believe that this shift

---

300 Letter from F. Yates to Oliver and Boyd, 10th June 1963, Oliver and Boyd Collection, NLS.
301 G. Gigerenzer et al. (1989), p. 118.

in the choice of the sources has been crucial for understanding the development of statistical methods for experimental research in and beyond agriculture. The role of the statisticians at Rothamsted is here framed in a bigger picture, in which research workers and farm staff interact with the mathematical consultants.

The account I have given of the set-up and organization of Fisher's department helps further to unravel the changes brought by the introduction of statistical methods at Rothamsted. Statistics had its own needs and the opening of Fisher's department required the acquisition of professional computing equipment and the appointment of a small group of human computers. Number crunching was necessary in the day-by-day application of Fisher's methods to the planning and analysis of the station experiments and in the development of computing tools for statistics (**chapter 2**, **chapter 4**).

The methods that Fisher devised for field trials in agriculture suited as well the needs of the laboratory investigations and statistics became an instrument for promoting the convergence of the research done at Rothamsted in the experimental fields and in the laboratories. Analysis of variance and experimental design won their way in the station border area between lab and field and required the cooperation of statisticians, plant physiologists, crop ecologists, laboratory workers. At Rothamsted Fisher's statistical methods, along with the computing activity they required, were enrolled among the tools of experimental research to increase the standards of accuracy in agriculture and biology.

APPENDIX 1.1: Staff of the Rothamsted statistics department, 1919-1933.

a. Statisticians

Name	Appointment	Education	Further career
Ronald Fisher	1919-1933 1933-1939 (honorary consultant)	M.A. (Cambridge) Sc.D. (Cambridge)	1933-1943 Galton professor of eugenics, University College London. 1943-1957 Balfour professor of genetics, Cambridge University.
Winifred Mackenzie	1920-1927	B.Sc (Econ)	------
John Wishart	1927-1930	M.A. (Edinburgh) D.Sc. (University of London)	1931-1956 Reader in statistics, School of Agriculture, Cambridge University. 1953-1956 Director of the statistical laboratory, Cambridge University.
Joseph Irwin	1928-1931	B.A. (Cambridge) M.Sc. (University of London) M.A. (Cambridge) D.Sc. (University of London) Sc.D. (Cambridge)	1931-1965 Statistician appointed to work for the Medical Research Council.
Margaret Webster	1930-1933	B.A.	------
Frank Yates	1931-1968	B.A. (Cambridge) Sc.D. (Cambridge)	1968-1978 Senior research fellow at Imperial College London. 1968-1993 Rothamsted honorary fellow.

b. Computers

Name	Appointment
W. D. Christmas (honorary)	1921-1931
A. D. Dunkley	1921-1932
Kathleen Abbott	1923-1926
Florence Pennells	1927-1937
Alice Kingham	1927-1930
Kitty Rolt	1929-1935
J. M. West	1933-1934
Margaret Dunckley	1933

c. Postgraduate and Voluntary Workers*

Name	Period	Education	Institution and sponsorship
E. Somerfield	Dec. 1922-Apr. 1923	-----	Assistant of W. S. Gosset, Guinness Brewery Dublin.
L. H. C. Tippett	1923-1925	B.Sc. A.R.C.S.	British Cotton Industries Research Association (Shirley Institute, Manchester).
J. E. James	1926 (one month)	-----	Sent by Colonial Office
T. N. Hoblyn	1925-1926	-----	East Malling Research Institute
(Prof.) B. Balmukand	Oct. 1927-Jul. 1928	-----	Agricultural College Lyallpur, Bengal
A. J. Page	Oct. 1927-1928	-----	I.C.S., Burmah
D. W. Boehme	Jun.-Aug. 1928	Ph.D.	Halle
W. H. Beckett	Sept.-Oct. 1928	-----	Assistant Superintendent, Department of Agriculture, Acra, Gold Coast.
J. B. Hutchinson	1928	M.A.	Empire Cotton Research Station, Trinidad
H. Hotelling	Jun.-Dec. 1929	Ph.D.	Stanford University, California
H. G. Sanders	1929	-----	-----
B. P. Scattergood (honorary)	1927-1929	M.A.	-----
J. Pepper	Aug.-Sept. 1929	M.A.	Postgraduate student at University College London.
G. W. Nye	Aug.-Sept. 1929	-----	Agricultural Department, Campala, Uganda
W. G. Eggleton	Sept.-Oct. 1929	-----	Agricultural Advisory Department, Imperial Chemical Industries.
R. J. Kalamkar	Sept. 1929-Apr. 1932	B.Sc., B. Agric.	Nagpur University, Central Provinces, India
Frances E. Allan	Oct. 1929-Jul. 1930	M.A.	University of Melbourne. Holder of travelling studentship from Council for Scientific and Industrial Research, Melbourne.
H. W. Jack	Nov. 1929	-----	Economic Botanist Agricultural Department, Kuala Lumpur.
J. W. Hopkins	1930-1932	M.Sc.	University of Alberta, Edmonton, Canada
C. H. N. Jackson	Jan.-Feb. 1930	M.Sc.	Zoologist, Tsetse Research Laboratory, Tanganyika Territory.
E. Anderson	Feb. 1930; Sept. 1930	-----	Missouri Botanical Garden (sponsored by the Rockefeller Foundation, International Education Board)
H. C. Arnold	May 1930	-----	Agricultural Department Salisbury, Rhodesia
(Prof.) A. de Oliveira Franco	May-Jun. 1930	Ph.D.	Chief of Technical Section, Bureau of Cotton, Ministry of Agriculture, Rio de Janeiro, Brazil.
B. Christidis	Jul. 1930	-----	Plant Breeding Station, Salonika
C. H. Goulder	Jul.-Aug. 1930	-----	Dominion Rust Research Laboratory, Manitoba

			Agricultural College, Winnipeg.
A. W. R. Joachim [worked also in the chemistry and bacteriology laboratories at Rothamsted]	Sept.-Oct. 1930	B.Sc.	Department of Agriculture, Ceylon
A. L. Murray	Oct.1930-Apr. 1931	B.A.	Assistant of W. S. Gosset, Guinness Brewery, Dublin.
F. R. Immer	Oct.1930-Jun. 1931	Ph.D.	Associate geneticist, U.S. Department of Agriculture, University Farm, St. Paul, Minnesota (Rockefeller Foundation Fellowship).
(Prof.) R. F. Summerby	Feb.-Jun. 1931	-----	Agronomy Department, MacDonald College, Quebec.
S. H. Justensen	Mar.-Jun. 1931	Ph.D.	The University, Wageningen, Holland
H. R. Hoskins	Apr. 1931	-----	Serere Experiment Station, Uganda
J. T. Campbell	Jul. 1931	-----	Fellowship from University of New Zealand
F. Billington	Jul.-Aug. 1931	-----	-----
H. B. Bescoby	Sept. 1931	-----	Wye Agricultural College
H. J. Buchanan-Wollaston	Nov. 1931	-----	Fisheries Laboratory, Lowestoft
T. Eden	1932 (arr. Feb.)	-----	Tea Research Institute, Ceylon
H. B. Bescoby	1932 (arr. Apr.)	-----	Wye Agricultural College
S. A. Stouffer	Apr.-Aug. 1932	-----	Department of Sociology, University of Wisconsin, Madison.
R. O. Iliffe	May-Jul. 1932	M.A.	Agricultural Research Institute, Coimbatore, India.
R. S. Koshal	May 1932-May 1933	M.Sc.	Senior Research Assistant, Technical Research Laboratory, Bombay.
I. Bachér	Jul.-Sept. 1932	-----	Agricultural Department, Central Experiment Station, Stockholm, Sweden.
P. E. Turner	Jul.; Sept. 1932	-----	Imperial College of Agriculture, Trinidad
J. Rasmussen	Jul.-Aug. 1932	Ph.D.	Seed Breeding Station, Lund University, Svalöf, Sweden.
C. Stuart Christian	Oct. 1932-Mar. 1933	B.Sc. (Agric)	Department of Genetics, Division of Plant Industry, Queensland University Brisbane (research fellowship from Council of Scientific and Industrial Research, Australia).
R. K. S. Murray	Nov. 1932	-----	Rubber Research Scheme, Neboda, Ceylon
R. A. Taylor	Nov-Dec. 1932	B.Sc.	St. Andrews

A. Bigot	Jan.-May 1933	-----	Agricultural High School, Wageningen, Holland, scholarship L. E. B. Foundation.
R. A. Scott	Feb-May 1933	-----	Department of Agriculture, Launceston, Tasmania.
S. S. Wilks	Jan.-Apr. 1933	Ph.D.	Columbia University, New York
H. L. G. Milne	May-Jul. 1933	-----	Department of Agriculture, Entebbe, Uganda
J. B. Hutchinson	1928; May-Jun. 1933	M.A.	Institute of Plant Industry, Indore, Central India
A. P. Malan	Jul.-Sept. 1933	M.Sc.	School of Agriculture Cambridge (working under Wishart). Previously University of S. Africa.
I. Zacopanay	1933-1934	B.A. (Agric)	-----
A. V. Coombs	1933-1934	-----	Appointed to work with Imperial Chemical Industries at Colombo, Ceylon.

* Sources used in the preparation of the table: 1) *Reports* (1919-1933) of Rothamsted Experimental Station; 2) *Records of the Rothamsted Staff Harpenden* (1929-1935); 3) Rothamsted Staff Council Minutes Volume 1 and 2, RR Library and Archive, Ref. STA 2.1; 4) List of Fisher's voluntary workers at Rothamsted prepared by N. S. Hall (N. S. Hall (2007), pp. 321-322).

Archive materials

Barr Smith Library, The University of Adelaide

1) R. A. Fisher Papers (digitized)

- Letter from R. A. Fisher to T. Eden, 1st December 1931 [accessed 10th June 2012 http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67664/6/1931-12-01.pdf].

2) R. A. Fisher Papers

- Correspondence with T. Eden

Letter from T. Eden to R. A. Fisher, 11th August 1930.

Letter from T. Eden to R. A. Fisher, 1st April 1931.

Letter from T. Eden to R. A. Fisher, 25th August 1942.

- Correspondence with F. Yates

Letter from F. Pennells and K. Rolt to R. A. Fisher, 3rd December 1931.

Letter from F. Yates to R. A. Fisher, 7th December 1931.

Letter from F. Pennells and K. Rolt to R. A. Fisher, 12th December 1931.

Bodleian Libraries, University of Oxford

The British Association Mathematical Tables Committee, 1871-1948, Minute Book of the Committee, 1928-1930, Ms. Eng. Min. d 1157.

Museum of English Rural Life, University of Reading

Thomas Eden's résumé, ca. 1946, E. J. Russell Correspondence, FR HERT 11/1/1.

The National Archives of the UK

1) Records of the Cabinet Office. Office of the Lord President of the Council: Registered Files, Correspondence and Papers, Ref. CAB/123.

- Report of the Committee on Agricultural Research Organisation, Economic Advisory Council, 29th April 1930, CAB/123/275.

2) Department of Scientific and Industrial Research and related bodies, Ref. DSIR 36

- Ministry of Agriculture and Fisheries: Agricultural Research Council. Annual report 1924-1925, DSIR/36/4239.
- Ministry of Agriculture and Fisheries, Agricultural Research Council. Annual report 1927-1928, DSIR/36/4239.

National Library of Scotland

Letter from F. Yates to Oliver and Boyd, 10th June 1963, Oliver and Boyd Collection, Acc.5000/Roneo System/Box 980.

Rothamsted Research, Library and Archive

1) E. J. Russell Papers

- Letter from E. J. Russell to R. H. Fowler, 29th January 1929, RUS 2.5.
- Letter from E. J. Russell to R. Biffen, 3rd October 1929, RUS 2.7.
- Letter from R. Biffen to E. J Russell, 4th October 1929, RUS 2.7.

- Interim report on the system of recording results at Rothamsted, Woburn and on the Farm, RUS 4.31.
- Letter from D. J. Watson to B. A. Keen, 28th July 1933, RUS 2.9.
- Letter from H. G. Miller to B. A. Keen, 2nd August 1933, RUS 2.9.
- Letter from B. A. Keen to H. G. Miller, 3rd August 1933, RUS 2.9.
- Letter from H. G. Miller to B. A. Keen, 4th August 1933, RUS 2.9.
- Rothamsted register of the colloquium meetings, RUS 4.34.

2) Laboratory Accounts, 1913-1947, Ref. LAT 34.
- Rothamsted Laboratory Cash Account, 1913-30th September 1919.
- Rothamsted Laboratory Cash Account, October 1919-January 1921.

3) Field Plots Committee, Ref. FX.
- Minutes 1st meeting, 31st January 1924, to Minutes 40th meeting, 1st July 1932, FX 1.1.1.
- Minutes 41st meeting, 30th November 1932, to Minutes 64th meeting, 2nd July 1940, FX 1.1.2.

4) Rothamsted Staff Council Minutes, Ref. STA 2.1.
- Volume 1, January 1921 – April 1928.
- Volume 2, May 1928 – October 1934.

5) Statistics, 20th century, Ref. STATS.
- Broadbalk wheat data, STATS 6.1, STATS 6.2.
- Letter from W. S. Gosset to A. D. Hall, 8th December 1910, STATS 12.
- Letter from R. A. Fisher to F. Yates, 9th May 1931, STF 31 (old catalogue reference).
- Appointment of a member on the staff of the Rothamsted Experimental Station: Frank Yates, 14th September 1931, STF 31 (old catalogue reference).
- Letter from R. A. Fisher to F. Pennells and K. Rolt, 5th December 1931, STF 31 (old catalogue reference).
- Letter from R. A. Fisher to F. Yates, 5th December 1931, STF 31 (old catalogue reference).
- Letter (copy) from E. J. Russell to R. A. Fisher, 21st December 1931, STF 31 (old catalogue reference).
- Letter from F. Pennells and K. Rolt to F. Yates, 29th December 1931, STF 31 (old catalogue reference).
- Report of the Meeting of the Sub-committee on animal husbandry, 2nd February 1932, STATS 6.3.
- Letter from R. A. Fisher to F. Yates, 29th May 1935, STATS 7.5.
- Letter from L. J. Comrie to F. Yates, 18th October 1938, STATS 8, D.A. Boyd Papers.
- Letter from J. Fisher Box to F. Yates, 15th October 1974, STATS 7.5.

<u>Special Collections Department, Iowa State University Library</u>

Letter from R. A. Fisher to G. W. Snedecor, 6th January 1934, G. W. Snedecor Papers, RS 13/24/51, Box 1, Folder 9.

Chapter 2

THE POLITICS OF THE *STATISTICAL TABLES FOR BIOLOGICAL, AGRICULTURAL AND MEDICAL RESEARCH*

2.1 *Introduction*

Mathematical tables deserve a place of their own among the computing tools that promoted the application of statistical methods to scientific research in the twentieth century. Available in a great variety of formats, up until the 1980s, collections of mathematical tables were a flexible and handy companion for statisticians and research workers engaged in the analysis of experimental data. Printed in books and journals, they represented a commodity in the community of statisticians and table making became often a side activity of statisticians and statistical laboratories. In a few cases it was a one-man enterprise, but more often it was the availability of computing facilities that made the effort possible.

Mathematical formulae and number crunching were the raw matter of table making for statistics, but the net outcome was far more than a linear combination of them. Striking differences could arise in the format of tables for the same function, but prepared with different statistical methods in mind, because making a new table required many subjective choices such as deciding the variables to be tabulated in rows and columns and the intervals to be used. If subjective choices were possible for a single statistical table, assembling a whole collection offered open spaces of negotiation in relation to the theoretical principles underpinning it, its arrangement and copyright, and its prospective users. In the context of these negotiations forms of power and authority could be embedded in the computing tool making it a value-laden artefact – that is a political artefact – within the community of its users.[302]

In their *History of Mathematical Tables*, Campbell-Kelly and colleagues are ready to acknowledge that empirical tables for statistics, like the ones of the census data, are strongly influenced by social conventions and by the aims of the political authority, but they regard the statistical tables computed from objective mathematical formulae as value-free.[303] Instead, for the reasons I have mentioned above, collections of mathematical tables cannot be considered neutral just because they are the outcome of mathematical formulae and number crunching. Quite the opposite, they

---

[302] The idea of investigating the politics of statistical tables is already in D. A. Mackenzie (1981), pp. 153-182. Mackenzie, however, examines the debated interpretation for the statistical association between data arranged in contingency tables, not the making of a whole collection of statistical tables.
[303] M. Campbell-Kelly et al. (2003), pp. 4-5.

should be carefully cross-examined because tacit aims and goals can be easily hidden in a collection of numbers.

The study of the political dimension of artefacts is an established field in the historiography of science and technology and in the past decades several contributions have been published on the topic.[304] Langdon Winner is one of the main authors in this field of research and in the present chapter I will refer in detail to the paper "Do artifacts have politics?", in which Winner examines two main categories of political technologies, the artefacts whose "invention, design, or arrangement [...] becomes a way of settling an issue in a particular community" and the technological systems that for their very nature are inherently political.[305]

The artefacts that Winner has in mind are bridges, manufacturing machines, tomato harvesters, the atom bomb, and the target of his analysis is the impact of these technologies on society at large and not on a scientific community. However, put into perspective, Winner's analysis can offer useful suggestions also for the examination of my case study, the *Statistical Tables for Biological, Agricultural and Medical Research*, a collection of statistical tables published for the first time in 1938 and co-authored by the statisticians Ronald Fisher and Frank Yates.[306] For over forty years, this collection was used by statisticians and research workers as a computing tool for the application of analysis of variance and experimental design, the statistical methods developed by Ronald Fisher in the 1920s (**chapter 1**).

Like the artefacts that Winner calls political for their "invention, design, or arrangement", also the *Statistical Tables* promoted through their careful planning and the liberal management of their copyright the dissemination of Fisher's statistical methods. Therefore, I will call them a political technology, although the primary impact of this collection of tables was not on society at large, but on the community of its users, that is the statisticians and the research workers interested in the application of statistics to the planning and analysis of field and laboratory experiments.

I will accept Winner's suggestion that the analysis of artefacts, political for their design, require to go beyond their immediate use and investigate "whether a given device might have been designed and built in such a way that it produces a set of consequences logically and temporally *prior* to any of its professed uses".[307] Thus, I am going to examine in detail the structure of the

---

304 See, for instance, the collection of essays *Technologies of Power* (M. T. Allen and G. Hecht (2001)). The table of contents of this book – in which there are contributions ranging from the role of the telephone in the making of the American middle class to operational research in Britain – suggests by itself how varied can be the study of the political dimension of artefacts.
305 L. Winner (1980), p. 123. Besides the paper mentioned here in detail, also in *Autonomous Technologies* (L. Winner (1983)) and *The Whale and the Reactor* (L. Winner (1986)), Winner has examined the political value of technologies and their impact on society.
306 R. A. Fisher and F. Yates (1963).
307 L. Winner (1980), p. 125.

Statistical Tables and the copyright management of the materials included in the collection to support my argument that Fisher and Yates' collection was a value-laden computing tool.

As the potential users of the *Statistical Tables* were both trained mathematicians and research workers with a limited statistical background, the values embedded in the book were perceived and acknowledged differently in relation to the statistical expertise of the two categories. I will argue, in fact, that research workers were unconcerned by mathematical debates over statistical significance or controversies on the interpretation of statistical distributions and just wanted mathematical formulae of easy application, while for the statisticians the preference for a table format or the decision to collaborate in the making of a new table was often a matter of loyalty or an opportunity for crafting alliances.

If my primary interest is for the *Statistical Tables* as a political artefact within the community of its intended users, I cannot disregard to mention that the statistical methods promoted by the book had consequences on society at large. The increasing value attributed by policy makers to Fisherian statistics throughout the twentieth century makes the *Statistical Tables* a political artefact also in a sense closer to Winner's own use and it sets its influence well beyond scientific research. As that is not my main concern, I am just going to suggest a few considerations on this aspect.

The chapter is divided into three sections. In the first one I will contextualize the *Statistical Tables* in the wider scenario of table making in Britain. I am going to deal, in particular, with the British Association Mathematical Tables Committee (from 1948 Royal Society Mathematical Tables Committee) that for almost a century influenced the computation of mathematical tables in the country and in which many statisticians were co-opted. On the making of statistical tables in Britain a special attention deserves Karl Pearson's Biometric Laboratory at University College London, where many relevant tables for statistics were computed since the beginning of the twentieth century. Pearson later edited these tables in book form as *Tables for Statisticians and Biometricians*, the collection that was the closer competitor of Fisher and Yates' book. In Pearson's collection it was reprinted the table of Student's distribution, whose first version was prepared by the brewer William Gosset in Pearson's Biometric Laboratory. I will discuss Student's table in some detail for the influence that Student's test had on the research agenda of Ronald Fisher, and for Fisher's contributions to the 1925 version of the table.

In the second section I am going to present in detail the *Statistical Tables.* I will briefly sketch the main features of analysis of variance and experimental design, the statistical methods that represented the theoretical backbone of the collection, and account the past experiences of the book authors, Ronald Fisher and Frank Yates, as computers. I will then give a detailed account of the book structure, describe its making throughout six editions and its editorial success.

In the final section I will show how the many research workers and statisticians who bought Fisher and Yates' collection as a computing tool were sold at the same time a peculiar vision of statistics. I am going to examine in detail the format chosen by Fisher for the chi-square table and the table of Student's distribution, and explain how the liberal policy adopted by Fisher and Yates for granting permission to reproduce materials in the book was influenced by Fisher's ideas on computing tools in statistics. Following Langdon Winner's paper I will then summarize my arguments on the values embedded in the *Statistical Tables*. In particular, I will argue that the careful planning of the book and the copyright choices of the authors made the *Statistical Tables* a political artefact and I will suggest how the methods promoted by the *Statistical Tables* had far-reaching consequences on society and not only on experimental research.

2.2 *Statistics and table making in Britain before the* Statistical Tables

2.2.1 *The statisticians' role in the British tradition of table making*

A "veritable culture of calculation" emerged in Britain during the mid-Victorian age promoting a commodification of numbers and their mass production. Numerical tables became widespread in scientific circles, government councils and among the lay public. Logarithms, trigonometric functions, multiplication tables were the working tool of the human computer, as well as the faithful laboratory companion of physicists, engineers, statisticians engaged in the solution of complex numerical problems.[308]

In scientific computation the wealth of production was not always synonym of quality or of an efficient use of resources. Mathematical tables published in journals and not in book form were difficult to locate, difficult to keep flat for regular use, difficult to access apart from in a reference library.[309] No list of errata for published books of mathematical tables was systematically available preventing improvements edition after edition.

Higher mathematical functions that could be of use in different scientific applications were often available in publications that little had to share with the discipline of many potential users making the search of the table difficult.[310] One such example is the integral of $exp(-x^2)$, a crucial function for astronomy as well as probability and statistics.

> The most readily accessible tables of the values for the integral of $\exp(-x^2)$ were published in two forms drawn from two different sources. They had first been published (together with their logarithms) in Strassburg in 1799 by Kramp in a book on astronomical refraction, from which they had

[308] A. Warwick (1995), p. 318. Warwick's essay gives a description of table making in Britain since the Victorian age.
[309] M. Croarken and A. Campbell-Kelly (2000), p. 45.
[310] A. Warwick (1995), p. 323.

been reprinted in the *Encyclopaedia Metropolitana* in an article on 'The Theory of Probability'. This article also included a table of the same function in the form more commonly used in the theory of probabilities (that is, multiplied by $2/\sqrt{\pi}$) reprinted from the *Berlin Astronomisches Jahrbuch* for 1834. The latter table was also reprinted, with extensions, in the 'Essay on Probabilities and Life Contingencies' in the *Cabinet Cyclopaedia* and in the article on 'Probability' in the *Encyclopaedia Britannica*.[311]

Lorraine Daston's account of the disciplinary migration of the normal curve ($exp(-x^2)/\sqrt{2\pi}$) during the nineteenth century makes sense of the multiple venues for the publication of the integral of $exp(-x^2)$ mentioned above and helps to understand how concrete was the risk that the table, published under a specialised heading, might be overlooked and recomputed in a different context with unnecessary labour.[312]

To bring order in this chaotic situation the British Association for the Advancement of Science set up in the 1870s a committee on mathematical tables with the task to survey the field and to draft proposals on the tables that could be useful to compute or reprint.[313] Throughout the following century, alternating periods of intense work to others of inactivity, the British Association Committee was a driving force of table making in Britain and many of the best British computers joined its ranks.[314]

Several statisticians contributed in a formal or informal capacity to the committee work (Appendix 2.1). Among them there were Karl Pearson, Ronald Fisher and Frank Yates, the authors of the collections of statistical tables that I am going to examine in this chapter. Karl Pearson is not listed among the official members of the Committee, but in the 1890s he supervised on behalf of the British Association the calculation of frequency distribution curves and related logarithms for use by statisticians and biologists.[315] One such example are the tables of integrals for curve fitting, published in the British Association Report for 1899 and prepared, under Pearson's supervision, by Alice Lee, one of the computers in his Biometric Laboratory at University College.[316]

Unlike Pearson, Ronald Fisher did not undertake any specific job for the British Association Committee, although he was one of its long-term members, having served between 1925 and

---

311 A. Warwick (1995), p. 322.
312 L. Daston (2008).
313 The integral of $exp(-x^2)$ is mentioned as example of function with widespread application also in the 1873 report of the British Association Mathematical Tables Committee prepared by J. W. L. Glashier. (*British Association for the Advancement of Science Report 1873*, p. 2).
314 Detailed accounts of the history of the British Association Mathematical Tables Committee are M. Croarken and M. Campbell-Kelly (2000) and M. Croarken (2003).
315 In M. Campbell-Kelly and M. Croarken (2000) and M. Croarken (2003) there is a comprehensive list of the committee members, but in both cases there is no mention of Karl Pearson's involvement with the British Association. On this point see A. Warwick (1995), p. 326. More details can be found in the "Final Report on calculation of mathematical tables, summer 1948", cited in bibliography as "The BAASMTC now RSMTC" in *MTAC*, 1949, edited by J. C. P. Miller.
316 *British Association for the Advancement of Science Report 1899*, pp. 65-120. Further information on Karl Pearson and computing, including Alice Lee, is in D. A. Grier (2007), pp. 109-112.

1949. In 1928 Fisher was appointed general editor of the British Association table making project – a position that he held until January 1931 – and he housed in his own laboratories, at first at Rothamsted Experimental Station and later at University College London, some of the computing machines (and their computers) that the committee purchased with the legacy of a former member. At Rothamsted, Fisher and his two assistants John Wishart and Joseph Irwin (**chapter 1**), also members of the Mathematical Tables Committee, had available a Brunsviga-Dupla. The Galton Laboratory (**chapter 3**), instead, hosted after 1936 the National Accounting Machine purchased by the Committee, and a few years later also one of its Brunsviga calculators.[317]

Even though Fisher was a long-standing member of the Committee, his relationship with the British Association was not always easy, especially when Leslie J. Comrie was secretary (1929-1937). Comrie, trained in astronomy and with a flair for computing, "seemed happy just making beautiful numbers" and was not really interested in the kind of tables he produced and in their user-friendliness.[318] Fisher, instead, was concerned by the usability of mathematical tables. He strongly supported the idea that also users without calculating machines, like many of the research workers in agriculture and biology that he advised in the application of statistical methods, could benefit from the tables produced by the British Association.[319] He was also an advocate of the most efficient systems of computation and harshly disputed with the other members of the Committee over the best interpolation formulae.[320]

Frank Yates, the junior author of the *Statistical Tables*, joined the British Association Committee only after WWII, when the body was reconstituted under the patronage of the Royal Society. At that stage, Yates was dealing with the heavy calculations involved in the analysis of the experiments conducted at Rothamsted Experimental Station and in several others British institutions, because his department became a general statistical advisory service for agriculture and biology after WWII (**chapter 4**).[321] There was "little enthusiasm" for mathematical tables in

317 For Fisher's appointment as general editor see the minutes of the meeting held on 14th December 1928 (Minute Book Ms. Eng. Min. d 1157, BL, University of Oxford) and for his resignation the minutes of the meeting held on 22nd January 1931 (Minute Book Ms. Eng. Min. d 1158, BL, University of Oxford). The computing machines were purchased with part of the legacy left by Allan Cunningham in 1928. According to the Committee reports, three full-time computers of the British Association followed one another in Fisher's Galton Laboratory from 1936 onwards: F. H. Cleaver (appointed January 1937, resigned May 1938), H. O. Hartley (appointed June 1938, resigned September 1938), R. St. H. Tysser (later Cashen) (appointed October 1938, resigned April 1940).

318 M. Croarken and M. Campbell-Kelly (2000), p. 54. L. J. Comrie (1893-1950) was a key figure in scientific computation in Britain in the first half of the twentieth century (see M. Croarken, 2000). He served as secretary of the British Association Mathematical Tables Committee from 1929 until 1936, when he set up his private computing venture, the Scientific Computing Service.

319 J. Fisher Box (1978), p. 237 and p. 247.

320 Correspondence between R. A. Fisher and A. C. Aitken (April 1931), R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

321 RES (1949), p. 45.

the Royal Society. If Pearson and Fisher had seen the golden age of table making, Yates who served in the period 1948-1957, saw its drastic change and decline due to the availability of digital computers.[322]

Even though Pearson, Fisher and Yates, along with many more statisticians, took an active role in the British tradition of table making for scientific research and contributed with their expertise to the British Association Mathematical Tables Committee, the aims of the Committee and the ones of the statisticians who served in it did not overlap. Making statistical tables was never a priority of the British Association and many of the projects funded by the Committee drifted towards computing for computing's sake. Instead, collections of tables for the application of statistical methods to experimental data, such as the *Tables for Statisticians and Biometricians* or the *Statistical Tables for Biological, Agricultural and Medical Research*, were not the gratification of the computing taste of their authors, but rather a necessary tool in order to transform statistical theory into practice.

For this reason table making for statistics took place where the reduction of experimental data was a daily practice – as in the Biometric Laboratory, the Rothamsted statistics department, the Galton Laboratory –, rather than in the context of general table making projects like the ones supported by the British Association. This is an important point to keep in mind for my argument on the values embedded in the *Statistical Tables*, because it sets the role of computing tools for statistics well beyond mere number crunching, making them key elements in the application of statistical theory to experimental research.

2.2.2 Tables for Statisticians and Biometricians

Karl Pearson worked for most of his life at University College London where he taught and researched in applied mathematics, statistics, computing, eugenics.[323] The department of applied statistics that he created there was the first institutional space overtly devoted to statistics in Britain, but well before the formal opening of his department in 1911, Pearson had started statistical teaching and research at University College and set up a Biometric Laboratory.

---

322 The last secretary of the Committee, the computer pioneer Maurice Wilkes, "felt his job was to bring the Mathematical Tables Committee to an end" (M. Croarken (2003), p. 259). The Committee was eventually dissolved in 1965.

323 First hand accounts of Karl Pearson's scientific career written by his co-workers can be found in the *Obituary Notices of Fellows of the Royal Society* (G. Udny Yule and L. N. G. Filon (1936)) and in *Biometrika* (E. S. Pearson (1936, 1938)). The historians T. M. Porter and M. E. Magnello has written extensively on Pearson (T. M. Porter (1986, pp. 270-314), (2006); M. E. Magnello (1999a, 1999b)). On the computing activity at University College under K. Pearson see D. A. Grier (2007), pp. 102-174.

Founded in 1903 with a grant of the Worshipful Company of Drapers, Karl Pearson's Biometric Laboratory trained postgraduate students in the application of statistical methods, supported researchers in the solution of biometric problems and did a substantial computing work preparing mathematical tables for statistics. The computing effort never stopped under Pearson and "by 1906, Pearson could report that the group had mastered the art of mathematical table making".[324]

Most of the tables computed there were published in the journal *Biometrika,* co-founded by Pearson in 1901 to provide statistical tools for dealing with biological problems. One of the aims of the journal was to publish tables that could diminish "the labour of statistical arithmetic" for biometry.[325] As mentioned before, however, mathematical tables published in journals were more difficult to locate and handle than tables in book form and, thus, since the beginning, Pearson planned to reprint the tables in *Biometrika* as an autonomous collection.

The *Tables for Statisticians and Biometricians*, appeared in 1914, edited by Pearson, and represented the first systematic collection of statistical tables published in Britain.[326] In the book there were tables already printed in *Biometrika*, in the *Draper's Company Research Memoirs* (another publication of the Biometric Laboratory), in the *Transactions of the British Association Report* (1899) and original materials.

The array of tables in the book was wide. There were tables of the normal distribution, tables for fitting curves and calculating correlation coefficients and probable errors, tables for testing the goodness of fit, the table of Student's distribution for making tests of significance with small samples, standard mathematical tables such as powers of natural numbers. The proportion with which these groups of tables were included in the collection was suggestive of the statistical methods endorsed by Pearson.

> It was the age of correlation and curve fitting. In *Tables for Statisticians and Biometricians* [...] 37 per cent of the tabular matter was concerned with curve fitting, and a further 18 per cent with various forms of correlation. The normal and Poisson distributions occupied 17 per cent, χ^2, and "Student's" z 5 per cent, the remaining 23 per cent being devoted to tables of basic mathematical functions and miscellaneous statistical tables.[327]

In the introduction of the book a brief description of each table, often taken from the preface of the original journal paper, was given, but this was not intended "to replace actual instruction in

---

[324] D. A. Grier (2007), p. 117.
[325] K. Pearson (ed.) (1914), p. *vi*.
[326] The *Tables for Statisticians and Biometricians* issued by Pearson had a second edition in 1924 and a third in 1930-31. The new editions came to include tables published in *Biometrika* up to that date and in 1930-31 a second volume was added to the collection. Ronald Fisher wrote for *Nature* a critical review of the edition published in 1930. He accused Pearson to take a partial or personal tone dealing with controversial topics, to deliberately ignore the idea of degrees of freedom, to have similar materials scattered in different parts of the two volumes and to indulge in the use of special terminology adopted only in the Biometric Laboratory (R. A. Fisher, 1933).
[327] F. Yates (1951), pp. 19-20.

the use of the tables such as is given in a statistical laboratory".[328] The book was thus aimed at a public of skilled users and it became especially popular among human computers and statisticians. To these users the extended edition in two volumes was presented as "a mine of numerical information" even a few decades later.[329] To Pearson's regret the *Tables for Statisticians and Biometricians* were not a self-consistent collection and *Barrow's Tables* of squares, cubes, square roots, cube roots, reciprocals and a set of tables of trigonometric functions were suggested as a necessary complement to the collection.[330]

Pearson's attitude towards his mathematical tables was ambivalent. As pointed out by his biographer, Theodore Porter, "Pearson did not believe in textbooks" and never wrote one.[331] However, Pearson's books of tables – not only the collection already mentioned, but also the series of *Tracts for Computers*, mainly booklets of mathematical tables issued after WWI –, although aimed at a public of skilled users, could not refrain, from a certain didactic approach intended to explain the fundamental aspects of Pearson's statistic.

> He introduced these books with sample problems and guidelines for reducing them to a suitable form and finding the appropriate entry in his tables. The introduction conveyed the basics of statistical reasoning as he understood it. In most cases this meant fitting a density curve or surface to the data, then perhaps determining whether a point or set of data is consistent with the distribution.[332]

The same ambiguity can be found in Pearson's attitude towards the economic side of table making. Pearson's appreciation as user went to "landmarks of computing science", but as table maker he could not disregard the requests of the market in terms of price and usability.[333] He was disdainful of cheap books of mathematical tables and blamed his co-worker Major Greenwood for the appreciation of Barlow's cubes and squares, a low-price collection, but on the other hand he advised to purchase this same book as a complement of his *Tables for Statisticians and Biometricians* and apologised with the users of his own collection for its high price. The first edition of Pearson's book, in fact, was quite costly (fifteen shillings) despite the revenues offered by the previous publication of the tables in *Biometrika* and despite the support given by the Worshipful Company of Drapers to the Biometric Laboratory.[334]

Pearson constantly stressed the financial problems linked to the making of the book, especially in relation to the computation of the tables and their printing, and ascribed to these causes the

[328] K. Pearson (1914), pp. *vi-vii*.
[329] A. Fletcher et al. (1962), Vol. I, p. 16.
[330] K. Pearson (1914), p. *v*.
[331] T. M. Porter (2006), p. 262. In contrast, R. A. Fisher largely entrusted to his own textbooks, *Statistical Methods for Research Workers* and the *Design of Experiments*, the dissemination of his statistical methods.
[332] T. M. Porter (2006), p. 262.
[333] T. M. Porter (2006), pp. 304-305.
[334] The tables in Pearson's book were printed from the same mould prepared for the journal. Being the typesetting of a mathematical table costly, this represented a relevant economy in the making of the *Tables for Statisticians and Biometricians*.

delay in the publication, announced since 1901, and the necessity to omit from the collection some tables of higher mathematical functions initially planned.[335]

The tension between aspirations and opportunities as table maker – one of the many tensions that crossed Pearson's life – helps to explain the structure of his *Tables for Statisticians and Biometricians*. If, on the one hand, the book was built with an encyclopaedic ideal in mind, including even the table of Student's distribution peripheral to Pearson's statistical agenda, on the other the intellectual property on the collection was jealously enforced as a way to support the revenues offered by the publication of the tables in *Biometrika*. In the preface of the book, in fact, Pearson stated clearly his intention to oppose any plagiaristic attempt or any request of reproduction in order that *Biometrika* would not be deprived "of such increased circulation as it obtained from being the sole *locus* of these Tables".[336] This attitude characterized also the other mathematical tables computed in Pearson's laboratory and edited by him and went so far as to hinder the work of the British Association Mathematical Tables Committee:

> We have been in correspondence with professor Karl Pearson, and we find that it is unlikely either that he will make substantial changes in Tracts for Computers No. 5 when a second edition is required or that a formal proposal for co-operation between the British Association and the Laboratory of Applied Statistics in the production of a volume of interpolation coefficients would be accepted. The issue of copyright was not settled in the letters exchanged. Professor Pearson evidently regards the legal question as doubtful, but he would resent bitterly any independent use of the material in the Tract. We recommend that the B. A. [British Association] volume should contain a critical account of the Everett tables that are in existence, but that the B. A. should not at present publish fresh Everett tables either in the volume under preparation or as a supplement.[337]

To this choice about the copyright of the tables and to the proprietary idea of computing tools that Karl Pearson supported, it is largely linked the history of the *Statistical Tables*. *Biometrika* was the leading journal for the publication of statistical work in Britain and no other collection of tables could be issued without infringing Pearson's copyright or facing the expensive and time-consuming task of computing new tables from scratch. Due to this constraint the *Statistical Tables* of Ronald Fisher and Frank Yates grew slowly throughout the years and relied on the previous experience and efforts in computation of the authors, on their opportunity to have assistant staff to employ for making calculations and on the help of their colleagues ready to give suggestions and lend tables to the collection. On the other hand, Pearson's attitude towards the reproduction of his tables indirectly contributed to the diffusion of the alternative format adopted by Ronald Fisher and Frank Yates, who, on the contrary, followed a liberal policy in granting permission to reproduce material from their book.

---

[335] K. Pearson (1914), p. *vi*.
[336] K. Pearson (1914), p. *vi*.
[337] British Association Mathematical Tables Committee, meeting 19th November 1930, BL, University of Oxford, Minute Book Ms. Eng. Min. d 1157.

Before describing in detail the structure of Fisher and Yates' collection I want to examine the making of the table of Student's distribution, prepared by the chemist William Gosset for dealing with the small experimental samples arising in his practice as brewer (**chapter 1**). The first version of the table was published in *Biometrika* and included in the *Tables for Statisticians and Biometricians*, but Pearson was never really interested in the small samples that could be tackled using Student's table, because "his [Pearson] primary concern was faithfulness to his abundant empirical data rather than causal inference".[338] In contrast Ronald Fisher's statistical methods addressed precisely small samples problems and Fisher contributed to the making of a new version of the table of Student's distribution suggesting the variable transformation, from z to t, which is today canonical. A closer look at the table prepared by Gosset offers thus the opportunity for a better understanding of the contrasting statistical ideas embedded in Karl Pearson's and Ronald Fisher's collections.

2.2.3 *The table of Student's distribution*

The computation of the table of Student's distribution was a one-man initiative stimulated by an industrial context. The author, William Gosset, was an employee of the Guinness brewery in Dublin. He was one of the university graduates that the company had hired at the turn of the twentieth century in order to transform the art of brewing in a scientific activity.

At Guinness Gosset "was mixed up with a lot of large scale experiments partly agriculture but chiefly in an Experimental Brewery" where analysis of malt and hops and tasting of the beer took "a day to each unit of the experiment, thus limiting the numbers".[339] The data accumulated in the experimental brewery were not easy to interpret because "the variation was high and the observations were few" and Gosset turned for help at first to the standard textbooks on the theory of errors and then to Karl Pearson, as the leading authority in statistics.[340]

In 1906-1907 he spent a few months in Pearson's Biometric Laboratory, working on the small samples problems that concerned his experimental work at Guinness. However, the statistical methods devised by Pearson's laboratory were unsuitable for solving Gosset's problems because their validity relied on the availability of large masses of data, while the brewer's goal was to draw inferences when only a few results – less than ten, often only four or five – were available.

> I find out the P. E. [probable error] of a certain laboratory analysis from n analyses of the same sample. This gives me a value of the P. E. which itself has a P. E. of P. E./$\sqrt{(2n)}$. I now have another sample analysed and wish to assign limits within which it is a given probability that the truth must lie. E.g. if n

[338] T. M. Porter (2006), p. 253.
[339] W. S. Gosset (1962), Letter No. 1, 15th September 1915.
[340] J. Fisher Box (1987), p. 48.

were infinite, I could say 'it is 10:1 that the truth lies within 2.6 of the result of the analysis'. As however *n* is finite and in some cases not very large, it is clear that I must enlarge my limits, but I do not know by how much.[341]

Despite his limited mathematical training – "[m]y mathematics stopped at Maths. Mods. [Mathematical Moderations] at Oxford" – Gosset was able to discuss the large/small sample boundary and to present an alternative test to employ with small samples in order to judge "whether a series of experiments, however short, have given a result which conforms to any required standard of accuracy or whether it is necessary to continue the investigation".[342] Although peripheral to Pearson's research programme, Gosset's work was published in *Biometrika* in 1908, under the pen name of Student.

The application of Gosset's test required alternative tables of the probability integral to be used with small samples and Gosset provided his readers in *Biometrika* with them.[343] For the 1908 paper he computed the values of the Student's distribution for number of observations as small as four and as large as ten. The quantity tabulated, however, was not t as in the formulation now standard of Student's test, but z "which is obtained by dividing the distance between the mean of the sample and the mean of the population by the standard deviation of the sample".[344]

A second extended table of the Student's distribution was published, again in *Biometrika*, in 1917. The new table was not computed directly by Gosset, but by another Guinness' employee, W. L. Bowie, who considered samples of growing dimension from two up to thirty.[345] It is no accident that this second table was prepared in the company's office, because at that stage Gosset was already in charge of the statistical work at Guinness and his test had been integrated in the working routine of the company.[346]

[341] E. S. Pearson (1939), p. 216. Assuming to have a set of experimental results (the sample) distributed according to a certain statistical population (normal, binomial, poissonian etc.) two fundamental values for the examination of the population are the mean and the standard deviation (the square root of the variation (variance) in the population). The probable error is a kindred concept of the standard deviation. It was very popular among the nineteenth century error theorists. Roughly the probable error is 0.67 times the standard deviation. The probability integral evaluates the probability that an experimental result fall within a certain distance – usually measured using the standard deviation as unit – from the mean. For the normal curve in one standard deviation above and below the mean there are about sixty-eight per cent of the data and within three standard deviations over ninety-nine per cent.

[342] The first quotation is taken from a letter of Gosset addressed to R. A. Fisher (W. S. Gosset (1962), Letter No. 2, 15th December 1918), while the second is in Student (1908), p. 25.

[343] Student (1908), p. 19. The aim of Gosset's 1908 paper was "to determine the point at which we may use the tables of the [Gaussian] probability integral in judging of the significance of the mean of a series of experiments and to furnish alternative tables for use when the number of experiments is too few." (Student (1908), p. 2)

[344] Student (1908), p. 2. The variable z here mentioned in relation to Student's table should not be confused with the distribution of z and the variance ratio, which is mentioned later in relation to the *Statistical Tables*. Student's test is a test of statistical significance that compares the means of two small samples to ascertain whether they are drawn from the same population.

[345] Student (1917), pp. 416-417.

[346] J. Fisher Box (1978), p. 50.

As mentioned before, Gosset was not a trained mathematician and despite his willingness and effort in computation, he did not stand up as an accurate computer. Later both the tables prepared for *Biometrika* "were found to be perfectly rotten". He frankly admitted why:

> The fact is that I was even more ignorant when I made the first table than I am now and thought I was going to be accurate to 4 places by taking 5 in the working! And the second was of course constructed on the same lines though not by me. I ought to have checked it myself, but must have been pretty casual about it. [347]

The making of the tables of Student's distribution as an essential companion for Gosset's theoretical achievement in the statistical analysis of small samples can be seen as a further evidence of my claim that mathematical tables were crucial in the application of statistical theory during the twentieth century. Moreover, as I am going to describe below, the new version of the table prepared in the 1920s proves the shifting alliances of Gosset, from Karl Pearson's to Ronald Fisher's statistical methods.

Ronald Fisher, the senior author of the *Statistical Tables*, was interested as much as Gosset in methods for dealing with small samples and since the 1910s began a scientific correspondence on these topics with the brewer. In September 1922 Gosset sent to Fisher a copy of the tables of the Student's distribution, as he was "the only man that's ever likely to use them".[348] Shortly afterwards Fisher proposed to the brewer a different formula for computing his integral and a few months later Gosset began to work on a new version of the table. In contrast to the first version the new one had as entry not the number of observations (n), but the number of degrees of freedom (n-1) – a concept that Fisher had introduced for the first time in 1922 as a correction of Pearson's chi-square test and that Pearson constantly objected to – and the quantity tabulated was now the variable $t = z/\sqrt{(n-1)}$.[349]

Gosset did the bulk of the computing work for the new table in his spare time. He worked on the table for some years with a calculating machine taken home from the office during the winter months, when computing machines were in less demand at the brewery.[350] Fisher contributed to the making of the table providing the mathematical formulae, a few calculations and constant

---

[347] W. S. Gosset (1962), Letter No. 36, 23rd November 1923.
[348] W. S. Gosset (1962), Letter No. 11, 21st September 1922.
[349] More details on the degrees of freedom controversy between Fisher and Pearson are given when I deal with the table of the chi-square distribution. Fisher pointed out the problem related to the degrees of freedom in Student's test since 1912, when he sent to Gosset a rigorous mathematical proof of his distribution. Karl Pearson, consulted by Gosset, replied: "Whether the proper formula for the S. D. [standard deviation] is $[S(x-m)^2/n]^{1/2}$ or $[S(x-m)^2/(n-1)]^{1/2}$ seems to be of very little practical importance, because only naughty brewers take n so small that the difference is not of the order of the probable error of the summation!" (Letter from K. Pearson to W. S. Gosset, 17th September 1912, quoted in E. S. Pearson (1990), p. 48). The controversy over the degrees of freedom is just an example of the scientific contrasts between R. A. Fisher and K. Pearson.
[350] "The tabulating season having commenced I took a calculating machine home on Saturday and began work last night on it" (W. S. Gosset (1962), Letter No. 34, 16th October 1923).

advice. His assistant of the time, Winifred Mackenzie, checked also part of the computations, because again Gosset did not stand up as a professional computer.[351]

The new version of the table, once completed, conserved little of the statistical training Gosset had originally received in Karl Pearson's laboratory, while evident was Fisher's intervention in relation to the degrees of freedom and the introduction of the variable t instead of z. Fisher was also to write an introductory note on the formulae employed in the calculation of the tables and on the uses of Student's test.

Loyal to Karl Pearson, Gosset offered him the opportunity to publish again the table in *Biometrika* – "I've come to the conclusion that I must offer it to K. P. first" –, but no final agreement was reached, possibly due to copyright issues with Fisher who wanted the right to include the table in a future book.[352] According to Fisher "after long delay, [Pearson] capriciously refused it [the table], publishing in the meanwhile under his own name a proof of one of our results."[353]

Eventually the table of Student's distribution recomputed using the variable t was printed in 1925 in the Italian journal of statistics *Metron*, alongside two contributions written by Fisher, an explanation of the use of Student's distribution and the expansion of Student's integral used for computing the table.[354] In the same year in which the *Metron* papers appeared, Fisher's own version of Student's table was published in *Statistical Methods for Research Workers*.[355] Fisher's table had a different layout from the one originally devised by Gosset. Fisher used the rows of the table for the degrees of freedom, while tabulated in the columns, not the values of t, but the associated probability values for certain fixed intervals. The layout of Student's table published by Fisher was comparable to the one that he had adopted for the table of the chi-square distribution printed in the same book and that I will discuss later in detail as an example of the political dimension of the *Statistical Tables*.

---

351 "Table III was calculated from Mr Fisher's formulae and I have to thank Miss Mackenzie, M.Sc., of the Rothamsted Statistical Laboratory for kindly checking this part of the work." (Student (1925), reprinted in E. S. Pearson and J. Wishart (1942), p. 117). Despite the support of Fisher and his assistant, mistakes were discovered also in the latest version of the table (W. S. Gosset (1962), Letter No. 86, 5th July 1927 and 23rd July 1927).

352 For the publication of the table in *Biometrika* see W. S. Gosset (1962), Letter No. 35, 2nd November 1923. On the copyright issues see W. S. Gosset (1962), Letter No. 38, 20th December 1923. Despite his loyalty to K. Pearson, Gosset did not refrain from writing to Fisher: "In most of your differences with Pearson I am altogether on your side and in some cases I have agreed to differ from him long ago" (W. S. Gosset (1962), Letter No. 5, 3rd April 1922). Moreover Gosset's statistical assistants at Guinness, E. Somerfield and A. L. Murray, had their training at Rothamsted with Fisher and not in the Biometric Laboratory with Pearson.

353 Letter from R. A. Fisher to G. S. Phillpotts, 7th July 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide. G. S. Phillpotts was the brother in law of Gosset.

354 Student (1925), R. A. Fisher (1925b), R. A. Fisher (1925c). Fisher was acquainted with the editor of *Metron*, the Italian statistician Corrado Gini. Fisher and Gini met for the first time in August 1924 at the International Mathematical Congress held in Toronto (F. Cassata (2006), p. 145).

355 A detailed description of the book is in A. W. F. Edwards (2005a).

The making and remaking of the table of Student's distribution proves that statistical tables are not neutral commodities and that different research programs, shifting alliances and new power relations within the community of statisticians can influence their format.[356]

2.3 *The* Statistical Tables for Biological, Agricultural and Medical Research

2.3.1 *Ronald Fisher's statistical methods in experimental research*

The theoretical backbone of the *Statistical Tables for Biological, Agricultural and Medical Research* was represented by analysis of variance and experimental design, the statistical methods developed by Ronald Fisher during the 1920s for the examination of the agricultural experiments conducted at Rothamsted Experimental Station (**chapter 1**).

The analysis of variance is a tool for analysing experimental data splitting the causes of variation in independent components. With this method it is possible to work out the error associated with the data and whether the discrepancies in different trials are due to chance or not.

A precondition for the examination of experimental data with the analysis of variance is the careful planning of experiments. Fisher singled out four main principles in arranging experiments: replication, randomisation, use of factorial experiments, confounding.

Analysis of variance and experimental design can be used in a wide array of experimental disciplines. The ones explicitly mentioned in the title of the book – biology, agriculture and medicine – are just some of the fields in which Fisher's statistical method were promptly adopted.[357]

---

356 On the history of the table related to Student's distribution, the choice of the variable transformation from z to t, and Fisher's own version of Student's table has written also the economist S. T. Ziliak (S. T. Ziliak (2008), published on the author's website; S. T. Ziliak and D. N. McCloskey (2008), pp. 227-233). I agree with Ziliak when he claims that in Fisher's version of the table there were embedded Fisher's ideas on statistical significance. However, unlike Ziliak, I do not find in published and archival sources any evidence that Fisher tried to write out of the history of statistics W. S. Gosset. In *Statistical Methods for Research Workers*, Fisher explicitly attributes the paternity of the Student's distribution to the chemist (R. A. Fisher (1925a, first edition), p. 17; R. A. Fisher (1946, tenth edition), p. 22); the same attitude is again in Gosset's obituary written by Fisher for the *Annals of Eugenics* (R. A. Fisher (1939a), pp. 3-4). Ziliak's suggestion that in *Metron* the authorship of Gosset's table was unclear and that the table was included in one of the two contributions published by Fisher in the same issue does not find any confirmation in the table of contents of the journal, where the three papers are clearly separated. Moreover, Ziliak refers constantly to Student's copyright on his own table, but this is not correct. The copyright on Student's table published in *Biometrika* did not belong to Gosset, but to K. Pearson. It was Pearson's policy that prevented the diffusion of the table of Student's distribution as published in *Biometrika*, while Fisher's liberal attitude towards his own table promoted his statistical ideas. The table recomputed by Fisher for *Statistical Methods for Research Workers* did not require any permission from Gosset or *Biometrika* because it was, technically, a new table. It had a different format and Fisher was, due to a special agreement with his publisher, the real owner of the copyright.

357 For the adoption of analysis of variance and experimental design in agriculture and biology see chapter 1. For statistics in medicine an interesting reading is H. M. Marks (2000). Marks focuses on the period from the 1950s onwards and argues that in the previous two decades "statistical analysis in clinical medicine [...] remained the rare

Ronald Fisher presented analysis of variance and experimental design in his textbooks, *Statistical Methods for Research Workers* and the *Design of Experiments*. In *Statistical Methods*, Student's distribution was explained in detail – even with the same example given in Student's original paper – and it was discussed in what sense the analysis of variance was an extension of Student's test.[358] The affinity between Fisher and Gosset's work went well beyond mathematical formalism: it was rooted in the experimental context in which both statistical investigations were conceived. Gosset's research was driven by his work as chemist in the Guinness brewery, while the development of Fisher's statistical methods was prompted by the problems that he met as statistical consultant of the research workers at Rothamsted Experimental Station.

Therefore the disciplines – biology, agriculture and medicine – explicitly mentioned in the title of Fisher and Yates' collection were not meant as theoretical disciplines, but as experimental practices. Interested in Fisher's statistical methods were the research workers that planned experiments in the field or in the laboratory and had to examine the resulting data and the statisticians who advised them in these tasks. The aim of the *Statistical Tables* was to reprint together the tables already published in *Statistical Methods for Research Workers* plus "some others in common laboratory use", to simplify the practical application of Fisher's statistical methods in experimental research.[359]

The making of the *Statistical Tables* as a political artefact should be understood in relation to this aim. In particular, Ronald Fisher wanted to promote his statistical methods among research workers and indeed the *Statistical Tables* stressed aspects as usability and practical utility for solving laboratory problems in order to bound these users to the adoption of analysis of variance and experimental design. The research workers were rather inclined to accept cookbook recipes for statistics without much questioning (**chapter 3**), as for the guidelines for significance testing that I am going to discuss in relation to the chi-square table and the table of Student's distribution.[360] *Statistical Methods for Research Workers*, on purpose, presented no mathematical

---

exception, not the rule" (H. M. Marks (2000), p.136). It is true, however, that already in 1938 the physician Donald Mainland published his first textbook on the application of statistics to medicine, *The Treatment of Clinical and Laboratory Data: An Introduction to Statistical Ideas and Methods for Medical and Dental Workers* (D. Mainland (1938)). The book was printed by the same publisher of the *Statistical Tables* and Fisher acted as referee for Oliver and Boyd in the publication of Mainland's book (Letter from Oliver and Boyd to R. A. Fisher, 19th March 1937, Oliver and Boyd Collection, NLS).

358 R. A. Fisher (1946), pp. 121-122.

359 Letter from R. A. Fisher to Oliver and Boyd, 20th May 1937, Oliver and Boyd Collection, NLS. In the table of contents submitted by Fisher to Oliver and Boyd are mentioned: decimal and natural logarithms, decimal antilogarithms, squares, square roots, reciprocals, sines, tangents, factorials, logarithmic factorials, random numbers, the table of the normal distribution, the table of the Student's distribution, the table of z and the variance ratio, the chi-square table, the table of the correlation coefficient, a table of constants, schemes for Latin squares and completely orthogonal squares.

360 F. Yates (1951), G. Gigerenzer et al. (1989), p. 97. Fisher's textbooks consistently emphasised significance testing (F. Yates (1951), pp. 32-33). For a philosophical approach to significance testing, including also Fisher's attitude on

proof, a fact that greatly shocked Fisher's fellow statisticians, but did not impair the editorial success of the book that sold in Fisher's lifetime thirty-six thousand copies (English edition).[361]

Research workers were not interested in the subtleties of table making as well. The dislike of biologists for interpolation was publicly stated in a review of the *Statistical Tables* by the statistician and biologist Chester I. Bliss – "since few biologists take readily to interpolation, several tables could be expanded to advantage" – and the computing equipment on which research workers could rely was basic if not poor, often limited to a slide rule.[362] Fisher and Yates as consultants of research workers were already aware of these constraints and carefully considered them in the making of their collection.[363]

2.3.2 *Ronald Fisher and Frank Yates as computers*

The leading figure in the project of the *Statistical Tables* was Ronald Fisher, who in the 1930s was already an established statistician and geneticist (**chapter 1**, **chapter 3**). His career in research had begun in 1919 at Rothamsted Experimental Station, where he had founded the local statistics department and developed analysis of variance and experimental design.

Since the early 1920s the staff of the Rothamsted statistics department managed by Fisher included a small group of human computers and Fisher always praised the computing skills of his assistants and of the visiting workers that came to Rothamsted for learning his statistical methods.[364]

Fisher himself was a skilled computer.[365] Even before his engagement with the British Association in 1925, he had already prepared the extensive set of tables published in *Statistical Methods for Research Workers*. He computed the tables from scratch with his Millionaire calculator and the help of just one assistant, Winifred Mackenzie. Fisher conceived statistics and computing

---

this issue, see I. Hacking (1976), chapter VI. See also the discussion in the following section on the five per cent threshold for statistical significance.

361 F. Yates (1951), p. 30. For the copies of *Statistical Methods for Research Workers* sold in Fisher's lifetime see Letter from Oliver and Boyd to F. Yates, 10th June 1963, Oliver and Boyd Collection, NLS.

362 C. I. Bliss (1943), pp. 346-347.

363 See for instance the discussion on the availability of computing equipment at Rothamsted Experimental Station before the opening of the statistics department in chapter 1 (1.5.1.c).

364 The assistant statisticians that Fisher had both at Rothamsted (W. A. Mackenzie, J. Wishart, J. O. Irwin, F. Yates) and at the Galton Laboratory (W. L. Stevens) were skilled computers and, apart from Mackenzie, they were all members at some point in their career of the British Association Mathematical Tables Committee. Fisher praised also the voluntary workers for their computing skills. For instance, on A. L. Murray, Gosset's assistant at Guinness in Dublin, he wrote "Murray strikes me as a remarkably quick computer." (W. S. Gosset (1962), Letter No. 128, 4th February 1931). Fisher asked Murray to test alternative interpolation methods in his controversy with the other members of the British Association Mathematical Tables Committee mentioned above. (Correspondence between R. A. Fisher and A. C. Aitken (April 1931), R. A. Fisher Papers (digitized), BSL, The University of Adelaide)

365 F. Yates and K. Mather (1963), p. 93. On Fisher and computation see also J. Fisher Box (1978), pp. 242; 245-246.

as two interrelated activities, and promoted the continuity between statistical methods and computing work not only in his own research, but also in the departments he directed.[366]

When Fisher left Rothamsted in 1933 he succeeded Karl Pearson as head of the Galton Laboratory at University College London.[367] Despite the change, throughout the 1930s Fisher remained strongly associated with Rothamsted as honorary consultant and the *Statistical Tables* were the outcome of the collaboration between the computing staff in the statistics department at Rothamsted and in the Galton Laboratory. Fisher relinquished his role of honorary consultant at Rothamsted only in the 1940s, when he was appointed to the chair of genetics at Cambridge University.[368]

The computing tradition of the Rothamsted statistics department continued with Frank Yates (**chapter 4**). Yates fully embraced Fisher's methods and they became lifelong friends: the *Statistical Tables* are the result of their fruitful scientific collaboration and of their personal friendship as well. Before his appointment at Rothamsted, Yates had been research officer and mathematical advisor to the Gold Coast Survey where he had especially supervised the computing activity of the department. The first experience of Yates in table making can be traced back to this period as he was in charge of a booklet of tables printed in 1929 to supplement the standard collections of mathematical tables used by the surveyors of the department.[369]

Yates was even more interested than Fisher in advanced computing technologies and the Rothamsted statistics department in the 1940s began to acquire punched-card equipment. As mentioned before, Yates took part in the Royal Society Mathematical Tables Committee after WWII and served in it for about a decade. His attendance in the Committee was justified also by his pioneering role in the computerization of scientific calculation in Britain. In the 1950s the Rothamsted statistics department, due to Yates' interest for computing technologies, was at the forefront in Britain in the acquisition of digital computers, a technology from which the later editions of the *Statistical Tables* benefited.[370]

---

366 F. Yates (1966), p. 233.

367 When Pearson retired from University College London his department was split in two sections. Fisher was appointed Galton professor of eugenics and head of the department of eugenics, while E. S. Pearson, K. Pearson's son, became reader in statistics and head of the department of statistics.

368 The department of genetics in Cambridge had almost no computing facility and thus, from the third edition of the *Statistical Tables* onwards, it was the staff of the Rothamsted statistics department that mainly contributed to the new computing work required by the book. The appointment at Cambridge coincided with Fisher's move away from Harpenden, the village in which Rothamsted Experimental Station is located and where he had lived since 1919.

369 Letter from F. Yates to R. A. Fisher, 23rd April 1931, RR Library and Archive, STF 31 (old catalogue reference); Gold Coast Survey (1929).

370 Yates was also one of the founding fathers of the British Computer Society and its president in 1960-61.

Fisher instead "never much concerned himself with electronic computers", although he was ready to request the help of their users when necessity arose.[371] For instance, since 1949 he contacted Maurice Wilkes, the head of the Mathematical Laboratory at Cambridge University, for the computation of a differential equation with the local mainframe, the EDSAC (**chapter 4**).[372] The calculation, programmed by Wilkes' co-worker, David J. Wheeler, was completed in 1950 and in the same year the table with the solutions of the differential equation appeared in a paper on gene frequencies authored by Fisher.[373] In 1953 Fisher asked again and again received Wilkes' support in computing the solutions of a differential equation and after 1954, when the Rothamsted statistics department acquired a mainframe of its own, Fisher resorted to Frank Yates for the solution of similar problems.[374]

2.3.3 *Constructing the* Statistical Tables

Ronald Fisher and Frank Yates did not plan the *Statistical Tables* as a comprehensive collection as Pearson's *Tables for Statisticians and Biometricians*, nor they expected that all their readers had been trained in a statistical laboratory. They conceived instead a "small book", designed and realised with Fisher's statistical methods in mind.[375] In fact, the first edition of *Statistical Tables* was a thin book of ninety pages (on the other hand Pearson's collection had over two hundred pages) with thirty-four tables and it had a rather long introduction (twenty-three pages) at the beginning that explained how the tables had been constructed and how they should be used. There were also a few numerical examples for guiding the reader. General rules for interpolation were stated and many tables were provided with specials tips for rapid interpolation in order to speed up the calculations.

Through many cross-references, the *Statistical Tables* were presented as a useful supplement for the readers of Fisher's textbooks – *Statistical Methods for Research Workers* and the *Design of Experiments* – strengthening the links between statistical theory and its application through the use of the computing instrument. For instance in presenting the possible procedures for randomization to the readers of *The Design of Experiments* Fisher suggested: "To save the labour of card shuffling use is often made of printed tables of random sampling numbers [...]. The first

[371] F. Yates (1966), p. 233.
[372] Letter from R. A. Fisher to M. V. Wilkes, 16th February 1949, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.
[373] R. A. Fisher (1950), pp. 354-355. See also in the R. A. Fisher Papers (digitized), BSL, The University of Adelaide: Letter from R. A. Fisher to M. V. Wilkes, 27th April 1950; Letter from R. A. Fisher to M. V. Wilkes, 26th September 1950; Letter from R. A. Fisher to M. V. Wilkes, 2nd October 1950.
[374] Letter from R. A. Fisher to M. V. Wilkes, 27th November 1953, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.
[375] Letter from R. A. Fisher to Oliver and Boyd, 20th May 1937, Oliver and Boyd Collection, NLS.

such table was published by Tippet [a British statistician who had been a visiting worker in Fisher's department at Rothamsted]; another is available in *Statistical Tables*."[376]

The complementary role of the collection of tables in relation to Fisher's statistical books was further emphasised by the fact that the same publisher, the Edinburgh company Oliver and Boyd, printed all the books and in their internal pages *Statistical Methods*, the *Design of Experiments* and *Statistical Tables* mutually advertised each other.

The opening tables of the collection (normal distribution, Student's distribution, chi-square distribution, distribution of z and the variance ratio, correlation coefficient) were indispensable tools for the application of the analysis of variance to experimental data. They were a key element for making tests of significance, i.e. deciding whether a set of experimental data confirmed or not a given hypothesis, and they were of general use in statistics applied to experimental research. This was the set of tables enlarged and reprinted from Fisher's *Statistical Methods for Research Workers,* a publication already well known in the 1930s.

In relation to experimental design the *Statistical Tables* offered to the reader non numerical tables of both Latin squares and randomised blocks, two types of arrangements at first developed by Fisher and Yates for agricultural trials, but of general applicability in experimental research. The table of random numbers included in the collection since the first edition was relevant for experimental design, because the chance allocation of treatments was a crucial element among the principles stated by Fisher in arranging experiments.

The *Statistical Tables* offered to the reader also standard tables of general mathematical functions like logarithms, sines, cosines, squares, square roots, reciprocals for facilitating the calculations linked to the analysis of variance without the necessity to turn to other collections, such as the *Barlow's Tables* recommended by Pearson. In the book there were also tables that offered an answer to specific research problems that could be addressed through Fisher's methods. For instance, the tables of probits were useful for dealing with problems in toxicology and tests of psychological preference could be examined through tables of ordinal and ranked data.

The prospective publics of the *Statistical Tables* were both the statisticians who acted as consultants of research workers and the research workers themselves, who usually could rely only on modest computing tools.[377] Therefore, the tables had been planned and computed in such a

[376] R. A. Fisher (1947), p. 50.

[377] The *Statistical Tables* were also a teaching aid in the classes on analysis of variance and experimental design, as evident in this letter addressed to Yates by the biologist C. I. Bliss: "My class urgently needs copy of the new edition of statistical tables. When will they be available […]?" (Letter from C. I. Bliss to F. Yates, 28th January 1948, RR Library and Archive, STF 31 (old catalogue reference)).

way that also a user supplied only with a slide rule or a poor desk calculator could work through his or her data using the tables. Not only Fisher, but also Yates emphasized this aspect: "I am definitely keen to have the variance ratio as well as *z*, because this ratio would be easier to remember than *z*, and because it can be computed directly from a slide rule. The reader could be explicitly informed that the *z* table should be used for purposes of interpolation."[378] The possibility to interpolate the values easily was a further component that made Fisher's and Yates's book effective and contributed to its popularity.

Moreover, despite the high cost for printing the first edition, the *Statistical Tables* were sold since 1938 at a price affordable even for research workers like biologists, agronomists and physicians that devoted only a small amount of their research budget to computing tools. The low price of the collection – twelve shillings and six pence for the first edition – was not a mere accident, but a careful marketing strategy and the authors suggested and warmly supported this choice, even though it meant a diminution of their royalties.[379] In conversation with the publisher about the price of the book, Yates clearly stated the position of the two authors:

> Professor Fisher is abroad at present and will not be back before the B. A. [British Association] I do not really like giving any opinion without consultation with him, but I do know he feels that the price should be kept as low as possible. You will remember that in our discussion in Edinburgh last year he expressed the hope that the book could be published at 5/- [five shillings]. It has grown somewhat since then, but I know he has still a low price in mind, say 7/6 [seven shillings and six pence]. I also agree that the price should be kept low. From what you say in your letter, and from the final format of the book, it seems that a price as low as 7/6 would not be possible, but I should regard 12/6 [twelve shillings and six pence] as the maximum, if we are to encourage a free sale. 15/- [fifteen shillings] is getting too near the £1 [twenty shillings] category.[380]

For Yates an affordable price would have limited copyright infringements and encouraged the research workers to purchase their own copy of the book. Moreover, better incomes could be expected from the following editions of the book because "as soon as the book becomes known, it should have as good a market as 'Statistical Methods' – no worker with the one will want to be without the other".[381]

The *Statistical Tables* had their official presentation at the British Association meeting held in Cambridge, in August 1938, where both the table makers of the British Association Committee and the research workers that attended the various sections of the meeting (botany, agriculture, psychology, engineering, chemistry, geology...) were present. The authors pushed the publisher to meet this deadline. "It would give the book a good start if it were ready for this meeting", wrote

---

378 Letter from F. Yates to R. A. Fisher, 13th May 1936, RR Library and Archive, STATS 7.5.

379 Fisher and Yates accepted for the first edition of *Statistical Tables* a royalty of ten per cent only (the standard was fifteen per cent) in order to concur with the publisher in the cost of printing the book (Letter from Oliver and Boyd to R. A. Fisher, 26th August 1938, Oliver and Boyd Collection, NLS).

380 Letter from F. Yates to Oliver and Boyd, 12th August 1938, Oliver and Boyd Collection, NLS.

381 Letter from F. Yates to Oliver and Boyd, 12th August 1938, Oliver and Boyd Collection, NLS.

Yates to Oliver and Boyd in May, and in August the publisher reported to him "that the booksellers in Cambridge appear to have made a good start with the sale of this work".[382]

2.3.4 *Computing the* Statistical Tables *throughout six editions*

Since 1938 up to the early 1960s six editions of the *Statistical Tables* appeared, one more or less every five years (1938, 1943, 1948, 1953, 1957, 1963). The series of new editions was interrupted by the death of Fisher in 1962 and after that moment only reprints of the sixth edition were published till the early 1980s, when the publication of the book eventually ended.

The planning for the first edition of Fisher and Yates' collection of statistical tables began in 1936, just two years before the publication of the book, but the materials in the collection had been computed over a longer time interval. For example, the tables of the Student's distribution, chi-square etc. had been already prepared in the 1920s for *Statistical Methods for Research Workers*; the table for making tests of significance on 2 × 2 contingency tables was rearranged from the version published in an earlier journal paper of Yates; some tables of probits were borrowed or adapted from the corresponding tables published in the *Annals of Applied Biology* by a former student of Fisher, Chester Bliss.

The tables that appeared in the first edition of *Statistical Tables* were computed by hand and with desk calculators by the human computers and statisticians in Fisher and Yates' laboratories. The correspondence among Fisher, Yates, Wilfred L. Stevens (Fisher's statistical assistant at the Galton Laboratory) and William G. Cochran (Yates' assistant at Rothamsted) unveils the hierarchical structure in the making of the book. Fisher and Yates consulted each other planning the tables and setting out the mathematical formulae for their calculation, they then left instructions for the computing staff and their assistants, who took directly part in the calculations and supervised the work of the computers.[383] In some cases the authors still computed values on their own and even tested the most effective way of employing the tables with different calculating machines, as in the case of Yates and the table of orthogonal polynomials: "you will remember I mentioned that I was timing some polynomial fitting to see how successive addition

[382] Letter from F. Yates to Oliver and Boyd, 5th May 1938, Oliver and Boyd Collection, NLS; Letter from Oliver and Boyd to F. Yates, 19th August 1938, Oliver and Boyd Collection, NLS (Material from the manuscript section of the National Library of Scotland Archives reproduced by kind permission of Pearson Education Limited).

[383] Letter from F. Yates to R. A. Fisher, 29th May 1936, RR Library and Archive, STATS 7.5; Letter from F. Yates to R. A. Fisher, 2nd June 1937, RR Library and Archive, STATS 7.5; Letter from W. G. Cochran to F. Yates, 23rd June 1937, RR Library and Archive, STATS 6.5; Letter from F. Yates to R. A. Fisher, 10th July 1937, RR Library and Archive, STATS 7.5; Letter from R. A. Fisher to F. Yates, 15th July 1937, RR Library and Archive, STATS 7.5; Letter from F. Yates to R. A. Fisher, 27th July 1937, RR Library and Archive, STATS 7.5.

compared with multiplication by the ξ's. As you will see by the enclosed summary, the ξ's have it hands down if no printing adder is available".[384]

The computation of the tables was just the first step. The manuscript should then be checked, typed and checked again and later also much proofreading was required for detecting misprints and improving the readability of the tables. The typesetting of a mathematical table was a crucial aspect for its final usability. Type, spacing, decimals and so on were carefully chosen by the author rather than by the publisher as they were an integral part of the table making enterprise and Yates was especially nit-picking on this point:

> Thank you for the revised proof of the Table which I am returning herewith. It is now quite legible, but the spacing of the columns still leaves something to be desired. Most of the difficulty seems to arise through keeping the whole of each column throughout the three panels in line. I don't think this is at all necessary, in which case each panel should be spaced on its own merits. Following the rule I gave in my last letter, namely inserting equal spaces between each number of the bottom line of the panel, panels 26 and 27 are now correctly spaced, but the remainder require some adjustment. Some of the fractions might also be drawn a little more to the right as shown.[385]

Despite the careful proofreading, complete accuracy remained only a theoretical goal. Edition after edition Fisher and Yates corrected the errors that they discovered or that were brought to their attention by the users of the tables. The fellow statistician Horace Norton, for example, made enquiries "about discrepancies between Colcord and Deming's original table and the excerpt in Fisher and Yates", but Yates did not "recall the detailed story of the table" and could only promise to "have the discrepancies looked into before the next edition is compiled".[386] In the same way Yates and David J. Finney, a former assistant of both him and Fisher, corresponded on a possible error in the *z*-transformation, reported in the computing journal *Mathematical Tables and Other Aids to Computation*. But Finney assured Yates "I believe that we later checked this and agreed that you were right and your critic wrong. After working it again, I still agree with you".[387]

The publisher prepared for the authors of the *Statistical Tables* an interleaved copy of the book so that they could keep trace of any mistake or necessary addition to the book for future

---

384 Letter from F. Yates to R. A. Fisher, 30th July 1937, RR Library and Archive, STATS 7.5. In the enclosed note the minutes needed for fitting a 5th degree polynomial are compared using successive addition and multiplication with a Millionaire calculator and a Burroughs adding machine. The table of orthogonal polynomials had been sent to Fisher a few days before: "Herewith the table of orthogonal polynomials, which I have now completed. […] I think we can take it that these tables are correct, since the sums of squares have been computed direct and from your formula." (Letter from F. Yates to R. A. Fisher, 27th July 1937, RR Library and Archive, STATS 7.5)

385 Letter from F. Yates to Oliver and Boyd, 4th October 1937, Oliver and Boyd Collection, NLS.

386 Letter from H. W. Norton to F. Yates, 6th February 1947, RR Library and Archive, STF 31 (old catalogue reference); Letter from F. Yates to H. W. Norton, 6th February 1947, RR Library and Archive, STF 31 (old catalogue reference).

387 Letter from F. Yates to D. J. Finney, 25th November 1946, RR Library and Archive, STF 31 (old catalogue reference); Letter from D. J. Finney to F. Yates, 27th November 1946, RR Library and Archive, STF 31 (old catalogue reference). *Mathematical Tables and Other Aids to Computation* was the first computing journal published worldwide. The publication of the journal began after WWII.

revisions. The use of the interleaved copy was an established practice for the books published by Ronald Fisher with Oliver and Boyd, but in the case of the collection of tables it turned out as a particularly handy solution. In the average time of five years that passed from one edition to the other, mistakes pointed out by the users could be easily forgotten without such systematic record.[388]

The *Statistical Tables* were constantly updated and enlarged throughout the years and the last edition counted fifty-four tables. The new tables that were added to the book were mainly tables useful for biological and medical research, like the tables for computing linkage (3rd and 5th edition); a new table for probits (3rd edition); tables with random permutations of ten and twenty numbers for general use in the construction of experimental arrangements (4th edition). Further experimental designs were also introduced.

After WWII the making of mathematical tables consistently changed due to the availability of electronic mainframes. On the one hand with digital computers a numerical value could be computed in a few milliseconds, offering unprecedented opportunities for tabulating mathematical functions in a quick and reliable way, but on the other, the increased availability of mainframes radically reduced the need of printed collections of mathematical tables to be employed in scientific work.

The *Statistical Tables* by Fisher and Yates could not escape this fate. The early use of mainframes in the Rothamsted statistics department largely computerised the making of the new tables that entered in the later editions of the book. For example, Michael Healy one of the programmers of the computer Elliott 401 available at Rothamsted since 1954 (**chapter 4**), prepared for the fifth edition, under Fisher's theoretical guidance, the table on the "Significance of difference between two means" for the application of the Behrens test.[389] When the 401, a machine with limited potentialities, was not suitable, other external computer resources were used, although this choice might cause some delay, as in the preparation of the last edition of the book.[390]

---

388 Oliver and Boyd prepared for each edition of Fisher's books an interleaved copy (see Oliver and Boyd Collection, NLS). The first one was sent to him in 1925 for *Statistical Methods for Research Workers* (Letter from Oliver and Boyd to R. A. Fisher, 25th June 1925, Oliver and Boyd Collection, NLS). In the case of the *Statistical Tables* both the authors, Fisher and Yates, received a copy for annotations. "An interleaved copy [of *Statistical Tables*] has been prepared and dispatched to you to-day by Parcel post. A similar copy has been posted to Professor R. A. Fisher." (Letter from Oliver and Boyd to F. Yates, 19th August 1938, Oliver and Boyd Collection, NLS. Material from the manuscript section of the National Library of Scotland Archives reproduced by kind permission of Pearson Education Limited)

389 Correspondence between R. A. Fisher and M. J. Healy (March-April 1956), R. A. Fisher Digital Archive; Letter from R. A. Fisher to F. Yates, 29th May 1956, RR Library and Archive, STF 31 (old catalogue reference); Letter from R. A. Fisher to F. Yates (with enclosed note on the computation of the table for the Behrens' test using the Elliott 401), 5th June 1956, RR Library and Archive, STF 31 (old catalogue reference).

390 "I have been awaiting the computation of a new table on a large electronic computer (not the Rothamsted computer) which should have been done some months ago, but owing to machine difficulties it got held up." (Letter

By the early 1960s more than one statistician volunteered for extending the *Statistical Tables* with the aid of an electronic computer, but the project was never realized and in the 1980s personal computers provided with off-the-shelf statistical software made the *Statistical Tables* anachronistic and the publication of the book ended, with much regret of Frank Yates.[391]

2.3.5 *A successful publication*

The collection of tables prepared by Fisher and Yates was a successful book: by 1963 twenty-two thousand volumes of its English edition had been sold and a Spanish translation had been issued.[392] The several editions of the book were constantly reviewed in general scientific journals, like *Nature* and *Science*, but also in publications addressed to specific audiences, like the *Journal of the Royal Statistical Society*, the *British Medical Journal*, the *Canadian Medical Association Journal*, the *Quarterly Review of Biology*, the *Eugenics Review*, the *Journal of Forestry* (Appendix 2.2). Many reviewers of the book were statisticians because Fisher and Yates decided to send the book with a presentation letter directly to them, as Yates wrote to Oliver and Boyd:

> I have been discussing with professor Fisher the question of Journals to which send copies, and we came to the conclusion that it would be best to send only a moderate number to the editors of the Journals, and in addition to send presentation copies to certain statisticians, with a request that they should review the work in whatever Journal they consider most suitable. We are preparing a list of such people, and we will let you have it in the course of a few days.[393]

The reviewers praised the quality of the publication – "printing, spacing and presentation and the large flat page, are according to the best canons", "paper, type and arrangement have been well chosen to minimize time and eye-strain" –, its modest cost that remained "remarkably low" even for the sixth edition, and the presence in one single collection of the whole set of tables for the application of analysis of variance and experimental design.[394] Statisticians especially appreciated the opportunity to interpolate the entry values for an efficient use of the tables.

Reviews were also used by different categories of users to formulate a wish list for further editions, suggesting possible additions to the book or the extension of current tables. Sometimes users wrote in person to Fisher, Yates or their publisher recommending a particular solution, as a

---

from F. Yates to Oliver and Boyd, 15th November 1962, Oliver and Boyd Collection, NLS)

391 "I told Grant [Oliver and Boyd's manager] that a number of people with electronic computers were thinking of extensions of our tables" (Letter from R. A. Fisher to F. Yates, 28th June 1961, RR Library and Archive, STF 31 (old catalogue reference)).

392 For the number of copies of *Statistical Tables* (English edition) sold from 1938 to 1963 see Letter from Oliver and Boyd to F. Yates, 10th June 1963, Oliver and Boyd Collection, NLS.

393 Letter from F. Yates to Oliver and Boyd, 25th August 1938, Oliver and Boyd Collection, NLS.

394 The quotations are taken respectively from: J. Wishart (1939), H. Hotelling (1938), I. D. Hill (1964).

Dr. Lipscomb who suggested the addition of a thumb index to the book, an advice enthusiastically received by the authors.[395]

Along with the reviews, also the requests for reprinting materials from the *Statistical Tables* offer an insight on the success of the book as a computing tool for bringing analysis of variance and experimental design within the grasp of research workers and statisticians. The people who approached Fisher and Yates or their publisher, Oliver and Boyd, with requests for reprinting tables in publications that dealt with Fisher's statistical methods were both authors engaged in publishing books for the market and teachers at college or university level willing to reproduce a few tables for their students. The tables that were in higher demand among research workers were the tables of the chi-square and Student's distribution, the basic tools for making tests of significance. Requests for reprinting these tables had been received by Fisher since the publication of *Statistical Methods for Research Workers* and since then he had adopted a liberal policy for granting permission to reproduce these materials.[396] For Ronald Fisher all these requests were an "excellent opportunity of getting the *Tables*" – and, by consequence, his statistical methods – "known to a new class of users".[397]

Some requests came to Fisher and Yates also from statisticians that had the competence, but were unwilling or unable due to time and money shortage to repeat the computational labour of making a whole new table. Moreover, reproducing and not making a new table increased accuracy overall, because the table borrowed had been already checked by its users.

Among the statisticians who approached Fisher and Yates for reproducing materials from the *Statistical Tables* there was also Egon Pearson, the son of Karl Pearson. In the 1950s he co-edited with Herman O. Hartley the *Biometrika Tables for Statisticians*, updating the collection of statistical tables prepared by the father Karl in the 1910s. Egon Pearson contacted Fisher and Yates asking permission to reproduce some of the values computed for both the z distribution and for Student's distribution, and some scores for ranked data. Yates considered that he and Fisher "should certainly accede to Egon's request with regard to $\underline{t}$ and $\underline{z}$ as we ourselves used their tables to prepare the 10 per cent points of z and e^{2z}", and, despite the animosity between Fisher and Karl Pearson's son, suggested Fisher that "it will be politic" also to grant permission for the

395 Letter from R. A. Fisher to Oliver and Boyd, 3rd July 1944, Oliver and Boyd Collection, NLS; Letter from Oliver and Boyd to R. A. Fisher, 6th July 1944, Oliver and Boyd Collection, NLS; Letter from R. A. Fisher to Oliver and Boyd, 15th July 1944, Oliver and Boyd Collection, NLS.
396 See, for instance, the correspondence between Fisher and Oliver and Boyd since 1927 (Oliver and Boyd Collection, NLS).
397 Letter from P. M. Dance (Fisher's research assistant and secretary in Cambridge) to R. Hunt (Yates' secretary at Rothamsted), 23rd February 1953, RR Library and Archive, STF 31 (old catalogue reference).

expected values of ranked normal deviates, which the senior author of the *Statistical Tables* had computed by himself.[398]

Indeed, mathematical tables were regarded among statisticians as a commodity that could travel through personal or professional acquaintance across their community. Tables could be traded for 'political' reasons, as in the example here quoted, and exchanged for other tabular matter. Usually, however, the statisticians who requested permission to reprint from the *Statistical Tables* were former students or co-workers of the authors and sometimes also contributors to the collection as David J. Finney, who wrote to Yates:

> I am sure that you must grow tired of requests for permission to reproduce material from 'Statistical Tables'. I have in preparation a fairly complete account of the design and analysis of biological assay, and hope to have it ready for the press in about six months time. […] amongst my special tables, I would like to include abridged versions of some of the more common ones.[399]

The first requests for reprinting materials from *Statistical Tables* started in the early 1940s and continued throughout the following decades. In 1960 over thirty requests were received. Permission was usually granted, subject to two conditions: that a proper acknowledgement to Fisher and Yates' collection was given in the preface of the publication and next to the tables taken from *Statistical Tables*, and that two copies of the book were sent to Oliver and Boyd.

Through the publisher, the authors suggested to their applicants the following formula of acknowledgement:

> Table ….. is reprinted/abridged from Table ….. of […] Fisher &Yates's Statistical Tables, Oliver & Boyd, Edinburgh, by kind permission of the Author (s) and Publishers.
>
> In the Preface
>
> I am (also) indebted to R. A. Fisher, F. Yates and Messrs Oliver & Boyd of Edinburgh for permission to reprint Tables Nos. …. From their book Statistical Tables (1938) which contains an extensive range of tables well designed for modern statistical methods.[400]

Once permission was given, Oliver and Boyd dispatched the complimentary copies to the authors. Such was the stream of requests for reprinting materials from the *Statistical Tables* that in sending the umpteen volume the publisher commented: "By this time you must be bringing

---

[398] Letter from F. Yates to R. A. Fisher, 19th September 1950, RR Library and Archive, STF 31 (old catalogue reference). If Fisher had felt that E. S. Pearson was hostile to him since the late 1920s, when Pearson had published critical reviews of Fisher's *Statistical Methods for Research Workers* (E. S. Pearson (1926), E. S. Pearson (1929)). Yates' relationship with E. S. Pearson, instead, should have been more amicable, as it was E. S. Pearson who seconded Yates' nomination to the Royal Statistical Society – Fisher was the proposer – in 1933 (Royal Statistical Society, Nomination Paper 4999).

[399] Letter from D. J. Finney to F. Yates, 6th November 1950, RR Library and Archive STF 31 (old catalogue reference).

[400] Letter from R. A. Fisher to Oliver and Boyd, 4th November 1940, Oliver and Boyd Collection, NLS. "I suppose the Weirton Steel Company should get our standard answer about the use of tables, in which, if you remember, we make some stipulation as to the form of acknowledgement deemed appropriate and ask for two copies of their publication and of any further editions it may attain to." (Letter from R. A. Fisher to Oliver and Boyd, 21st June 1943, Oliver and Boyd Collection, NLS).

together quite an interesting collection of books on this [the volume mailed was a Swedish monograph for biologists] and allied subjects."[401]

Beyond the irony, the requests received by Fisher and Yates for reprinting materials from the *Statistical Tables* offer the opportunity to trace the diffusion of analysis of variance and experimental design throughout three decades and well beyond the fields that represented the original target of the book. Requests for reprinting tables, in fact, reached Fisher and Yates not only from biology, agriculture and medicine, but also from psychology, economics, quality control in industry, education studies, sociology, aviation, the military, institutes of actuaries, just to mention a few. A selection of the requests collected during my archival research is presented in Appendix 2.3.

2.4 *The making of a political artefact*

2.4.1 *Table making and statistical significance*

In order to instance my point on how power and authority can be embedded in a computing tool, I want to examine in detail the format chosen by Ronald Fisher for the tables of the chi-square distribution printed in the *Statistical Tables*. The same format of the chi-square distribution was adopted also for Student's distribution, which was the other crucial tool for making tests of significance and thus particularly useful for the research workers interested in the analysis of their data.

The tables of the chi-square and Student's distribution in Fisher and Yates' collection had been reprinted and enlarged from the corresponding tables computed in the 1920s by Ronald Fisher for *Statistical Methods for Research Workers*. In order to avoid copyright issues with Pearson, Fisher had to choose a different layout for his own tables and the comparison between Pearson's and Fisher's tables of the same statistical distributions offers an insight into the political dimension of table making for statistics.

Karl Pearson introduced the chi-square test in 1900 as a quantitative standard for testing goodness-of-fit for frequency curves. The first extended table of the chi-square distribution was prepared by a co-worker of Pearson in the Biometric Laboratory, W. Palin Elderton, published in the first volume of *Biometrika*, and reprinted in the *Tables for Statisticians and Biometricians*. In the table computed in Pearson's laboratory the values of chi-square were tabulated in the rows, while

[401] *Biologisk Variationsanalys* by G. Bonnier and O. Tedin (Letter from Oliver and Boyd to F. Yates, 6th November 1945, Oliver and Boyd Collection, NLS. Material from the manuscript section of the National Library of Scotland Archives reproduced by kind permission of Pearson Education Limited).

the columns gave the number of parameters in the statistical problem. The entries of the table were the corresponding probability values. In Fisher's table of the chi-square distribution, instead, the probability was given in the columns, while in the rows there were the degrees of freedom.[402] In Pearson's table, thus, the user could read continuous probabilities, while in Fisher's version the chi-square values were tabulated only for discrete probability values ranging from .99 to .01, in agreement with the guidelines for significance testing presented in *Statistical Methods for Research Workers*. In his textbook, in fact, Fisher had suggested that

> In preparing this table we have borne in mind that in practice we do not want to know the exact value of P for any observed χ^2, but, in the first place, whether or not the observed value is open to suspicion. [...] We shall not often be astray if we draw a conventional line at .05, and consider that higher values of χ^2 indicate a real discrepancy.[403]

In Fisher's version of the chi-square table the column corresponding to a probability value of five per cent could be immediately singled out for each degree of freedom without any need of interpolation, making the application of Fisher's guidelines for significance testing straightforward. This format of Fisher's table – published in *Statistical Methods*, reprinted in the *Statistical Tables* and reproduced in many other sources – contributed, more than any theoretical consideration, to spread the five per cent threshold for statistical significance that he had recommended in his textbook in order to accept or reject the null hypothesis with the chi-square test. As pointed out by Theodore Porter "[t]his particular number [the five per cent], an indication of the probability that the results might have occurred by chance, is clearly no more than a convention".[404] Nevertheless it is today largely accepted and among psychologists it even became a criterion for determining whether an experimental result should be published or not.

Paradoxically, Fisher's own attitude towards the five per cent threshold changed throughout his life and became much more flexible that the one showed by the research workers who employed his statistical methods.

> Fisher's early writings encouraged this as a 'convenient convention', though the criticism of Neyman and Egon Pearson drove him to a more nuanced view. This later view was that 'no scientific worker has a fixed level of significance at which from year to year, and in all circumstances he rejects hypotheses; he rather gives his mind to each particular case in the light of his evidence and his ideas'.[405]

---

[402] "[E]very free parameter in the distribution formula reduced by one the degrees of freedom of the test." (T. M. Porter (2006), p. 256). Fisher's position on this point is nowadays widely accepted. The fierce controversy between Ronald Fisher and K. Pearson on the chi-square distribution is described in J. Fisher Box (1978), pp. 84-88. For a re-assessment of the controversy on degrees of freedom between Pearson and Fisher see S. M. Stigler (2002), chapter 19.

[403] R. A. Fisher (1946), p. 80. In chapter 3 I give a concrete example of the application of the chi-square test in the analysis of blood group data.

[404] T. M. Porter (1996), pp. 211-212.

[405] G. Gigerenzer et al. (1989), p. 78.

However, Fisher's tables and textbooks, prepared at the beginning of his career, carried with them the bias towards the five per cent threshold, influencing their readers for generations. This is a clear evidence of the power that an *ad hoc* computing tool can have in the dissemination of statistical ideas.

In *Statistical Methods for Research Workers* Fisher recomputed also the table of Student's distribution with the same format adopted for the chi-square table.[406] Again in Fisher's version were emphasized the discrete probability values and the column of the five per cent probability was well evident among them, while Gosset had prepared a table with continuous probabilities even for his 1925 version.[407] Fisher, thus, assured consistency to his guidelines for significance testing and the tables of the chi-square and Student's distribution – the latter synthetically indicated by Fisher as *t* distribution – played a key role in disseminating his ideas. That is what I mean for politics in relation to the *Statistical Tables*.

With liberality, Fisher authorized research workers and statisticians to reproduce his tables of the chi-square and Student's distribution and this copyright policy contributed to the diffusion of Fisher's format for these tables over the version printed in *Biometrika* and in the *Tables for Statisticians and Biometricians*. Fisher's format was so influential to be reproduced also in the *Biometrika Tables for Statisticians*, the 1950s update of Karl Pearson's collection.[408] In the preface of the book, the editors, Egon Pearson and Herman Hartley, on the one hand, attempted to claim independence from Fisher and Yates' work

> [t]he plan for the computation of tables of percentage points of the Beta distribution and of more extensive tables than then existed of the percentage points of Student's *t*, of χ^2 and of the variance ration, *F*, had been sketched out before *Statistical Tables for Biological, Agricultural and Medical Research* by Professor Fisher and Dr Yates appeared in 1938

but on the other, they had to acknowledge that "it is evident in these and other tables how much we have owed to the scheme of tabular presentation first used by Fisher in his *Statistical Methods for Research Workers*."[409] In the 1940s Fisher's format for the tables of the chi-square and Student's distribution had already become canonical and it is still used in current textbook of statistics.

To a certain extent, however, the acquisition of Fisher's format in the *Biometrika Tables for Statisticians* reduced the citations of Fisher and Yates' collection, as Egon Pearson did not follow

---

406 R. A. Fisher (1946), p. 119.

407 "For the purpose of the present book we require the values of *t* corresponding to given values of *P* and *n* [degrees of freedom]" (R. A. Fisher (1925a), p. 22). Gosset's third version of the table is Student (1925).

408 Before the publication of the *Biometrika Tables for Statisticians* (first edition 1954), tables computed according to Fisher's format had been already printed in *Biometrika*.

409 E. S. Pearson and H. O. Hartley (1966), p. *ix*. The affinity between the outline of the tables chosen by Fisher and the one adopted by Pearson and Hartley is particularly evident in the table that gives the percentage points of the chi-square distribution. The variables tabulated (degrees of freedom and probability values) are the same as in Fisher's version, the only difference being the choice of the intervals. The column corresponding to the probability of five per cent is present (E. S. Pearson and H. O. Hartley (1966), Table 8, pp. 136-137).

the jealous copyright policy of the father, but granted permission to reproduce materials from both *Biometrika* and the *Biometrika Tables for Statisticians*. In the 1950s some of the applicants who approached Fisher and Yates for reprinting materials from the *Statistical Tables* reproduced alongside these materials also tables derived from *Biometrika*.[410] For some requests listed in Appendix 2.3c, I discovered that the applicants eventually reproduced only (or mostly) materials from Egon Pearson's collection and not from Fisher and Yates' book. For instance, Carl Bennett chose to reprint the distribution of chi-square, the Student's distribution and the distribution of the variance ratio from the set of tables recomputed *a-la*-Fisher by Pearson and his co-workers and printed in *Biometrika* during the 1940s, despite the permission to reprint the same materials gained also from Fisher and Yates.[411]

2.4.2 *Copyright and authority*

In order to understand the copyright choices made for the *Statistical Tables* it is necessary to examine at first the attitude taken by Ronald Fisher in relation to the tables that he computed for *Statistical Methods for Research Workers* and that were reproduced in the main collection. Through the publisher, Oliver and Boyd, Fisher requested to Karl Pearson the permission to reprint the same tables already published in *Biometrika* (Student's distribution, chi-square etc.) in *Statistical Methods for Research Workers*, but Pearson denied this permission.[412] Fisher, faced with the choice to eliminate the tables from the book or calculate them anew, decided for the second option convinced that computing tools could promote his new statistical methods:

> In relation to my book [*Statistical Methods for Research Workers*] they [the tables] appear to me to be valuable, for the book consists principally of examples of the use of the newer statistical methods for which these tables are needed, and if I had left the reader merely with formulae, or with references to tables elsewhere, I should have seriously restricted the utility of the book.[413]

The computation of the six tables to be included in *Statistical Methods* represented a consistent labour and Fisher explicitly requested to the publisher to maintain the copyright on them, using the same tables, in their integrity or in abridged form, "whenever and wherever" he wished and authorizing their use at his discretion.[414] The publisher accepted Fisher's request but suggested

---

[410] See in Appendix 2.3c, for instance, the requests of Virginia L. Sanders (March, 1957) and Bernard W. Lindgren (September, 1957).

[411] See for instance the requests of M. W. Tate (September, 1951), C. A. Bennett (June, 1952) and H. M. Walker (May, 1953).

[412] Letter from Oliver and Boyd to R. A. Fisher, 21st August 1923, Oliver and Boyd Collection, NLS; Letter from Oliver and Boyd to R. A. Fisher, 25th August 1923, Oliver and Boyd Collection, NLS.

[413] Letter from R. A. Fisher to Oliver and Boyd, 29th May 1924, Oliver and Boyd Collection, NLS.

[414] Letter from R. A. Fisher to Oliver and Boyd, 23rd May 1924, Oliver and Boyd Collection, NLS.

him to follow the example of Karl Pearson in relation to the tables, refusing potential applications for reprinting materials in order to increase his royalties.[415]

To the suggestion Fisher replied with a lengthy letter in May 1924 stating his ideas on the copyright and the use of the tables. At first, Fisher remarked the difficulty to defend the copyright because mathematical formulae and methods are public property and minor alterations – such as changing the interval chosen for the tabulation or the variable tabulated – might easily allow plagiarism. Moreover, he felt that it was not in the public interest to restrict the circulation of primary scientific work, such as mathematical tables.

The political argument on the use of the tables, i.e. the possibility to use the computing tool as an element for promoting his statistical methods, comes at the conclusion of the letter. Fisher wrote, in fact,

> The sale of my book depends, if I am not mistaken, principally on the speed with which certain relatively new statistical processes become familiar to research workers, and university teachers. One of the greatest helps in this direction would be in the wider knowledge of these tables. At present my book is the only attempt to explain these processes to the non mathematical research worker, and they are for the most part unknown to the mathematicians, for some years it will probably be the only book of its kind. In these circumstances I suggest that the sale of the book would stand to gain, by the inclusion of the tables in (i) standard collections of mathematical tables, or (ii) even elementary text books of statistics, written by authors with no real knowledge of how to use them.[416]

This idea of copyright stated in 1924 stood for the future and it was not changed by the growing personal fame of Fisher and of his statistical methods. In fact, when in 1939 Oliver and Boyd complained at a new request of permission for reprinting tables from *Statistical Methods* – "I wonder sometimes when will these applications come to an end. It is all very good and nice that some should do the work and other reap, at least, the monetary benefit" – Fisher again defended his position:

> copyright claims may be more defensible than I have imagined; but, if so, I should still be very unwilling to restrict the fullest use being made of mathematical work already done by the incorporation of the tables in books or bulletins intended to recommend the methods to new bodies of users.[417]

Fisher's attitude towards the tables in *Statistical Methods for Research Workers* and the procedure established for granting permission to reproduce them influenced the choices made for the *Statistical Tables*. The request of acknowledgement in the preface and on the page of the table and the two complimentary copies for the publisher were the same procedure adopted for *Statistical*

---

415 "That in itself will further the sale of the book, and as the royalty is a reasonable one, we hope, it will tend to increase the sums payable to you each year". (Letter from Oliver and Boyd to R. A. Fisher, 27th May 1924, Oliver and Boyd Collection, NLS. Material from the manuscript section of the National Library of Scotland Archives reproduced by kind permission of Pearson Education Limited)

416 Letter from R. A. Fisher to Oliver and Boyd, 29th May 1924, Oliver and Boyd Collection, NLS.

417 Letter from Oliver and Boyd to R. A. Fisher, 14th February 1939, Oliver and Boyd Collection, NLS ((Material from the manuscript section of the National Library of Scotland Archives reproduced by kind permission of Pearson Education Limited); Letter from R. A. Fisher to Oliver and Boyd, 16th February 1939, Oliver and Boyd Collection, NLS.

Methods. Moreover, after the publication of the *Statistical Tables* Fisher asked people that requested permission about the tables in *Statistical Methods* to mention in the preface of their publication even the more extensive collection at that point available.

> I am sending an application from a Swedish statistically-minded plant breeder, Olaf Tedin, who seems to have undertaken to write a statistical introduction to a new German handbook of plant breeding. […] so far as his suggestions for small tables drawn from *Statistical Methods* are concerned, they seem to me eminently harmless in respect of their influence on the sale of *Statistical Methods*. I shall, if you permit publication, suggest the following note: 'This table is condensed from the table of chi-square (or *z*, or *r* and *z*, as the case may be) given in Fisher's *Statistical Methods for Research Workers*, 7th edition, 1938, Oliver and Boyd, Edinburgh, and in *Statistical Tables for Biological, Agricultural and Medical Research* by Fisher and Yates, in which a more extensive collection of tables has been produced by the same publishers.[418]

Such was the long-lasting influence of the tables printed in *Statistical Methods for Research Workers* that some authors – one evident example in Appendix 2.3c is Donald Mainland's application in July 1950 – requested to reproduce tables from the new collection co-authored by Fisher and Yates, but acknowledged the tabular matter with the numbering adopted for *Statistical Methods*. This generated some discrepancies between the tables officially requested and the ones eventually reproduced, as the tables of the chi-square and Student's distribution in *Statistical Methods* and in the *Statistical Tables* were swapped over.

For granting permission in reprinting from *Statistical Tables*, Fisher's consensus was not enough. Also the co-author, Frank Yates, should be consulted and be agreeable. From the letters exchanged between the two authors and from their correspondence with Oliver and Boyd it does not emerge any idea of Yates about the use of the tables for the dissemination of Fisher's statistical methods.[419] Rather the letters convey his scepticism about the possibility to enforce copyright on this computing tool. According to Yates, mathematical tables could be recomputed from one or many sources – or claimed to be so – and if major and minor errors might be traced as a proof of plagiarism, this attitude would be regarded "as an unworthy and undignified procedure" in the scientific community. Furthermore, for Yates mathematical tables were a commodity in the community of statisticians. He therefore emphasized that Student's, *z* and chi-square values were "in the public domain, rather like logarithms and trigonometrical functions" and that some tables in the collection had been computed by other statisticians and research workers who were the real owners of the copyright. According to Yates, thus,

> [t]he point that has to be decided […] is whether it is better to let them re-compile their own values, doubtless using our tables as a check, or give them permission to reproduce with the resultant acknowledgment to our tables. On this issue I am quite prepared to abide by your own and Professor

[418] Letter from R. A. Fisher to Oliver and Boyd, 15th June 1939, Oliver and Boyd Collection, NLS.

[419] As mentioned before, however, even Yates conceived the collection of tables as a companion to Fisher's statistical textbooks, in particular *Statistical Methods for Research Workers*.

> Fisher's views, but I thought it would be worth setting out what I think to be the position on the copyright issue.[420]

On the final point raised by Yates, Fisher had a clear idea since the 1940s. He believed that "the acknowledgement of origin which we have asked for is perhaps the best return we can expect for the use of the tables".[421] The increasing number of citations was evidently perceived by Fisher as a way to strengthen the authority of his collection.

The only applications that Fisher and Yates assessed in a stricter way were the requests whose commercial value seemed to the authors greater than their scientific importance or the requests for reprinting the random numbers, a table for which Fisher was reluctant to give permission, possibly because the constant reproduction of the same numbers might impair their intrinsic chance. He believed that "it would be far better if everyone developed their own", but even in this case he recognised that he was "probably too late in making difficulties".[422]

Therefore, not for quite the same reason, but nonetheless in agreement, Fisher and Yates concurred in giving permission to reproduce from the book to almost all their applicants. Reference to the *Statistical Tables*, thus, could then be found in very different publications – home made textbooks to be used in class seminars, business and military pamphlets, commercials books – that just had in common the use of Fisher's statistical methods.

2.4.3 *The politics of the* Statistical Tables

As mentioned in the introduction of the present chapter, Langdon Winner distinguishes between the artefacts political for their design and arrangement and the technological systems that require certain power relations in the society in order to justify their existence. In my argument on the politics of the *Statistical Tables*, I have been mainly interested in the first category of artefacts examined by Winner. In this category he mentions three main examples: the low-hanging overpasses designed by the architect Robert Moses in order to prevent the buses of the public transport – used by racial minorities and low-income social groups – from accessing Jones Beach, his most famous New York public park; the pneumatic moulding machines that allowed the entrepreneur Cyrus McCormik to outwit the National Union of Iron Moulders in his

---

[420] All the quotations are taken from the Letter from F. Yates to Oliver and Boyd, 11th August 1953, Oliver and Boyd Collection, NLS.

[421] Letter from R. A. Fisher to Oliver and Boyd, 4th November 1940, Oliver and Boyd Collection, NLS.

[422] Letter from R. A. Fisher to Oliver and Boyd, 16th January 1952, Oliver and Boyd Collection, NLS. Another set of requests for both *Statistical Methods for Research Workers* and *Statistical Tables* that could not be met were the requests coming from countries in which a translation of the books – *Statistical Methods* was translated in French, German, Italian, Spanish ad Japanese, while the *Statistical Tables* only in Spanish – had been issued by a local publisher who was also the owner of the book copyright for that country. Fisher did not pay much attention to this point and it was the publisher, Oliver and Boyd that constantly reminded him the limitation.

company during the nineteenth century; the mechanical tomato harvester that since the 1940s drastically reshaped tomato production in rural California transforming the tomato business in a large industry.

Winner considers all these artefacts as good cases of how power and authority can be embedded in a technology. Not in all of them there is an explicit plan for infusing political qualities into the artefact, but nonetheless in all these cases political categories are felt necessary for a complete understanding of the technology at stake.[423] The examples proposed by Winner are all chosen looking at the 'hardware' – the height of the overpasses, the performance of the moulding machines, the features of the tomato harvester –, while in examining computing tools such as mathematical tables it is the 'software' – i.e. intangible aspects like planning, usability and copyright – that especially matters. In particular, in the case study that I have examined I link the political qualities of the *Statistical Tables* to two main aspects, the design of the collection and the idea of copyright endorsed by its authors.

The making of the *Statistical Tables*, in fact, was not a straightforward application of mathematical formulae. Careful choices were made in relation to the inclusion/exclusion of tables from the collection, their format and usability. The introduction of the book emphasized which problems could be solved through Fisher's statistical methods and the *Statistical Tables* were indeed presented as a companion to Fisher's textbooks on analysis of variance and experimental design and closely followed the approach given in these publications. The affordable price of the book and the constant inclusion of new tables of interest for experimental research contributed to the success of the collection among research workers, who gained ready-made statistical tools to employ in their field or laboratory investigations.

In relation to my second point, the political use of the copyright, it is evident how the intellectual property on computing tools can be conceived in very different ways comparing the attitude taken on this issue by Karl Pearson, Ronald Fisher and Frank Yates. Pearson's main concern was preserving through the copyright the revenues offered by the tables published in *Biometrika*. Fisher, instead, recognised that the copyright could be an instrument for the dissemination of his statistical methods rather than an opportunity for further profits and in the former capacity he valued and exploited the copyright on the *Statistical Tables*. The co-author of the book, Yates, did not share Fisher's visionary attitude, but mainly conformed to his policy for lack of faith in the copyright on mathematical tables. Nonetheless the net outcome was a liberal

423 "As far as I know, no one has argued that the development of the tomato harvester was the result of a plot. Two students of the controversy, William Friedland and Amy Barton, specifically exonerate both the original developers of the machine and the hard tomato from any desire to facilitate economic concentration in that industry." (L. Winner (1980), p. 126)

attitude in granting permission to reproduce materials from the book and the dissemination of the tables computed by Fisher and Yates in tens of other publications.

In the arrangement and copyright management of the *Statistical Tables*, nothing was left to chance and through the choices of the authors I claim that the collection of tables became a value-laden technology for the propaganda of Fisher's statistical methods in scientific research, and therefore a political technology. The political qualities of the collection become even more evident when we consider that consumers and makers of technologies do not usually share the same degree of power or awareness, as Langdon Winner reminds us.[424] Therefore, if the authors of the *Statistical Tables* perfectly knew the bias towards Fisher's methods in the collection, the research workers who purchased the book accepted with the computing tool also this peculiar vision of statistics on which they had no control, as I have discussed in relation to Fisher's guidelines for significance testing.

My examination of the *Statistical Tables* has focused on the collection of tables as a political artefact within the community of its users, but it is also true that in a broader perspective Fisher and Yates' book affected society at large and thus, the *Statistical Tables* can be included even in the second category of political artefacts defined by Langdon Winner. For instance, I have argued that the format of the tables of the chi-square and Student's distribution reproduced in Fisher and Yates' collection contributed to the diffusion of the conventional threshold of five per cent in accepting or rejecting the null hypothesis, while making tests of significance. The effects of this convention were immediately perceived in scientific research, however, if at stake there was a medical trial for the assessment of a new drug or an agricultural experiment on the best way to increase the productivity of a crop, the scientific results translated also in public health or economic issues.

The economists Stephen Ziliak and Deirdre McCloskey's have raised this point in *The Cult of Statistical Significance*, in which they harshly criticize Fisher's ideas on statistical significance for their social cost. Ziliak and McCloskey's book resembles too much an unilateral condemnation to be taken at face value, but certainly Fisher's emphasis on tests of significance and his fixed conventions for accepting or rejecting an experimental result had and have social and political implications, well beyond the methodology of scientific research.

I believe that this second political dimension of the *Statistical Tables* is intrinsically linked to the first one I have examined at length in the present chapter, that is the decision of Fisher and Yates to craft a computing tool for statistics easy to use also for research workers that had only a limited mathematical training. Unlike Fisher and Yates' fellow statisticians, the research workers

[424] L. Winner (1980), p. 127.

accepted unquestioningly Fisher's methods and did not realize their limits, nor the undeclared intertwining between statistical ideas and computing tools.

The last remark I want to make on the politics of the *Statistical Tables* concerns my choice to follow Langdon Winner, rather than adopt more broadly the approach of the social construction of technologies. Winner himself has been challenged to defend his political interpretation of the New York overpasses built by the architect Moses against the interpretive flexibility promoted by social constructionists. To his critics Winner pointed out that

> What makes the conclusion that Moses' bridges are inegalitarian political artifacts a strongly defensible proposition [...] can be seen in the role that the bridges play in the social and political history of a particular community at a particular time, as well as in the personal history of a power broker notorious in his willingness to use all possible means, including public works projects, to shape social patterns to match with his vision of what was desirable.[425]

In the same way I claim that the making of the *Statistical Tables* was driven by the intention to craft a computing tool *ad hoc* for analysis of variance and experimental design, an instrument in which Ronald Fisher's statistical ideas were represented in every detail, from the format of the tables to the arrangement of the collection and the management of its materials. My interpretation is supported by the key role that Ronald Fisher attributed to computing tools in statistics. He believed that statistics and computing were two sides of the same coin, therefore an authoritative computing tool was for Fisher a concrete asset in the dissemination of his statistical methods and therefore it was a political tool. Otherwise it would be difficult to explain the time and effort that Fisher spent in computing statistical tables since the beginning of his career and the concerned letters he exchanged with his publisher, Oliver and Boyd, to retain the copyright on the tables he had computed.

2.5 *Conclusion*

Throughout the twentieth century statistics "has become one of the most important sources of scientific expertise and guarantors of objectivity in the modern world."[426] This aura of objectivity has enshrouded also the computing tools that were instrumental for its application. In particular, mathematical tables have been considered value-free, because computed from objective formulae. As I have discussed in the case of the *Statistical Tables for Biological, Agricultural and Medical Research* this claim should not be taken for granted.

---

[425] L. Winner (1993), p. 374.
[426] T. M. Porter (2006), p. 4.

If Karl Pearson's collection of statistical tables, influenced by the Victorian fascination for numbers, grew by accretion rather than selection including also tables, like the one of Student's distribution, peripheral to his agenda, Fisher and Yates' book was born in a more pragmatic context and with a political goal in mind. It was a carefully crafted instrument for presenting analysis of variance and experimental design as the right solution for dealing with statistical problems in experimental research. The political qualities of the collection were emphasised further by the copyright policy, strongly influenced by Ronald Fisher: tens of permissions were given for reprinting tables from the book and in the recipient publications mention was made to the original source circulating Fisher and Yates' work even beyond the initial ambitions of the authors. The commanding vision of statistics embedded in the *Statistical Tables* was also evident in the tabulation of specific functions, like the chi-square or Student's distribution. I have argued that their format in Fisher and Yates' collection was instrumental to support the guidelines for significance testing suggested by Fisher in his own textbooks.

How power and authority imbued the *Statistical Tables* becomes even more evident when we consider the reception of the book among its prospective publics of statisticians and research workers. Research workers were users, but not makers of tables, if we set aside specific examples like the brewer William Gosset. It is quite understandable, therefore, that many contributors and reviewers of Fisher and Yates' book were statisticians and not research workers. To statisticians number crunching was a relevant, and often enjoyable, part of their duties. They understood quibbles, problems and limitations in the computation of special functions and in the interpolation of numerical values reported in the tables, while research workers usually complied to rules, without much questioning, and limited computing as much as possible, as suggested by the distaste of biologists for interpolation. This class of users took the *Statistical Tables* at face value assimilating their ideas with even more loyalty than the authors themselves, as discussed about the conventional threshold for statistical significance. With them the *Statistical Tables* as a political technology proved successful indeed.

If the *Statistical Tables* impacted at first on the community of their potential users, the statistical methods they promoted had far-reaching impact on society, as Fisher's ideas on statistical significance contributed to decision-making throughout the twentieth century. In this sense the *Statistical Tables* are a political artefact also in the second meaning suggested by Winner, but a detailed examination of its implications would require another chapter and here I just briefly mentioned it.

In conclusion, I would like to remark that a key element in understanding the making of Fisher and Yates' collection has been their correspondence with the publisher Oliver and Boyd, not only on the *Statistical Tables*, but also on the other books written by Fisher and printed by the same firm, in particular *Statistical Methods for Research Workers*. The twentieth century author-editor correspondence of Oliver and Boyd has been preserved and the wealth of material available on the making of Fisher's books opens up a unique perspective on the aims and expectations that prompted the publication of the *Statistical Tables*.[427] For example, the discussion on the management of the book copyright could not be found in other archival sources, let alone in the secondary literature.

I claim therefore the relevance of archival sources on book publication for a deeper understanding of the scientific enterprise and of the aims embedded in it, but not explicitly declared. As a medium for the diffusion of knowledge, books, filled with words or figures, are rarely written without an agenda. Collections of mathematical tables are no exception, because mathematical formulae might not have politics, but the way in which tables are constructed, arranged and managed can, as I have discussed in the case of the *Statistical Tables*.

[427] NLS, Oliver and Boyd Papers, Inventory Acc. 5000.

APPENDIX 2.1: Statisticians members of the British Association Mathematical Tables Committee (later Royal Society Mathematical Tables Committee).

Name	Membership	Career in Statistics and Computing
Karl Pearson (1857-1936)	* 1894-1899 (*) Special Committee formed by R. Harley (chairman), A.R. Forsyth (secretary), J.W.L. Glaisher, A. Lodge and K. Pearson for the calculation of Pearson Integrals *G*(*r*, *v*), published in the British Association Report in 1896 and 1899	1903 Founder of the Biometric Laboratory, University College London. 1907 Director of the Galton Laboratory for National Eugenics. 1911 Head of the department of applied statistics that incorporated the Biometric and Galton laboratories. ❖ 1901 Co-founder of the journal *Biometrika* with F. Galton and F. R. Weldon. Editor of the *Tables for Statisticians and Biometricians* (first edition 1914). From 1919 editor of the series *Tracts for Computers*.
Ronald A. Fisher (1890-1962)	1925-1948; 1948-1949	1919-1933 Chief statistician, Rothamsted Experimental Station. 1933-1943 Galton professor of eugenics, University College London and head of the department of eugenics. 1943-1957 Balfour professor of genetics, Cambridge University. ❖ General editor of the British Association Mathematical Tables project (1928-1931). Senior author of the *Statistical Tables for Biological, Agricultural and Medical Research* (first edition 1938).
Alexander J. Thompson (1885-?)	1928-1948; 1948-1954; 1956-1965	1920-1945 Member of the statistical staff of the General Register Office, Somerset House, London. ❖ Author of the *Logarithmetica Britannica*, tables of logarithms from 10,000 to 100,000 published in nine parts from 1924 to 1952. The computation of the table started in 1922 in the Biometric Laboratory under Karl Pearson's supervision.
John Wishart (1898-1956)	1928-1948; 1948-1954	1924-1927 Research assistant to Karl Pearson, University College London. 1928-1931 Assistant statistician to Ronald Fisher,

		Rothamsted Experimental Station. 1931-1956 Reader in statistics, School of Agriculture, Cambridge University. ❖ Contributor to the *Tables of Incomplete Beta Function* supervised by Karl Pearson Secretary of the British Association Mathematical Tables Committee (1937-1946). Director of the statistical laboratory associated to the School of Agriculture, Cambridge University.
Joseph O. Irwin (1898-1982)	1928-1948	1921-1925, Appointed to work with Karl Pearson, University College London. 1928-1931 Assistant statistician to Ronald Fisher, Rothamsted Experimental Station. 1931-1965 Biostatistician, Medical Research Council, London School of Hygiene and Tropical Medicine. ❖ Contributor to Karl Pearson's series *Tracts for Computers.* Editor of the table of integrals and derivatives of the normal probability integral computed by J. R. Airey for the British Association. Editor of the tables of the probability integral computed by W. F. Sheppard and published by the British Association.
James F. Tocher (1864-1945)	1928-1945	1911-1941 Lecturer in statistics, University of Aberdeen ❖ "A man of wide and varied interests, he found himself engaged early in life on statistical computations in connection with Karl Pearson's population studies. He attended regularly the annual meetings of the Association, and it was at Glasgow in 1929 that he was invited to join the Committee. He lived too far away to attend the meetings of the Committee in London, but its work always interested him." (Extract from Tocher's obituary, published in the Report of the Committee, *The Advancement of Science*, 1948)
Egon S. Pearson (1895-1980)	1930-1933	1921-1933 Lecturer, department of applied statistics, University College London. 1933-1935 Reader in statistics and head of the department of applied statistics.

		1935-1960 Professor of statistics, University College London. ❖ 1936-1966 Managing editor of *Biometrika*. 1954 Co-editor of *Biometrika Tables for Statisticians* (the revised edition of Karl Pearson's *Tables for Statisticians and Biometricians*).
Wilfred L. Stevens (1912-1958)	1936-1958	1935-1941 Assistant Statistician to Ronald Fisher, Galton Laboratory, University College London. 1941-? Member of the statistics department, Rothamsted Experimental Station. ---- Lecturer in statistics at Coimbra University 1944-1947 Statistician at Imperial Chemical Industries Ltd, Billingham. 1947-1948 Statistician in the Admiralty statistics department. 1948-1958 Professor of mathematical statistics, University of Sao Paulo, Brazil. ❖ Joined the British Association Mathematical Tables Committee in 1936 as manager of the Galton Laboratory computing room. Contributor as table maker and proof-reader (first edition) to the *Statistical Tables*.
Rose O. Cashen (?-?)	1941-1948	October 1938-April 1940 Full time computer for the British Association Mathematical Tables Committee working in Ronald Fisher's Galton Laboratory at University College. ---- PhD in statistics 1948- ---- Statistician in the Admiralty ❖ Computer of the British Association Mathematical Tables Committee.
Frank Yates (1902-1994)	1948-1957	1931-1933 Assistant statistician to Ronald Fisher, Rothamsted Experimental Station. 1933-1968 Head of the statistics department, Rothamsted Experimental Station. ❖ 1927-1931 Computing officer for the Gold Coast Survey. Junior author of the *Statistical Tables for Biological, Agricultural*

		and Medical Research (first edition 1938). 1960-1961 President of the British Computer Society.
Alexander C. Aitken (1895-1967)	1948-1950; 1956-1960	1925-1946 Lecturer and reader in statistics and mathematical economics, University of Edinburgh. 1946-1965 Professor of mathematics, University of Edinburgh. ❖ Revised the mathematical tables prepared by John B. Clark, published in Edinburgh by Oliver and Boyd. Computer in E. T. Whittaker Mathematical Laboratory in Edinburgh.
Harold Jeffreys (1891-1989)	1949-1954	1946-1958 Plumian professor of astronomy and experimental philosophy, Cambridge University. Author of *Theory of Probability* (1939), a textbook on Bayesian principles in statistics. ❖ Co-author of a book on seismological tables published in 1940 by the British Association (Jeffreys' main scientific career was in geophysics).
Maurice S. Bartlett (1910-2002)	1962-1965	1933-1934 Assistant Lecturer in statistics, University College London. 1934-1938 Statistician, Imperial Chemical Industries, Ltd. 1947-1960 Professor of mathematical statistics, University of Manchester. 1960-1967 Professor of statistics, University College London. 1967-1975 Professor of biomathematics, University of Oxford.

APPENDIX 2.2: Selected reviews of the *Statistical Tables for Biological, Agricultural and Medical Research.*

Date	**Edition of the Book Reviewed**	**Journal**	**Reviewer**
1938	First Edition	*Science* (Vol. 88, No. 2295, pp. 596-597)	Harold Hotelling (statistician)
1939	First Edition	*Nature* (Vol. 144, No. 3647, p. 533)	John Wishart (statistician)
1939	First Edition	*Journal of the Royal Statistical Society* (Vol. 102, No. 2, p. 298)	Maurice G. Kendall (statistician)
1939	First Edition	*Journal of the American Statistical Association* (Vol. 34, No. 205, pp. 206-207)	George W. Snedecor (statistician)
1939	First Edition	*Journal of Forestry* (Vol. 37, No. 4)	Besse B. Day (statistician)
1939	First Edition	*The Canadian Medical Association Journal* (Vol. 40, No. 1, p. 104)	Anonymous
1939	First Edition	*Eugenics Review* (Vol. 30, No. 4, pp. 298-299)	F. R. Simpson (eugenicist)
1943	Second Edition	*Science* (Vol. 98, No. 2546, pp. 346-347)	Chester I. Bliss (biologist and statistician)
1943	Second Edition	*Annals of Eugenics* (Vol. 12, No. 1, pp. 147-148)	Kenneth Mather (geneticist)
1943	Second Edition	*Eugenics Review* (Vol. 35, No. 1, pp. 16-17)	F. R. Simpson (eugenicist)
1944	Second Edition	*Journal of the American Statistical Association* (Vol. 39, No. 228, pp. 523-524)	Lowell J. Reed (biometrician)
1944	First Edition	*Mathematical Tables and Other Aids to Computation* (Vol. 1, No. 8, pp. 360-361)	William G. Cochran (statistician)
1949	Third Edition	*Mathematical Tables and Other Aids to Computation* (Vol. 3, No. 25, pp. 316-320)	William G. Cochran (statistician)
1950	Third Edition	*The Quarterly Review of Biology* (Vol. 25, No. 1, p. 125)	Alphonse Chapanis (psychologist)
1953	Fourth Edition	*Journal of the Royal statistical Society (Series C)* (Vol. 2, No. 3, p. 203)	Leonhard H. C. Tippett (statistician)
1956	Fourth Edition	*Econometrics* (Vol. 24, No. 1, p. 90)	Herman O. Hartley (statistician)
1958	Fifth Edition	*Journal of the Royal statistical Society (Series A)* (Vol. 121, No. 2, p. 240)	Keith D. Tocher (statistician)

1958	Fifth Edition	*The Incorporated Statistician* (Vol. 8, No. 4, pp. 193-194)	W. R. B. (?)
1958	Fifth Edition	*Canadian Journal of Comparative Medicine* (Vol. 22, No. 1, p. 8)	Anonymous
1963	Sixth Edition	*Review of the International Statistical Institute* (Vol. 31, No. 3, p. 449)	G. Goudswaard (statistician)
1964	Sixth Edition	*The British Medical Journal* (Vol. 1, No. 5376, p. 172)	Ian David Hill (statistician)
1965	Sixth Edition	*Journal of the Royal Statistical Society Series A* (Vol. 128, No. 1, p. 146)	G. B. Wetherill (statistician)
1965	Sixth Edition	*Biometrische Zeitschrift* (Vol. 7, No. 2, pp. 124-125)	Anonymous
1971	Sixth Edition	*Biometrische Zeitschrift* (Vol. 13, No. 4, p. 285)	H. Toutenburg (statistician)

APPENDIX 2.3: Selected requests for reprinting materials from the *Statistical Tables for Biological, Agricultural and Medical Research.*

The figures presented in Table I and II are a qualitative evidence of the requests for reprinting materials received by the authors and the publisher of the *Statistical Tables*. They cannot be considered a comprehensive list of all the applications received, nonetheless they are suggestive of the diffusion in different disciplines of the tables originally published in the Oliver and Boyd's collection.

The information used for filling in the tables is extracted from the Oliver and Boyd Papers at the National Library of Scotland, from the Fisher-Yates correspondence held at Rothamsted Research and at the Barr Smith Library, University of Adelaide. In these sources I have counted over two hundreds requests, but in some cases only the name of the applicant is reported. Once excluded these doubtful cases, one hundred and sixty-seven requests were left and they represented the data I used in preparing Table I and II. In Table I, I have counted more than once the requests that were explicitly addressed to different fields. For instance, the request for reprinting tables in the book *Elementary Statistical Methods in Psychology and Education* has been counted in both education and psychology. In making the list of requests I have considered not only the letters of application, but also the cover letters for complimentary copies of the books in which materials from the *Statistical Tables* had been reprinted and that were sent to the authors. When possible, data related to published books have been checked against library catalogues.

a. Distribution of the requests per discipline

Discipline	
Agriculture, Botany and Plant Breeding	9
Biology [Including Anthropometry and Biometry]	21
Chemistry and Research Methods	4
Economics, Management and Accounting	16
Education	8
Engineering, Quality Control, Operations Research	21
Industrial Research and Applications	22
Medicine, Psychiatry, Pharmaceutics	8
Meteorology	1
Military Sector	10
Psychology	13
Sociology	5
Statistics, Mathematics and other collections of tables	33

b. Distribution of the requests per year

Time Interval	Number of requests
1939-1945	8
1946-1950	39
1951-1955	46
1956-1960	74

c. Sample requests

Date	Applicant	Table(s) requested	To be published in
December, 1939	Houghton Mifflin Company, on behalf of Everett F. Lindquist, Professor of Education, the State University of Iowa	Random numbers.	*Statistical Analysis in Educational Research* (1940), Houghton Mifflin Company
September, 1940	W. S. Robinson, Instructor in Sociology, Columbia University	Distribution of chi-square, random numbers (all abridged).	*Statistical Inference in Social Research* (booklet to be mimeographed for experimental class work)
October, 1941	Kenneth Mather, Geneticist	Normal deviate, Student's distribution, distribution of chi-square, variance ratio (all abridged).	*Statistical Analysis in Biology* (1943), Methuen
March, 1948	Imperial Chemical Industries Ltd	Square roots.	Book on analytical methods issued for internal use
May, 1948	Quinn McNemar, Professor of Psychology, Statistics and Education, Stanford University	Student's distribution, distribution of chi-square, distribution of z and the variance ratio, transformation of r to z (all abridged).	*Psychological Statistics* (1949), John Wiley and Sons
October, 1948	David. J. Finney, Lecturer in Design and Analysis of Scientific Experiments, Oxford University	Student's distribution, distribution of chi-square, distribution of z and variance ratio (all abridged).	Revised edition of *Biological Standardization* by J. B. Burn (1950), Oxford University Press
November, 1948	Frank Yates	Random numbers (abridged).	*Sampling Methods for Censuses and Surveys* (1949), Charles Griffin and Co.
September, 1949	Alan Brearley and David R. Cox	Random numbers.	*An outline of statistical methods for use in the textile industry* (1949), Wool

			Industries and Research Association, Torridon, Headingley, Leeds
November, 1949	A. Trevor Haynes, Institute of Actuaries, Glasgow	Student's distribution (abridged).	Intermediate textbook on statistics and a collection of tables for use of the candidates in the examination room
November, 1949	Robert R. Race	Distribution of chi-square (abridged).	*Blood Groups in Man* (1950), Blackwell Scientific Publications
December , 1949	Howard A. Boyle, Professor of Biology, George Fox College, Newberg, Oregon	Distribution of chi-square.	*Laboratory Exercises in Genetics* (Privately printed publication)
Undated, 1949	His Majesty's Stationery Office	Various tables in abridged form.	*Industrial Experimentation*, booklet printed by H.M.S.O
March, 1950	Robert L. Petersen, Department of Psychology, University of Miami	Student's distribution, distribution of chi-square, distribution of z and the variance ratio.	Request to mimeograph one hundred copies of the mentioned tables for a class in psychology
July, 1950	Donald Mainland, Dalhousie University, Halifax	Student's distribution, distribution of chi-square (all abridged).	*Elementary Medical Statistics* (1952), W. B. Saunders Company NB. The table requested are from *Statistical Tables*, but they are numbered wrongly because the author uses the scheme adopted in *Statistical Methods for Research Workers*, in which the table of the chi-square and Student's distributions are swapped over
August 1950	Egon S. Pearson, Professor of Statistics, University College London	Student's distribution (abridged) and corresponding values for r, certain intervals in the distribution of z and the variance ratio, scores for ordinal or ranked data.	*Biometrika Tables for Statisticians* (1954), Cambridge University Press
September, 1950	Oscar Kempthorne, Statistics Department, Iowa State College	Distribution of z and the variance ratio (abridged).	*The Design and Analysis of Experiments* (1952), Wiley
November, 1950	David J. Finney	Student's distribution (abridged), distribution of chi-square,	*Statistical Method in Biological Assay* (1952), Charles Griffin and Co.

		distribution of *z* and variance ratio (abridged), fiducial limits for a variance component (abridged), complementary LogLog transformation.	
December, 1950	Cedric A. B. Smith, Galton Laboratory, University College London	Student's distribution and the variance ratio (all abridged). [The significant values of the correlation coefficient and the distribution of chi-square are instead reproduced from *Statistical Methods for Research Workers*]	*Biomathematics* (1954) [main author W. M. Feldman] (revised edition by C.A.B Smith), Hafner Publishing
April, 1951	Edwin L. Crow, US Naval Ordnance Test Station, California	Normal distribution, Student's distribution and orthogonal polynomials (all abridged).	Outline manual of statistical methods and tables
June-July 1951	Silvio Vianelli Istituto di Statistica, Università di Palermo	Example 7, Latin squares (abridged), balanced incomplete blocks (abridged), orthogonal polynomials (references only).	*Metodologia statistica delle scienze agrarie, 1954 (vol. 1), 1956 (vol. 2), Edizioni Agricole Bologna* (in Italian)
July 1951	Chester I. Bliss, Connecticut Agricultural Experiment Station	Student's distribution, scores for ordinal (or ranked) data (abridged), scores for ordinal data – sums of squares of mean deviations tabulated (abridged).	*An Outline of Biometry* [with D. W. Calhoun] (1954), Yale Co-operative Corp.
September, 1951	Merle W. Tate, Associate Professor of Education, University of Pennsylvania	Student's distribution and correlation coefficient (all abridged). [Actually the significance values for the correlation coefficient (abridged) were taken from *Statistical Methods for Research Workers*] [The values for the distribution of chi-square and the distribution of the variance ratio are taken instead from the new tables recomputed in *Biometrika* (1941, 1943) by E. S. Pearson and his	*Statistics in Education* (1955), Macmillan Company

		co-workers]	
October, 1951	Masami Ogawara, Meteorological Research Institute, Tokyo	Latin squares.	*Diagrams and Tables for Statistical Analysis*
October, 1951	R. L. Anderson, Institute of Statistics, State College Station, University of North Carolina (co-author with Professor T. A. Bancroft, Iowa State College)	Student's distribution (abridged).	*Statistical Theory in Research* (1952), McGraw-Hill
Undated, 1951 or early 1952	W. L. Gross, London	Normal curve, student's distribution (abridged), distribution of chi-square (abridged), distribution of *z* and the variance ratio, correlation coefficient.	Pamphlet entitled *The Fundamentals of Statistical Analysis* to be published by the Association of Engineering and Shipbuilding Draughtsmen
May, 1952	R. Cavé, Paris, Ingenieur militaire en chef de l'Armement	Ordinates of the normal distribution, Student's distribution, distribution of chi-square, distribution of *z* and the variance ratio, significance of difference between two means, probits, factorials (abridged), random numbers (abridged).	*Le Controle Statistique Des Fabrications* (1952), Eyrolles Editeur-Paris (Book about quality control and techniques of statistical analysis)
June, 1952	C. H. L. Oliver, Director of Naval Ordnance, Admiralty Bath	Factorials, Student's distribution, distribution of chi-square, correlation coefficient.	*The Planning and Conduct of trials*, part of *The Gunnery Manual*
June, 1952	Carl A. Bennett, Statistics Unit, Nucleonic Division, Hanford Works, Washington (co-author with Norman L. Franklin, University of Leeds)	Distribution of *z* and the variance ratio (abridged). [Despite the request made to Fisher and Yates and the positive answer of the authors all the tables reprinted in the book (distribution of chi-square, Student's distribution, variance ratio) are taken from the set of	*Statistical Analysis in Chemistry and the Chemical Industry* (1954), Wiley

		tables recomputed *a-la*-Fisher by E. S. Pearson and his co-workers and printed in *Biometrika* during the 1940s.	
June, 1952	Mamoru Nakamoto, Japanese Standards Association, Tokyo	Distribution of chi-square.	*Method of Estimation and Test of Significance by Attributes*
August, 1952	A. J. Skey, Canadian Armament Research and Development Establishment	Distribution of chi-square, Student's distribution, random numbers, variance ratio (all abridged).	Paper on the use of statistical methods in the analysis of armament firing trials. To be published as a classified report for workers in defence.
October, 1952	J. Pedoe, Department of Mathematics, the Technical College, Sunderland	Ordinates of the normal distribution (abridged).	Book for Higher National Certificate Students
December, 1952	Irwing W. Burr, Professor of Mathematics and Research Associate in the Statistical Laboratory, Purdue University, Indiana	Student's distribution. [Actually the table reproduced in Burr's book is the table of chi-square]	*Engineering Statistics and Quality Control* (1953), McGraw-Hill
February, 1953	Acheson J. Duncan, Associate Professor of Statistics, School of Engineering, John Hopkins University	Distribution of chi-square (abridged).	Paper on the chi-square test to be given at the National Convention of the American Society for Quality Control
February, 1953	Myra T. Grenier, Librarian Aerojet Engineering Corporation, California	Student's distribution, distribution of z and the variance ratio.	*Aerojet Design Handbook*, publication mainly for internal circulation
May, 1953	Helen M. Walker, Professor of Education, Teachers College Columbia. Co-author with Joseph Lev, senior statistician, Department of Civil Service State of New York	Student's distribution, distribution of chi-square*, distribution of z (all abridged). [*The main part of the chi-square table is reprinted from *Biometrika* (1941), and only three columns are taken from the table in Fisher	*Statistical Inference* (1953), Holt, Rinehart and Winston

		and Yates' collection]	
March, 1957	Virginia L. Senders, Department of Psychology University of Minnesota	Student's distribution (abridged). [The distribution of chi-square is borrowed from Pearson & Hartley, *Tables for Statisticians*, while the distribution of the variance ratio is reprinted from *Biometrika*]	*Measurement and Statistics* (1958), Oxford University Press
April, 1957	Alva M. Tuttle, Associate Professor of Economics, the Ohio State University	Student's distribution (abridged).	*Elementary Business and Economic Statistics* (1957), McGraw-Hill
June 1957	Samuel B. Richmond	Student's distribution.	*Principles of Statistical Analysis* (1957), Ronald Press
September, 1957	Thomas L. Saaty, Office of Naval Research, Maryland	Normal distribution. Student's distribution, distribution of chi-square, distribution of *z* and the variance ratio, random numbers (all abridged).	*Mathematical Methods of Operations Research* (1959), McGraw-Hill
September, 1957	Bernard W. Lindgren, Department of Mathematics, G. W. McElrath , Department of Mechanical Engineering, Institute of Technology, University of Minnesota	Ordinates of the normal distribution. [The request is for the table of the normal distribution, but actually, reprinted in the book is Student's distribution (abridged)] The chi-square distribution is reprinted instead from Pearson and Hartley's *Biometrika Tables for Statisticians*]	*Introduction to Probability and Statistics* (1959)
April, 1960	George G. Simpson, Anna Roe and Richard C. Lewontin	Student's distribution, values of the correlation coefficient for different levels of significance. [Preface and table caption mention that the materials on the correlation coefficient are taken from table VI, while the table actually reproduced in the book	*Quantitative Zoology,* rev. edition (1960), Harcourt, Brace and Co.

		is table VII from Fisher and Yates' book]	
May, 1960	Karl F. Schuessler, Indiana University, Department of Sociology (co-author with John H. Mueller, Professor of Sociology, Indiana University)	Distribution of chi-square.	*Statistical Reasoning in Sociology* (1961), Houghton Mifflin Company
June, 1960	Georg H. Zimny, Professor of Psychology, Marquette University	Student's distribution, random numbers (all abridged).	*Method in Experimental Psychology* (1961), Ronald Press Company
August, 1960	Eastman Kodak Company, Rochester Institute of Technology	-----	*Symbols, Definitions and Tables for Industrial Statistics and Quality Control*
October, 1960	Howard W. Alexander, Earlham College, Richmond, Indiana	Student's distribution (abridged).	*Elements of Mathematical Statistics* (1961), John Wiley & Sons
November, 1960	R. R. Umarji, MSC, Principal, Government College, Mangalore, and Professor of Mathematics, Bombay Educational Service, India	Student's distribution, distribution of chi-square (all abridged).	*Probability and Statistical Methods* (manual for censuses) (1962), Asia Publishing House

Archive materials

Barr Smith Library, the University of Adelaide

R. A. Fisher Papers (digitized)

- Letter from R. A. Fisher to G. S. Phillpotts, 7th July 1939 [accessed 25th October 2012 http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67919/3/1939-07-07.pdf].
- Correspondence between R. A. Fisher and A. C. Aitken (April 1931) [accessed 25th October 2012 http://digital.library.adelaide.edu.au/dspace/handle/2440/10706/].
- Letter from R. A. Fisher to M. V. Wilkes, 16th February 1949, [accessed 30th October 2012 http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68111/8/1949-02-16.pdf].
- Letter from R. A. Fisher to M. V. Wilkes, 27th April 1950, [accessed 30th October 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68111/6/1950-04-27.pdf].
- Letter from R. A. Fisher to M. V. Wilkes, 26th September 1950, [accessed 30th October 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68111/5/1950-09-26.pdf].
- Letter from R. A. Fisher to M. V. Wilkes, 2nd October 1950, [accessed 30th October 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68111/4/1950-10-02.pdf].
- Letter from R. A. Fisher to M. V. Wilkes, 27th November 1953, [accessed 30th October 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68111/2/1953-11-27.pdf].
- Correspondence between R. A. Fisher and M. J. M. Healy (March-April 1956) [accessed 25th October 2012, http://digital.library.adelaide.edu.au/dspace/handle/2440/67738].

Bodleian Libraries, University of Oxford

- The British Association Mathematical Tables Committee, 1871-1948, Minute Book 1928-1930, Ms. Eng. Min. d 1157.
- The British Association Mathematical Tables Committee, 1871-1948, Minute Book 1930-1935, Ms. Eng. Min. d 1158.

National Library of Scotland

- Letter from Oliver and Boyd to R. A. Fisher, 21st August 1923, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 240.
- Letter from Oliver and Boyd to R. A. Fisher, 25th August 1923, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 240.
- Letter from R. A. Fisher to Oliver and Boyd, 23rd May 1924, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 240.
- Letter from Oliver and Boyd to R. A. Fisher, 27th May 1924, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 240.
- Letter from R. A. Fisher to Oliver and Boyd, 29th May 1924, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 240.
- Letter from Oliver and Boyd to R. A. Fisher, 25th June 1925, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 253.
- Letter from Oliver and Boyd to R. A. Fisher, 19th March 1937, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 347.

- Letter from R. A. Fisher to Oliver and Boyd, 20th May 1937, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 347.
- Letter from F. Yates to Oliver and Boyd, 4th October 1937, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 356.
- Letter from F. Yates to Oliver and Boyd, 5th May 1938, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 369.
- Letter from F. Yates to Oliver and Boyd, 12th August 1938, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 369.
- Letter from Oliver and Boyd to F. Yates, 19th August 1938, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 369.
- Letter from F. Yates to Oliver and Boyd, 25th August 1938, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 369.
- Letter from Oliver and Boyd to R. A. Fisher, 26th August 1938, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 360.
- Letter from Oliver and Boyd to R. A. Fisher, 14th February 1939, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 373.
- Letter from R. A. Fisher to Oliver and Boyd, 16th February 1939, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 373.
- Letter from R. A. Fisher to Oliver and Boyd, 15th June 1939, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 373.
- Letter from R. A. Fisher to Oliver and Boyd, 4th November 1940, Oliver and Boyd Collection, Acc.5000/Roneo System/Box 386.
- Letter from R. A. Fisher to Oliver and Boyd, 21st June 1943, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 418.
- Letter from R. A. Fisher to Oliver and Boyd, 3rd July 1944, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 428.
- Letter from Oliver and Boyd to R. A. Fisher, 6th July 1944, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 428.
- Letter from R. A. Fisher to Oliver and Boyd, 15th July 1944, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 428.
- Letter from Oliver and Boyd to F. Yates, 6th November 1945, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 441.
- Letter from R. A. Fisher to Oliver and Boyd, 16th January 1952, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 600.
- Letter from F. Yates to Oliver and Boyd, 11th August 1953, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 650.
- Letter from F. Yates to Oliver and Boyd, 15th November 1962, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 949.
- Letter from Oliver and Boyd to F. Yates, 10th June 1963, Oliver and Boyd Collection, Acc. 5000/Roneo System/Box 980.

Rothamsted Research, Library and Archive

- Letter from F. Yates to R. A. Fisher, 23rd April 1931, STF 31 (old catalogue reference).
- Letter from F. Yates to R. A. Fisher, 13th May 1936, STATS 7.5.
- Letter from F. Yates to R. A. Fisher, 29th May 1936, STATS 7.5.
- Letter from F. Yates to R. A. Fisher, 2nd June 1937, STATS 7.5.

- Letter from W. G. Cochran to F. Yates, 23rd June 1937, STATS 6.5.
- Letter from F. Yates to R. A. Fisher, 10th July 1937, STATS 7.5.
- Letter from R. A. Fisher to F. Yates, 15th July 1937, STATS 7.5.
- Letter from F. Yates to R. A. Fisher, 27th July 1937, STATS 7.5.
- Letter from F. Yates to R. A. Fisher, 30th July 1937, STATS 7.5.
- Letter from F. Yates to D. J. Finney, 25th November 1946, STF 31 (old catalogue reference).
- Letter from D. J. Finney to F. Yates, 27th November 1946, STF 31 (old catalogue reference).
- Letter from H. W. Norton to F. Yates, 6th February 1947, STF 31 (old catalogue reference).
- Letter from F. Yates to H. W. Norton, 6th February 1947, STF 31 (old catalogue reference).
- Letter from Chester I. Bliss to F. Yates, 28th January 1948, STF 31 (old catalogue reference).
- Letter from D. J. Finney to F. Yates, 6th November 1950, STF 31 (old catalogue reference).
- Letter from P. M. Dance to R. Hunt, 23rd February 1953, STF 31 (old catalogue reference).
- Letter from R. A. Fisher to F. Yates, 29th May 1956, STATS 7.5.
- Letter from R. A. Fisher to F. Yates (with enclosed note), 5th June 1956, STATS 7.5.
- Letter from R. A. Fisher to F. Yates, 28th June 1961, STATS 7.5.

Royal Statistical Society Archive

Royal Statistical Society Nomination Paper, 4999: Frank Yates, elected 21st November 1933.

Chapter 3

STATISTICS, COMPUTING AND INFORMATION TECHNOLOGIES IN THE STUDY OF THE ABO BLOOD GROUPS

3.1 *Introduction*

At the beginning of the twentieth century, the physician Karl Landsteiner discovered the first human blood group system, conventionally indicated as ABO. Landsteiner's discovery proved to be of interest in medicine, for the safety of blood transfusion; in the study of human inheritance, because blood groups were the first genetic marker; in anthropology, as blood groups are present with different proportions in different populations. Following Landsteiner's discovery, more blood groups systems were identified and today thirty major ones are known.[428]

In the understanding of human blood groups and in their application to medicine, genetics and anthropology, statistics offered a major contribution. Blood groups, in fact, are inherited according to Mendel's laws and are "genetically determined at conception and remain fixed for life", but serological tests enable only to group people according to their phenotype and "cannot as a rule determine the genotype of an individual".[429] Statistical methods applied to blood groups data, instead, allow to "calculate the frequencies of the genes and of the genotypes in the population from the frequencies of the phenotypes".[430]

In the 1920s the mathematician Felix Bernstein, comparing expected and observed frequencies of the ABO blood groups, proved the correct hypothesis on their inheritance and provided the mathematical formulae for estimating the frequencies of the genes A, B, O. Analogous formulae were needed for all the blood group systems that were gradually discovered and statisticians constantly acted as consultants of serologists, not only providing mathematical formulae, but often also computing gene frequencies and comparing expected and observed values on their behalf.

If statistics was a key tool in the study of human blood groups, and number crunching an unavoidable corollary, data collection was a necessary precondition, not free from obstacles and tiring routines. The blood groups data employed in the statistical analysis were derived from both family materials and from population studies. A third potential, but unintended, source of data

[428] G. Daniels and I. Bromilow (2010).
[429] A. E. Mourant (1954), p. 2; p. 4.
[430] A. E. Mourant (1954), p. 4.

for the ABO system and later also for the Rh system was represented by the records of the transfusion services that began to be organised since the 1920s.

I want to discuss the interplay of statistics, computing and information technologies for data collection in the study of human blood groups, and the demarcation of mathematical and laboratory expertise in serology, examining the long-lived collaboration between the statistician and geneticist Ronald Fisher and the serologists of the Galton Laboratory Serum Unit that he set up in 1935, with funding provided by the Rockefeller Foundation, while he was professor of eugenics and head of the Galton Laboratory at University College London.

Fisher's official connection with the unit was severed in 1939 at the outbreak of WWII, but throughout his life he never lost touch with his former co-workers, both during the warfare and in its aftermath, and constantly acted as their statistical consultant in the study of human blood groups. I am not going to write a comprehensive account of this collaboration, but in the present chapter I will rather focus on a specific episode, the survey of the ABO blood groups in Britain undertaken by Fisher during WWII in unofficial co-operation with the Galton Laboratory Serum Unit and using the records of the Emergency Transfusion Services.

The chapter is arranged in six sections. I will begin introducing the role of the ABO blood groups in the study of human heredity and describing the creation of Ronald Fisher's serum unit in the 1930s and the dispersion of the Galton Laboratory staff during WWII. In relation to the ABO blood group survey, which is the main part of the chapter, I will examine at first the management of the donor records adopted by the Emergency Transfusion Services and the previous practices which inspired it, employed by the British Red Cross in the administration of its voluntary panel of blood donors. In particular, I will deal with the enrolment forms and index cards that prompted the survey and I will discuss the role that Fisher had as recordkeeper, ensuring the preservation and the non-duplication of the data extracted from the transfusion records.

I will use a double frame for my analysis of Fisher's survey. On the one hand I am going to account the employment of statistical expertise in wartime Britain, Fisher's failure to take part with his staff in computing work for the warfare and his gradual recognition of the computational and statistical labour involved in the blood group survey as his work of national importance. On the other hand I am going to examine how Fisher's statistical outlook contributed to the survey. In particular I will examine how the hundreds of thousands of medical records of the Emergency Transfusion Services were transformed in genetical data through the application of Fisher's statistical methods for determining the frequencies of the genes A, B, O. Even though the anthropological and ethnological dimension of the survey is not of direct concern for my argument, I will sketch the outcome of Fisher's work and how it can be set in the context of the previous researches on the distribution of the ABO blood groups in the human population.

The final point that I will discuss in the chapter is Fisher's lifetime collaboration with serologists. I want to examine the unbalance between Fisher's mathematical expertise and the one of the serologists that he advised. If Fisher was amused by the "evasive action" taken by serologists "when anything heavy in the way of mathematics seems imminent", the latter were relieved by the fact that the study of blood groups certainly required statistics, but the methods were "most fortunately simple to apply" and demanded "no understanding of their subtlety".[431] On the other hand, Fisher needed the advice of his fellow serologists when technical matters affected the quality of the data gathered for statistical analysis. Thus, the collaboration was a mutual exchange of expertise rather than a one-way adoption of statistical tools.

In conclusion, in relation to my case study I will re-assess the role of mathematical tools, information management and practical computation in serology and I will discuss the demarcation between statistical knowledge and laboratory practice in the study of the ABO blood groups.

3.2 *The ABO blood groups in the study of human heredity*

During the first half of the twentieth century in the study of human genetics blood groups had a role comparable to the one that DNA gained in the second half of the century.[432] They were an ideal marker for the mapping of the human chromosomes, they were independent from "age, disease and the influence of other genes in the body", and intermediate forms were extremely rare.[433] Blood groups were also a powerful tool for anthropology because their distribution changed from population to population and such changes could be accounted in terms of geography and human migrations, as proved for the first time by the serologists Ludwik and Hannah Hirszfeld during WWI. The ABO and Rh blood groups were also relevant in blood transfusion, as the incompatibility of donor and recipient, resulted in harmful consequences, even the death, for the recipient.

The ABO blood group system was the first discovered and studied. It consists of two possible antigens – A and B – on the red cells and two possible antibodies in the serum – α (anti-A) and β

[431] R. R. Race and R. Sanger (1954). Quotations p. *vi*, p. 8.
[432] W. H. Scheneider (1996a), p. 275. For an assessment of the role of blood groups in the study of human genetics and anthropology see also the contributions of G. Garratty (1996) and J. Marks (1996) published alongside Scheneider's work in the monographic issue on blood groups of the journal *History and Philosophy of the Life Sciences*.
[433] R. R. Race and R. Sanger (1954), pp. 8-9.

(anti-B). The antibodies are also called agglutinins because they provoke the clumping of red cells when a testing serum with the corresponding antibody is added to a suspension of red cells. Four phenotypes are possible A, B, AB and O.[434] People belonging to the blood group A have anti-B in their serum, which agglutinates B cells, the reverse is true for people belonging to the B group. People of group AB have both antigens A and B on the red cells and no antibody in their serum, while people of group O, on the contrary, have no antigens on their red cells and both anti-A and anti-B antibodies in their serum.

Therefore, in blood transfusion people of blood group A and B can receive blood only from people of the same group or from O donors, people from the group AB can receive blood from donors of all the blood groups, while people of blood group O that have no antigen on their red cells can receive blood only from same group donors, but can give blood to all the others.[435] For this reason, O donors were especially appreciated during the warfare, as their blood could be transfused in emergency conditions, when it was not possible to check beforehand the blood group of the receiver.

In 1910 the serologists Emil von Dungern and Ludwik Hirszfeld proved that the inheritance of the ABO blood groups followed Mendel's laws and that the characters A and B were linked to dominant genes, while O to a recessive factor. They suggested also a system of inheritance for the ABO blood groups based on two independent pairs of genes. Inconsistencies in this scheme (Appendix 3.1) were pointed out in the 1920s by the mathematician Felix Bernstein, head of the Institute for Mathematical Statistics at Göttingen.[436] Bernstein explained the inheritance of the ABO blood groups supposing not two independent pairs of genes, but three alleles at one locus, and in order to tackle the mathematical problem he adopted the Hardy-Weinberg formula for the equilibrium of two alleles in a population and extended it to the case of three alleles.[437] He then made use of the several population and family materials available by mid-1920s to support his hypothesis against von Dungern and Hirszfeld's theory. Population studies offered a good

---

434 In 1930 O. Thomsen, V. Friedenreich and E. Worsaae proved that there was a bipartition of the blood group A in A_1 and A_2, the former dominant to the latter. Thus the phenotypes of the ABO system are not four, but six: A_1, A_2, B, A_1B, A_2B, O. Blood transfusion, however, is not influenced by the difference between A_1 and A_2 donors and thus the donor records that I am going to discuss at length were arranged and managed in relation to the four phenotypes A, B, AB and O.

435 This is a simplified description of the incompatibilities in blood transfusion neglecting the Rh factor of the donor and the recipient. The Rh blood group system was discovered in 1940 by the serologists Karl Landsteiner and Alexander Wiener immunizing rabbits and guinea pigs with the blood of the monkey *Macacus Rhesus*. Landsteiner and Wiener discovered that the resulting antibodies could also agglutinate human red cells and they called Rh positives the cases where agglutination occurred and Rh negative the others. In blood transfusion a Rh negative donor can give blood to a Rh positive recipient – compatible for the ABO group – without consequences, while if a Rh negative donor receives Rh positive blood, he or she can develop Rh antigens.

436 For more biographic information on Bernstein see H. Nathan (1970) and N. Schappacher (2005). An obituary of Bernstein (in Italian) is C. Gini (1957). Bernstein's statistical work in blood group genetics is accounted in P. M. H. Mazumdar (1995), chapter 14; P. M. H. Mazumdar (1996), and H. O. Lancaster (1994), pp. 53-57.

437 For the Hardy-Weinberg law see W. B. Provine (2001), pp. 131-136.

agreement between observed and expected frequencies according to Bernstein's hypothesis, while the family materials presented some exceptions that were blamed over technical mistakes, illegitimacy, or peculiarities of the red cells or the testing sera.[438]

Pauline Mazumdar has observed that Bernstein's calculations of the gene frequencies were guided by two assumptions: a one-to-one relationship between gene and antigen and a one-to-one relationship between antibody and antigen. To the mathematician Bernstein the convenience of such choices should not have gone unnoticed as they allowed a straightforward reduction of a complex system into a manageable set of expectations for gene frequencies. Mazumdar comments that precisely such assumptions made blood group genetics amenable to statistical treatment and stirred up "a wave of Mendelian algebra [...] over blood group serology, as the triple-allele hypothesis was debated in the world literature".[439] In less than a decade Bernstein's theory for the ABO blood groups and his mathematical analysis were generally accepted and adopted to extract the frequencies of the genes A, B, O from the phenotype data available.[440]

The acquisition of data on blood groups, however, was limited by the difficulties of fieldwork. In anthropology blood group data should be gained in large quantity from a random sample of the population – hundreds or thousands of data were usually necessary – and it should be possible to compare the values of gene frequencies all over the globe. The pioneering research of Ludwik and Hannah Hirszfeld on the distribution of the ABO blood groups in different populations was made possible only by the events of WWI that blocked for two years a multiracial contingent of soldiers in Salonica offering the Hirszfelds an unexpected opportunity to test several racial groups at once.[441]

Constraints were also evident in the work of the Galton Laboratory Serum Unit before the mobilization for the national emergency. Only a few hundreds people were tested and the choice of the subjects was mainly due to chance, proximity and feasibility. Many people approached by the unit, in fact, were workers and students at University College London, their families and acquaintances.[442] The investigation of heritable conditions through blood group genetics was also complex because the serologists had to travel extensively to collect from the family members

438 F. Bernstein (1966b), pp. 123-124. Bernstein (1966a, 1966b) are the English translations of Bernstein's 1924 and 1925 papers on the ABO blood groups originally published in German.
439 P. M. H. Mazumdar (1995), p. 299.
440 P. M. H. Mazumdar (1995), pp. 299-300.
441 For the pioneering research of Ludwig and Hannah Hirszfeld see W. L. Schneider (1996b) pp. 280-282.
442 G. L. Taylor and A. M. Prior (1938a), 344.

blood samples and possibly test them for other inherited abilities, such as the ability to taste phenyl-thio-carbamide.[443]

The constraints experienced in the research on blood groups are a recurrent issue in human genetics because records in this field are often fragmentary and limited in their number, as human beings have few progenies and information on inherited conditions can be gathered directly only for three or four generations. Due to these limitations Phillip Thurtle has pointed out that "classical genetics was a science of record keeping. [...] Records needed to be kept over time, and an organized form of access and retrieval needed to be instituted" in order to study inheritance.[444] Thurtle mentions the Eugenics Record Office started by Charles Davenport at Cold Spring Harbor as "the best example of the application of recordkeeping innovations to the study of human heredity", and emphasises Davenport's choice to invest in advanced office equipment and borrow from the twentieth century business companies the then innovative strategies of information management.[445]

The connection suggested by Thurtle between information technologies, information management and the fostering of genetic knowledge suits well my narrative of the ABO blood group survey undertaken by Ronald Fisher. In my account I emphasise the materiality of the records adopted by the Emergency Transfusion Services and the strategies of information management employed for blood transfusion in Britain as decisive elements for the study of human biological diversity. Before discussing the blood group survey, however, I will sketch the involvement of Ronald Fisher in serology and the constitution of his serum unit at the Galton Laboratory.

3.3 *Galton Laboratory Serum Unit, 1935-1939*

Sponsored by Francis Galton, founder of eugenics and promoter of its statistical twist, the Galton Laboratory for the Study of National Eugenics started in 1904 as Eugenics Record Office. In 1907 it became a proper laboratory not only for the collection of data, but also for their statistical analysis under the auspices of the Senate of University College London and the supervision of the statistician Karl Pearson, already responsible in the same institution for the Biometric Laboratory, a centre for statistical training and computing work.[446] In 1911, at the

443 An interesting example is the collection of blood group data on Friedereich's ataxia, an inherited condition of the spinal cord, organised in the 1930s by the British geneticist Lancelot Hogben and described in P. M. H. Mazumdar (2011), p. 175.
444 P. Thurtle (2007), p. 3.
445 P. Thurtle (2007), p. 290.
446 D. J. Kevles (1995), pp. 3-40; T. M. Porter (2006).

death of Francis Galton, University College received the bequest of his estate to support the work of the laboratory that he had created and to establish a new professorship in eugenics for Karl Pearson. In the same year the Biometric and Galton laboratories became the two components of the department of applied statistics, under Karl Pearson's directorship.

Research in eugenics and human genetics, as the two were synonym in the first half of the twentieth century, was the primary goal of the Galton Laboratory.[447] Its tasks, as envisioned by Francis Galton, were to "act (i) as a storehouse for statistical material bearing on the mental and physical conditions to inheritance and environment, (ii) as a centre for the publication or other form of distribution of information concerning National Eugenics".[448] To the fulfilment of this mission contributed on the one hand the pedigrees prepared by the workers in the laboratory and on the other its publications, *The Treasury of Human Inheritance* and the *Annals of Eugenics.*

Under Pearson's management the Galton Laboratory not only applied statistical methods to the study of data on human inheritance, but became also a centre for statistical teaching and computation, and notable was the contribution that the Galton and Biometric laboratories gave to computing work of national importance during WWI.[449]

Despite Galton's bequest, the budget of the laboratory was tight. In the 1920s Pearson complained that "the war deducted at least 50% from the real value of our income. The result has been that there are no funds for the upkeep of the Museum or of the Anthropometric Laboratory, or for carrying on by field-workers any form of social investigation" and, above all, he regretted the policy of University College that did not provide further funding as "we are not a paying department, because we have only postgraduate students".[450]

In 1933, at Karl Pearson's retirement, the department of applied statistics was split in two independent units, statistics and eugenics, the latter assigned to the statistician, geneticist and eugenicist Ronald Fisher, who promoted the creation of the Galton Laboratory Serum Unit. Before his election to the Galton professorship, Ronald Fisher had been chief statistician at Rothamsted Experimental Station, where he developed the statistical methods of analysis of variance and experimental design for agricultural and biological research (**chapter 1**). At Rothamsted Fisher had also the opportunity to pursue his interests in genetics and eugenics; he

447 On the reciprocal relation of eugenics and human genetics see P. M. H. Mazumdar (2011), p. 58.
448 National Eugenics, as conceived by Galton was, "the study of the agencies under social control that may improve or impair the racial qualities of future generations either physically or mentally". Both the quotations in the text and in the footnote are from the Appeal Fund, 1907-1920, UCL Record Office, AR 179, Folder 4.
449 For the computing activity at the Galton Laboratory see D. A. Grier (2007), pp. 126-133. The mission of teaching statistical skills to research workers was established since 1907: "Short courses of instruction will be provided for those engaged in social, anthropometric, or medical work and desirous of applying modern methods of analysis to the reduction of their observations." (Appeal Fund, 1907-1920, UCL Record Office, AR 179, Folder 4)
450 Letter from K. Pearson to C. M. Gayley (dean of the American University Union), June 1925, RA, Box 16, Folder 218.

started his breeding experiments with mice, snails and poultry and in 1930 he published the *Genetical Theory of Natural Selection*, his main contribution to the birth of population genetics and a manifesto of his eugenic commitment.[451]

When he moved to University College, Fisher re-organised the research at the Galton Laboratory "to go beyond Pearson [...], to create a laboratory of mathematical genetics, with attention given to both words in the phrase [... and] made Mendelian genetics an intrinsic feature of work at the Galton".[452] The craniometric studies that were extensively conducted under Pearson were gradually abandoned and the collection of anthropometric data cut down to make space for Fisher's Mendelian work on mice, snails, grouse locusts and *Artemia Salina*, the brine shrimp.

Not all the research programme set up by Pearson was discarded. For instance, the preparation of pedigrees of rare anomalies was carried on by Julia Bell and the results published in the *Treasury of Human Inheritance*; Mary Karn, a researcher and computer in the Galton Laboratory since Pearson's time, undertook an extensive anthropometric research on the physical features of the British schoolchildren to be published in the *Annals of Eugenics*.

To the postgraduate students in the Galton Laboratory Fisher gave courses on the logic of experimentation and the genetics of quantitative characters. His lectures concerned the study of inheritance, population genetics and the teaching of the statistical methods useful in planning biological experiments. During his classes Fisher presented his own methods, analysis of variance and experimental design, which became the new statistical tools of the eugenics department, instead of the calculation of correlation coefficients, a hallmark of Pearson's statistics (**chapter 2**).[453]

The Galton Laboratory under Fisher lived on a tight budget as much as it had done previously. The expenses for the publications of the laboratory, the *Annals of Eugenics* and *The Treasury of Human Inheritance*, were consistent. "I see that you have done wonders with the publication fund, though I am embarrassed at having to undertake to do wonders after you" wrote Fisher to Karl Pearson, and to meet the costs of the *Annals of Eugenics* Fisher had to obtain an annual grant from the Eugenics Society.[454]

---

451 R. A. Fisher (1930). On Fisher and population genetics see in W. B. Provine (2001).
452 D. J. Kevles (1995), p. 201.
453 A brief account of research, teaching and finances in the Galton Laboratory under Fisher's management is enclosed in the application for funding sent to the Rockefeller Foundation in 1934. (Prospectus of the Eugenics Department, 1934, RA, Box 16, Folder 219)
454 Letter from R. A. Fisher to K. Pearson, 5th March 1934, R. A. Fisher Papers (digitized), BSL, The University of Adelaide. For the contribution of the Eugenics Society to the *Annals of Eugenics* see J. Fisher Box (1978), p. 262.

University College, including also the Galton bequest, contributed about £4,000 to the department of eugenics and did not want to add further funding to support new strands of research, again as "the income from fees in a purely research Department of this kind is almost entirely negligible".[455]

Money for further researches should be sought from external bodies and Fisher turned to the Rockefeller Foundation as a possible sponsor for his new research programme in serology.[456] Fisher's decision to approach the Rockefeller Foundation seems a natural choice: in the 1920s the Foundation had generously sponsored the re-organisation of the medical school at University College Hospital and in 1927 another consistent grant was given to the department of zoology at University College.[457]

However, by 1934 when Fisher drew up his proposal, the funding policy of the Foundation had radically changed compared to the previous decade. "The mainstay of the new system was the individual project grant. These were [...] systematic programs of investment, through individuals, in new fields of research [...]."[458] Fisher's plan to set up a unit for the investigation of the serological reactions liable of genetic interpretation fit well in the new funding policy. It was of limited duration – five years –, limited costs – thirty-five thousand dollars allocated in 1935 supplemented by other ten thousand dollars in 1937 – and certainly experimental, as pointed out by the Rockefeller's officers that reviewed Fisher's application.[459]

The research proposal emphasised the possibility to study the inheritance of mental defects through serological techniques, a suggestion that resonated with the psychobiology programme of the Rockefeller Foundation Medical Sciences Division and more broadly with the "sciences of man agenda" – ambiguously eugenics – endorsed by the Rockefeller Foundation Natural Sciences Division and by its influential director Warren Weaver.[460] Fisher hinted to its potential sponsors that

> the usefulness of serological methods would not be confined to purely medical investigations. Factors capable of detection in this way may indeed often exert a positively beneficial influence on health,

[455] Letter from A. Mawer to D. P. O'Brien, 26th October 1934, RA, Box 16, Folder 19.
[456] For the Rockefeller Foundation's involvement in science during the twentieth century see G. Gemelli et al. (1999) for the influence on medical research in Europe and R. E. Kohler (1991) for an overview of the Foundation's work with a US perspective.
[457] J-F. Picard and W. H. Schneider (1999), pp. 24-32.
[458] R. E. Kohler (1991), p. 233.
[459] For the money invested by the Rockefeller Foundation in Fisher's project see the Financial Prospectus, 7th December 1938, RA, Box 16, Folder 221. The experimental nature of Fisher's research programme in serology is discussed in various items of the internal correspondence of the Rockefeller Foundation, i.e. Memorandum from D. P. O'Brien to A. Gregg, 1st March 1935, RA, Box 16, Folder 220; Memorandum from W. E. Tisdale to W. Weaver, 5th March 1935, RA, Box 16, Folder 220.
[460] L. E. Kay (1993), pp. 41-50.

intelligence, artistic appreciation, sensory discrimination, longevity, etc., or at least on factors linked with these.[461]

Eventually, the Medical Sciences Division guided by Alan Gregg – not without some internal conflict with the Natural Sciences Division – granted Ronald Fisher the money to set up a small serum unit in April 1935.[462] The Galton Laboratory provided the workspace and "the services of Dr Fisher as director of the studies", while the Medical Research Council administered the grant on behalf of the Foundation.[463] Fisher set the research agenda of the unit and liaised with the Rockefeller Foundation and the Medical Research Council, but he did not have any technical competence in serology and his contribution to the work of the unit was only the statistical examination of the results.

In correspondence with Daniel P. O'Brien, Rockefeller officer in Paris, Fisher had presented his research proposal as "a great opportunity for giving to the subject [human genetics] a solidly objective foundation, under strict statistical control, thinking, of course, in terms of what the Galton Laboratory would be capable of undertaking."[464] However, if the expertise in statistics and genetics and the collection of pedigrees of human anomalies was certainly a prerogative of Fisher's laboratory, the technical competence in serology should be acquired from scratch, hiring new personnel for the laboratory research. This was a costly enterprise and the money provided by the Rockefeller Foundation was mainly spent for meeting the salaries of the new workers, rather than for the purchase of equipment.[465]

Fisher strongly requested to hire George L. Taylor as chief of the new unit. A painstaking research worker, Taylor had extensive practice in serology and constantly refined his laboratory techniques during the first years spent at the Galton Laboratory, visiting foreign institutions – for several months he worked with V. Friedenreich at the Universitetets Retsmedicinske Institut in Copenhagen – where research on blood groups was more advanced than in Britain.[466] The serum

---

461 New Scheme of Research in Serological Genetics (written by R. A. Fisher), 1934, RA, Box 16, Folder 220.

462 "Since sending you my memo on aid to the MRC for genetics heading towards the study of mental diseases, WET [W. E. Tisdale] has raised the question of doubt in his mind as to the wisdom of procedure on this item at this time [...]. In view of WET's doubt on the matter, I feel it would probably be better to postpone consideration of aid until WW has the opportunity of coming to England. While this is essentially an NS project, I am nevertheless convinced that the arguments presented in the recommendation for consideration still hold and that all the opinions I have from Mellanby, Landsborough Thomson, Fisher and Adrian who are directly concerned with the matter, are in support not only of the item as it stands but of the significance of having it done at the present time [...]." (Memorandum from D. P. O' Brien to A. Gregg, 6th March 1935, RA, Box 16, Folder 220).

463 Rockefeller Foundation Report 1935, p. 82.

464 Letter from R. A. Fisher to D. P. O'Brien, 18th July 1934, RA, Box 16, Folder 219.

465 Cost of the New Scheme in Serological Genetics, 1934, RA, Box 16, Folder 219.

466 Memorandum from D. P. O'Brien to A. Gregg, 1st March 1935, RA, Box 16, Folder 220. Taylor worked on serology for both his medical degree at Manchester University and his PhD in Cambridge and he spent ten years with H. Dean at the department of pathology at Cambridge University as a researcher and demonstrator acquiring an intensive training in serology (Letter from H. R. Dean to R. A. Fisher, 4th October 1934, RA, Box 6, Folder 17). A brief scientific biography and a bibliography of Taylor are in H. R. Dean (1946). Letter from G. L. Taylor to R. A. Fisher, 28th August 1937, R. A. Fisher Papers, BSL, The University of Adelaide. Correspondence between G. L.

unit started in 1935 with just one serologist, Taylor, one laboratory assistant Aileen M. Prior, and a boy attendant for the animal house. In 1936-1937 a new laboratory assistant, Elizabeth W. Ikin (later Dobson), an assistant serologist, Robert R. Race, and a further boy attendant completed the personnel of the unit.

Fisher reported to the Rockefeller Foundation a first positive outcome of the research programme in serology already in 1936 claiming that Taylor had "detected, and repeatedly confirmed, a very remarkable series of reactions on the blood of patients in institutions for the mentally deficient, reactions which it has not been possible, so far, to parallel using the blood of normal persons".[467] Eventually, instead, the serological study of the inheritance of mental defects proved unsuccessful. No linkage between serological reactions and heritable mental diseases could be confirmed. However, the pre-war investigations of the Galton Laboratory Serum Unit were essential to build the technical and statistical expertise in blood group research – before WWII this field was a rather marginal subject in Britain and competence in serology was sparse in the country (Appendix 3.4) – that qualified the workers of the Galton Laboratory Serum Unit as the only 'professional' serologists during the warfare.[468]

3.4 *The Galton Laboratory in WWII*

The expertise of the Galton Laboratory Serum Unit became of national importance in the months that preceded the outbreak of WWII and during the warfare. By 1939 the Spanish Civil War had proved the relevance of blood transfusion for treating casualties provoked by military operations and air raids and the usefulness of blood banks to cope with a national emergency.[469] The effects of bombing over the main towns, London in particular, were feared and, overestimating the effects of air raids, the British government anticipated millions of civilian casualties since the early stages of the conflict.[470]

In order to cope with this scenario an Emergency Medical Service was organised to meet the needs of the civilian population. The Emergency Transfusion Services were a component of such organisation.[471] They were set up in just a few months, with only limited resources and almost from scratch because no organised network of transfusion centres was already in operation.

Taylor and V. Friedenreich, 1936-1937, WL, SA/BGU/A1/2.

[467] Letter from R. A. Fisher to D. P. O'Brien, 18th May 1936, RA, Box 16, Folder 221. For the inconclusive results of the serological investigation of mental defects see J. Fisher Box (1978), p. 349.

[468] Letter from G. L. Taylor to R. A. Fisher, 5th April 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

[469] *Life Blood* (1945), p. 11. Booklet produced by the Ministry of Information.

[470] D. Edgerton (2011), pp. 35-37.

[471] An account of the Emergency Transfusion Services is J. M. Vaughan and P. Panton (1952).

Before WWII blood transfusion in Britain had been employed in medical practice only on a limited scale and blood donors were usually relatives of hospital patients or volunteers recruited by charitable organisations, such as the Red Cross.

The first attempts to organise panels of blood donors for the warfare began during the Munich crisis of September 1938. The Galton Laboratory Serum Unit did several blood-grouping tests at the request of University College Hospital and the head of the unit, George Taylor, offered since then the services of his staff in case of a national emergency.[472]

The organisation of the Emergency Transfusion Services began in the spring of 1939. The Medical Research Council, on behalf of the Ministry of Health, set up four depots – at Luton, Slough, Sutton and Maidstone – "for the collection, storage and supply of blood for transfusion purposes" and coordinated a network of empanelling centres in the London area and in the Home Counties.[473] In the South-West of England a special transfusion service to collect blood for the Army was established and from 1940 a Regional Transfusion Service covered the needs for blood transfusion in the rest of Great Britain and in Northern Ireland.

Within the Emergency Transfusion Services, George Taylor and his co-workers received the task of preparing testing serum for empanelling centres and hospitals. In August 1939 the unit was formally detached from the Galton Laboratory and taken over by the Medical Research Council.[474] Taylor moved with staff, equipment and his laboratory animals from London, considered an unsafe location, to Cambridge. There he found hospitality in the department of pathology, where he had worked before his appointment at University College.

The remaining staff of the Galton Laboratory, the genetic and computing component including Fisher, had a more troubled destiny. In September 1939 University College was evacuated and the College authorities encouraged the remaining staff of the Galton Laboratory to seek war work. Fisher fiercely opposed this idea. He found alternative accommodation at Rothamsted Experimental Station, his previous institution, and he moved there with the remaining staff, the calculating machines of the department and some of the animal stocks.[475]

Fisher envisioned for the Galton Laboratory in WWII the same role as a computing centre of national importance that the laboratory had had with Karl Pearson in WWI. This is evident in the

---

472 Letter from G. L. Taylor to E. Mellanby, 28th September 1938, TNA, FD/1/5845.
473 J. M. Vaughan and P. Panton (1952), p. 334.
474 Letter from A. Landsborough Thomson to G. L. Taylor, 25th July 1939, TNA, FD/1/5845.
475 In 1939, beyond the Serum Unit, the staff of the Galton Laboratory was formed by the geneticist Alexandre C. Fabergé and the genetical assistant Sarah B. North, the statisticians Wilfred L. Stevens and Horace W. Norton, Julia Bell in charge of the *Treasury of Human Inheritance*, the computer Mary N. Karn, Fisher's secretary Barbara E. Simpson and the computer of the British association Rose O. Tysser. For the controversy between Fisher and University College see J. Fisher Box (1978), pp. 373-375.

public appeal that Fisher made from the columns of the *Times* in September 1939 during his controversy with the authorities of University College.

> During the last war our administrators learned, though perhaps with some reluctance, that men trained in research were essential for the success of the national effort. The remaining nucleus of my department, if I may speak in its praise, constitutes a unit for heavy mathematical computations as efficient, both in machines and men, as the country can command.[476]

Fisher remarked this plan also in correspondence with the Rothamsted director, John Russell, writing him that his "whole idea in suggesting a move to Rothamsted was to maintain the computing unit intact for its possible future value during the war".[477]

Instead, no job of national importance was assigned to Fisher and his staff and the members of the Galton Laboratory at Rothamsted were dispersed in the following months. Fisher's failure to secure for his unit work of national importance stands out as an exception in WWII, where scientific expertise – in particular, statistical expertise in the context of operational research – was largely deployed.[478] His exclusion is even more surprising if we consider that Rothamsted Experimental Station hosted two teams engaged in wartime work, led respectively by Frank Yates and J. B. S. Haldane.

Frank Yates, head of the station statistics department, former co-worker of Fisher and his successor at Rothamsted, contributed to operational research as scientific advisor on bombing strategies in the staff of the scientist and public servant, Solly Zuckerman. Moreover, on behalf of the Ministry of Agriculture, Yates' department undertook several surveys to provide reliable information for planning an optimal use of the agricultural resources available under the wartime restrictions (**chapter 4**).

In 1940 also J. B. S. Haldane, professor of biometry at University College London, moved with his department to Rothamsted Experimental Station. Haldane, a colleague of Fisher and only two years younger, was actively engaged with his department in wartime work since the beginning of the conflict. He did physiological work for the Royal Navy and the Royal Air Force and with the collaboration of Helen Spurway, his colleague and later his second wife, was engaged in statistical investigations for the Royal Air Force, the Ministry of Aircraft Production

---

476 R. A. Fisher (1939b). According to J. Fisher Box six members of the Galton Laboratory, Fisher included, "volunteered for national service as a computing unit and were entered as such on the Royal Society's register of scientific personnel". (J. Fisher Box (1978), p. 373) It has not been possible, however, to countercheck this information with archival sources.

477 Letter from R. A. Fisher to E. J. Russell, 27th September 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

478 For the deployment of scientific expertise in WWII in Britain see D. Edgerton (2011). In particular, on operational research in Britain see D. Edgerton (2011), pp. 140-147 and pp. 290-293; M. Fortun and S. S. Schweber (1993). For more information on operational research see chapter 4.

and the Army.[479] Spurway, who was a geneticist and not a professional computer, did most of the calculations involved in the statistical work, as Haldane claimed to be "better as an algebraist than a computer", in sharp contrast with Fisher who was a skilled computer and had trained part of the Galton Laboratory staff as a computing unit.[480]

Joan Fisher Box suggests that Fisher's exclusion from war work – an exclusion he bitterly resented – was motived by the fact that "he appeared as a senior biologist".[481] But that would not explain why Haldane, very close to Fisher for age and scientific expertise, was instead engaged in war work.

David Edgerton in his examination of the role of scientific experts in WWII emphasises two aspects that help to understand Fisher's exclusion. On the one hand Edgerton remarks the political divisions among scientific experts in WWII, and on the other he points out that the influential figures of wartime advisors were such for a previous association with the military establishment, often in WWI.[482] On both fronts Fisher was doomed to fail. He had not taken part in WWI due to his poor eyesight, while Haldane had served as officer in the Scottish regiment of the Black Watch, and thus he could not claim any relevant connection with a military service. Furthermore, Fisher was a conservative, while it was a group of left wing scientists that did more to advocate the necessity for planning and operational research, both activities in which statistical knowledge was mainly deployed. Frank Yates and J.B.S. Haldane belonged instead to this group.

Fisher's failure to take part in the warfare was beneficial for the blood group survey. Excluded from war work and bereaved of his department, Fisher could devote a consistent share of his time to the survey, sorting and counting the donor records, corresponding with the transfusion officers across the country, calculating gene frequencies and comparing results gathered from different areas of Britain.

3.5 *The ABO blood group survey: information technologies and computing equipment*

3.5.1 *Data collection*

With the establishment of the Emergency Transfusion Services started a massive propaganda for the enrolment of blood donors. Appeals were issued in the national and local press and on the wireless, posters distributed all over the country and blood transfusion films screened in

[479] R. Clark (1984), p. 151.
[480] Letter from J. B. S. Haldane to J. D. Bernal, 10th January 1944, Haldane Papers, UCL Special Collections.
[481] J. Fisher Box (1978), p. 373.
[482] D. Edgerton (2011), p. 140 and p. 147.

cinemas. By 1945 a million and four hundred thousand people had enrolled as blood donors all over Britain.[483] For each volunteer suitable as donor the empanelling centres recorded and archived, besides the blood group, personal and contact details in order to recall him or her for a bleeding. The medical records collected by the Emergency Transfusion Services regarded unrelated donors distributed on a local basis as every empanelling centre enrolled people that could be easily contacted in case of necessity.

Ronald Fisher promptly realized that the mobilization for WWII and the creation of the wartime transfusion services had temporarily broken the constraints on the acquisition of blood group data in the British population. He could gain access to the records of the Emergency Transfusion Services through his former unit, engaged in the warfare, but now at stake there was not the inheritance of mental defects and Fisher discouraged the transfusion officers that proposed to undertake a family examination of the wartime records.[484] The new data available under the Emergency Transfusion Services were suitable for a survey of the geographic distribution of the ABO blood groups in the country, a survey that could offer information of ethnological and anthropological value on the British population.

By August 1939 Fisher and his co-workers guided by George Taylor had already collected about sixty thousand ABO blood group data with the collaboration of the depots and empanelling centres in London and the Home Counties and by mid-September Fisher had already contacted the main empanelling centres in Scotland – Edinburgh, Aberdeen, Glasgow, Dundee – to ask for their cooperation, as this could be an area of particular ethnological interest.[485] Therefore, Fisher and Taylor promptly followed the suggestion of the physician Edward Billing to issue a public appeal in *The British Medical Journal* inviting the officers of the Emergency Transfusion Services to offer their data for "ethnological deductions as to the racial origins – Celtic, Saxon or Danish – of the inhabitants of various parts of Great Britain from the percentages of the blood groups".[486]

In October 1939 Fisher and Taylor jointly called the empanelling centres "to co-operate by sending in from time to time numbers [of volunteers enrolled] classified in eight classes".[487] The information that Fisher sought from the transfusion records was at first only blood group and

---

483 Ministry of Information (1945), p. 7.

484 "I ought to say that it would be wise to regard any family investigation that you attempted as completely independent of the direct enumeration of volunteers which I am compiling." (Letter from R. A. Fisher to D. F. Cappell, 18th October 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide)

485 Letter from R. A. Fisher to J. R. Copland, 13th September 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

486 E. Billing (1939).

487 R. A. Fisher and G. L. Taylor (1939). It was Taylor who drew Fisher's attention towards Billing's suggestion in *The British Medical Journal* (Letter from G. L. Taylor to R. A. Fisher, 5th October 1939, R. A. Fisher Papers, BSL, The University of Adelaide).

sex of the donor – whence the eight classes giving male and female donors for each phenotype of the ABO system – but later also the age became an element for classification, in the attempt to test whether "differential death rate between the ages of twenty and sixty might alter the blood group proportions in this age range".[488] Fisher carried on the survey until 1943, when he left the Galton professorship for the chair of genetics in Cambridge.[489] In Cambridge his collaboration with the former Galton Serum Unit focused on the Rh blood group system whose unravelling represented Fisher's main contribution to serology during the warfare.[490]

By the spring of 1943 Fisher had gathered with the help of his co-workers over three hundred thousand data on the ABO blood groups (Appendix 3.2a), but was still trying to collect more evidence on many areas of Britain with the help of Taylor:

> As to places of which we lack knowledge, they are innumerable. Here is a list of some: The neutral South and West of Ireland/All inhabited islands near Great Britain, even Anglesey, Orkneys, Shetlands and Faroe/Any part of Scotland other than Edinburgh, Dundee, Aberdeen [...]/ The rural North of England, Westmoreland, Cumberland and Northumberland/ Industrial Tyneside/Industrial Merseyside. After that I am less inquisitive as the facts seems simple, but it would be good to be able to compare Grimsby with Hull; also East Anglia is very isolated geographically and may not be homogeneous with the Midlands and South. Any way we know scarcely anything about it.[491]

Despite Fisher's efforts, however, the survey could not be comprehensive, because it had to follow the enrolment campaigns of the Emergency Transfusion Services that were somehow haphazard in relation to their geographic location and decreased over time, considering that the needs for blood transfusion had been overestimated at the beginning of the war.[492]

Fisher's concern in the survey was not only related to the results that could be gathered from the analysis of the data. Rather, he wanted first of all to collect and preserve information susceptible of statistical treatment and of intrinsic interest for human genetics, convinced "that more good will be done by aiming at completeness of record than at the limitation of bad data".[493] In so doing Fisher certainly carried on the mission of the Galton Laboratory as a

---

488 Letter from R. A. Fisher to W. C. Boyd, 8th May 1953, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

489 The last letter in the Blood Group Survey correspondence is dated February 1943. From the letters exchanged by Fisher and Taylor in the same year it is clear that the survey went on for the following months, but little by little gave way to more pressing issues, such as Fisher's moving to Cambridge and the related arrangements, and the research on the Rh blood group system. At the conclusion of WWII Fisher tried once more to complete and revise his data collection – see the correspondence (January-February 1946) between Fisher and A. E. Mourant on the figures of the donors enrolled by the Luton Depot, WL, SA/BGU/E7/1 –, but no new publication came out of this.

490 For Fisher's contribution to the Rhesus blood group system see J. Fisher Box (1978), pp. 356-370; A. W. F. Edwards (2007).

491 Letter from R. A. Fisher to G. L. Taylor, 27th March 1943, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

492 "I was interested to see a recent appeal for some enormous number of blood donors from Yorkshire. [...] I cannot imagine why there should be a drive for volunteers at present, as I presume not more than one in 10 who have already registered have yet been asked for blood". (Letter from R. A. Fisher to G. L. Taylor, R. A. Fisher Papers (digitized), 3rd February 1941, BSL, The University of Adelaide)

493 For the data collection aim of the survey see Letter from R. A. Fisher to J. R. Copland, 13th September 1939, R.

"storehouse for statistical material" related to human inheritance, but it was also a private ambition of Fisher to employ his departmental resources to save data amenable to statistical treatment, regardless of their nature. [494]

While at Rothamsted Experimental Station Fisher had transformed his statistics department in the repository of the yields and weather records collected at the station since the nineteenth century (**chapter 1**). Even beyond the data that related to his work in agricultural statistics, Fisher had been willing to deploy his departmental resources in order to save collections of records that he believed of interest for genetics, as the set of measurements on the eggs of the British birds that Fisher used to compare abundance and variability of the different species.[495] The data that prompted the study were collected by an amateur ornithologist, Rev. F. C. R. Jourdain, copied in the Rothamsted statistics department, and eventually deposited by Fisher in the archives of the Museum of Natural History in London. In the correspondence between Fisher and Jourdain there is already the routine that will become a feature of the blood group survey: receiving and sending back records, press for the dispatch of new ones, making enquiries about data that seemed incongruous.[496]

The formula of the blood group survey, thus, had been already tested in Fisher's career. In wartime, however, the network of people, institutions and records involved was much more complex. Appendix 3.2b gives an overview of the Emergency Transfusion Services that took part in Fisher's survey, as it can be reconstructed from archival and published sources.

3.5.2 *Information technologies for the management of blood donor records*

Up to WWII the Red Cross managed the most efficient system of blood donors and represented the main source of volunteers for the British hospitals. Its involvement in blood donation had started in the early 1920s when the honorary secretary of the Camberwell Division of the Red Cross, a civil servant, Percy Lane Oliver, set up a panel of voluntary blood donors.[497] Until 1925 the service, called the London Blood Transfusion Service, had its headquarters in Oliver's home and was supported only by voluntary contributions. In 1926 the Red Cross took over the service that became the British Red Cross Blood Transfusion Service, while Oliver conserved his leading role as organiser.[498]

A. Fisher Papers (digitized), BSL, The University of Adelaide. The quotation is in the Letter from R. A. Fisher to J. Vaughan, 5th December 1939, R. A. Fisher Papers, BSL, The University of Adelaide.

494 Appeal Fund, 1907-1920, UCL Record Office, AR 179, Folder 4.

495 R. A. Fisher (1937).

496 Birds data correspondence, RR Library and Archive, STATS 6.4.

497 A biographical sketch of P. L. Oliver is G. W. G. Bird (1992).

498 H. Hanley (1998) and H. Dodsworth (1996) offer an account of blood transfusion in Britain, including the

Despite its denomination, the Red Cross Blood Transfusion Service undertook only the clerical work of storing the data of the volunteers suitable as donors, receive the calls of requests from hospitals, retrieve the name of a possible donor compatible with the request and contact him or her. Below an account of the clerical activity of the Red Cross office prepared by Oliver in the late 1930s.

> Practically all calls come by telephone, perhaps 1 in 500 arriving by letter. We require the name of the hospital, name and group of patient and name of surgeon who will carry out the transfusion. These details are immediately entered on [a] form [...]. Donors' records are kept on cards, a different coloured card being used for each blood group. The whole of the information that can be obtained regarding the donor is typed upon the front, including times available, special objections or preferences for particular hospital, and such memoranda as 'Do not mention blood transfusion' or 'All correspondence in sealed envelopes', etc. The back of the cards contains a record of dates and hospitals at which the donor has given transfusions. Metal indicators clipped in various positions on the card denote – (a) Novice not yet used, (b) Keen for an early call, (c) Particularly available for night calls, (d) On telephone at home and (e) Possesses car or motorcycle. [...] In addition to the card index and the main register, an additional register of donors is kept, arranged in the various postal districts of London and the suburbs. This is particularly useful after business hours as it gives at a glance the number of members available in the vicinity of a particular hospital.[499]

Since 1929 the information technology at the core of the Red Cross Office was the filing drawer of coloured index cards with names, contact details, blood group of the donor according to the Moss classification (and embedded in the colour of the card) and time and places of previous bleedings.[500]

The cards adopted by the service were humble office technologies: made in plain cardboard, without any pre-set ruling, they were typewritten according to necessity by the staff of the office (Fig. 3.1a, 3.1b). They were cheap and easy to use as the colour code (blue for group O, yellow for group A, orange for group B and white for group AB) made a clerical mistake on the blood group of the volunteer almost impossible. The Red Cross conserved this system for decades. It was certainly still used after WWII, but at that stage the cards were divided in two filing drawers, one for donors with positive Rh factor, and the other with negative Rh.[501]

---

involvement of the Red Cross. First hand information on the Red Cross Blood Transfusion Service and its statistics is provided by the *Blood Transfusion Service Quarterly Circular*, the periodical newsletter started in October 1933 as an attempt "in keeping the individual members together as an organisation" (*Blood Transfusion Service Quarterly*, No. 1).

499 "Report on the work of the Blood Transfusion Service" (by P. L. Oliver), about 1937, WL, P. L. Oliver Papers, uncatalogued.

500 Before WWII the classification used for the ABO blood groups in Britain was the numerical classification suggested by W. L. Moss (I = group AB, II = group A, III = group B, IV = group O). For the several classification systems adopted for the ABO blood groups before WWII see W. H. Scheneider (1983), p. 548. For the index cards in the Red Cross Office see Minutes of the Meeting of the Committee of the London Blood Transfusion Service, 9th April 1930, WL, P. L. Oliver Papers, uncatalogued. A useful source on index cards and other office equipment popular in the 1930s and mentioned hereinafter is J. Robson (1929).

501 "Red Cross Blood Transfusion Service Telephone Duties", anonymous and undated report, WL, P. L. Oliver Papers, uncatalogued. From correspondence with the archivist of St. Bartholomew, London, an hospital mentioned in the report, it is likely that the document should be dated to the 1960s (e-mail communication with Kathie Ormerod, Barts Archive). The same report mentions the different set of colours chosen for the donors. The only discrepancy in the choices of the British Red Cross Blood Transfusion Service and of the Emergency Transfusion

The donor index cards were in no way the only information technologies used by the Red Cross Transfusion Service. People were recruited filling in an enrolment form in which the candidate stated name, contact details and the times at which he or she was available for donation (Fig. 3.2). Furthermore, the Red Cross monitored the outcome of the transfusion – with the reports sent in by the hospitals after each bleeding – and the condition of the volunteer – who was expected to report to Oliver's office within a day from the bleeding. In the economy of the Red Cross office it was crucial to know constantly and exactly all the passages underwent by the volunteers and by the blood transfused and the widespread use of information technologies, such as forms and index cards, was instrumental to the achievement of this result.

Two factors can account for such centrality of control and communication, or better "control through communication".[502] Blood transfusion in the 1920s and largely in the 1930s was still considered a risky practice. Many volunteers were forced to resign under the pressure of their family and relatives concerned for their well-being. Donors might claim negative effects due to the bleeding or hospitals might attribute health problems of the donor after the bleeding to the bad health of the volunteer. The Red Cross service had to face all these contentious issues. The transfusion could also be risky for the recipient because death resulted when a donor was wrongly grouped and his or her blood given to an incompatible recipient. In order to limit controversies a constant monitoring of the donors and their bleedings was required.[503]

The second reason is the voluntary nature of the system set up by the Red Cross. In order to bind the volunteers to the service, a constant relation with the donors was required and reply cards, hospital reports and certificates of service were to establish it. After each bleeding, in fact, the volunteer should re-enrol for the service and a stamped addressed reply card was sent to him or her for this purpose (Fig. 3.3). Along with the card the volunteer received an hospital report recording date and place of transfusion, name of the donor and the surgeon, sex of the patient and its medical conditions, amount of blood drawn and outcome of the transfusion, and a certificate of thanks hand decorated by the clerical staff of the Red Cross Blood Transfusion Service. In the 1930s donors who had given ten or more transfusions received also a bronze badge in recognition of their service. Reports, certificates, badges represented a moral recognition for the unpaid service given by the donor.[504]

Services is in the colour of the B donor cards – orange in the former case, pink in the latter. It might be that before the war also the Red Cross adopted the pink colour for B donors.

502 As it will be evident below, I am drawing on the analysis of the role of information technologies in management presented by J. Yates in *Control through Communication* (1993).

503 "Report on the Work of the Blood Transfusion Service" (by P. L. Oliver), about 1937, WL, P. L. Oliver Papers, uncatalogued.

504 A sample of a 1938 hospital report is in H. Dodsworth (1996), p. 460. On the certificates of thanks see the Minutes of the Meeting of the Committee of the London Blood Transfusion Service, 9th April 1930, WL, P. L.

The net result of these two requirements was that an idea of "scientific management" of the donors and by extension of the blood they had donated informed the office organisation of the Red Cross. This managerial attitude was linked to the massive use of information technologies – the telephone, the typewriter, the standardised form, the filing system – in the Red Cross Office, instead of more informal arrangements that might have been expected considering the voluntary and charitable nature of the enterprise. As argued by Joanne Yates for the development of American corporations, communication and information technologies "contributed to the specialization of office skills and consequently created an opportunity for applications of scientific and systematic management to the office as well as to the factory floor".[505]

The Red Cross Transfusion Service, of course, was not a business organisation, but like the companies that Yates considers, had experimented a quick growth in its activity. If in 1921 the service had provided only a donor for blood transfusion, ten years later the donors had increased to two thousand and by 1938 the donors provided for blood transfusion had been more than six thousand.[506] Such increase was only partially matched by the increase of the office staff, therefore a managerial attitude and the adoption of efficient information technologies were required to cope with the expansion of the service.

The Emergency Transfusion Services inherited the managerial attitude that had characterised the Red Cross Blood Transfusion Service and the centrality of information technologies in the organisation of blood transfusion, as the choice of paper tools inexpensive, readily accessible, available in large quantities and easy to use was decisive for the wartime services that had limited resources and should rely not only on qualified personnel, but also on volunteers and people trained on the spot.

The Emergency Transfusion Services borrowed the Red Cross system of colour cards for the internal administration of donors, but also the use of forms to transfer information into the service and the issue of certificates of acknowledgements. Moreover, Oliver, the peacetime organiser of the Red Cross service, was actively involved in the set up of the wartime transfusion scheme and the British Red Cross advanced even two hundred and fifty pounds "for printing, postage, clerical assistance and other administrative expenses" that were required in the organisation of the Emergency Transfusion Services.[507]

Oliver Papers, uncatalogued. For the donor badges see the Minutes of the Meeting of the British Red Cross Society Blood Transfusion Committee, 27th June 1932, WL, P. L. Oliver Papers, uncatalogued.

505 J. Yates (1993), p. 64.

506 The figures of the volunteers provided by the British Red Cross Blood Transfusion Service are in F. Hanley (1998), p. 69.

507 Minutes of the Meeting of the British Red Cross Society Blood Transfusion Committee, 13th April 1939, WL, P. L. Oliver Papers, uncatalogued.

At the core of the collaboration between Ronald Fisher and the empanelling centres there were the information technologies related to the management of the donor records. In the Emergency Transfusion Services the flow of information related to the donor was so structured: the prospective donor filled in an enrolment form for volunteering, the data of the donor were acquired and at the same time or in a second moment a few drops of the donor's blood were taken for grouping. After the laboratory test a card was sent to the donor stating his or her blood group and with this card in hand the donor should come back when summoned for a bleeding.

Once the potential donor had been grouped and was considered suitable for the task his or her data were transferred to another information system. In a few instances this was a register or an alphabetic list, but in most cases the choice was for a system of movable index cards.[508] All the blood depots in the London area recorded the donors in their panel with such a system. P. L. Oliver suggested the format of the cards adopted by these depots (see the quotation below). The information to be recorded was name, age, private and business contact details, health conditions, eventual war work of the volunteer, blood group and condition of the veins. In the index cards were also recorded the various bleedings of the volunteer in order to avoid an excessive blood demand on a single person.

> Mr Oliver had sent some specimen cards for indexing etc. The following was finally decided, should be sent to Mr Oliver for his approval, with the suggestion that it should be printed on a card 8×5 ins.
>
> Name in full Age
>
> Private address
>
> Business Address
>
> Telephone Number Private Business
>
> Have you entered any national service obligations that may interfere with your availability for blood transfusion purposes?
>
> Are you in good health?
>
> Have you had any serious illness recently?
>
> Group Character of Arm vein.[509]

[508] Bits and pieces of such lists survive in the correspondence of the Galton Laboratory Serum Unit at the department of pathology in Cambridge, because they were reused for keeping copy of the outgoing mail. One such bit is the portion of a list from the Regional Blood Transfusion Service in Sheffield dated December 1942. The list is organised in six columns: ABO blood group of the donor, serial number, name, sex, age, Rh grouping. In the mid of the warfare, in fact, the Emergency Transfusion Services began to test the donors also for the Rh factor (TNA, FD/8/6).

[509] Minutes of the Meeting of the Sub-Committee on Blood Transfusion of the MRC Emergency Pathological Services, 17th April 1939, WL, GC/186/1.

The documentary film *Blood Transfusion*, produced in 1941 by the British Ministry of Information on behalf of the Ministry of Health, offers an insight into the record administration of the Emergency Transfusion Services. The documentary, shot for propaganda, included a short sequence with a live show of the clerical activity in the Slough Depot. In the film sequence appeared enrolment forms (Fig. 3.4) and index cards (Fig. 3.5), with an explanation of the enrolment procedure in voice over.

> Men and women come to the Depot in their own time on their way to and from work. Clerks, typists, housewives, factory workers, shopkeepers, city workers. To become a donor you must be between the ages of 18 and 65. When you arrive at the depot to register, we take your name and address, details of your health, and a drop of blood from your ear or finger. This drop of blood is taken to the laboratory, where we find out which group you belong to. After registry, you can come to be bled at the depot. You are given your index card, which records your blood group, health details, where you can be found in an emergency, and how often your blood has been given.[510]

The enrolment forms and the index cards in the documentary are very similar, for the information provided, to their counterparts in the Red Cross service that evidently came to represent a standard for the clerical organisation of blood transfusion in Britain. As pointed out by Geoffrey Bowker and Susan Leigh Star "a standard spans more than one community of practice (or site of activity). It has temporal reach as well as in that persists over time".[511] In this case the clerical organisation established by the Red Cross suited the charitable service as well as the emergency one and lasted well after the warfare. Both the panel of voluntary donors carried on by the Red Cross during and after WWII, and the National Transfusion Service that developed from the Emergency Medical Services after the war, again relied on the paper support of forms and cards with consistent analogies to the pre-war and wartime organisation discussed so far.[512]

The Emergency Transfusion Services directly managed by the Medical Research Council in the London area and in the Home Counties, inherited the same colour code used by the Red Cross to distinguish the blood groups. The colour code of the depot card cannot be gathered from the black-and-white film, but it is confirmed in the correspondence between Ronald Fisher and the director of the Slough Depot, Janet Vaughan.[513] The colour system is also evident in the

---

[510] *Blood Transfusion Service*, UK Ministry of Information, 1941. Some parts of the documentary have been made available online http://catalogue.wellcome.ac.uk/record=b1675865~S3 by the WL as part of their film section. The scene shot in the Slough Depot is in the third segment, from min. 1:00 to min 1:45.

[511] G. C. Bowker and S. Leigh Star (2000), p. 13.

[512] For the Red Cross long term use of the same information technologies see "the Red Cross Blood Transfusion Service Telephone Duties", anonymous and undated report, WL, P. L. Oliver Papers, uncatalogued. Some colour-coded labels and some cards adopted by the National Transfusion Service are conserved at the WL in London, Ref. GC/107/2.

[513] "If the forms do not materialise I think we can still probably save the situation by making a fresh list of all our ABs. This can be done by picking out the white cards from the other colours, and if necessary we will arrange to have this done." (Letter from J. Vaughan to R. A. Fisher, 10th November 1939, R. A. Fisher Papers, BSL, The University of Adelaide)

index cards still surviving among the correspondence of the Galton Laboratory Serum Unit in Cambridge (Fig. 3.6a, Fig. 3.6b).[514] The Galton unit, as mentioned before, was not an empanelling centre for blood transfusion, but it set up a donor panel for preparing the testing serum. The cards used by the unit had the same basic structure of the ones used by the British Red Cross. Unlike the depot card presented in the film there was no preliminary printed space, but they were simply typewritten according to necessity. The data were the same already mentioned: name, sex and age of the volunteer, the donor number, contact details and conditions of the veins.

In the Emergency Transfusion Services run by the Medical Research Council the colour code of the donor index cards was linked to another set of records, the cards mailed to volunteers after they had been grouped. Again the colours white, yellow, pink – but almost indistinguishable from the orange adopted by the Red Cross – and blue stood for the blood groups AB, A, B and O.[515] Hugh F. Brewer, medical officer for the Luton depot, criticised the original format of these cards, as "the wording on the original Medical Research Council cards gave the impression that Groups A, B and AB were to some extent of negative value compared with O and donors frequently assumed, with disappointment, that they were of little value belonging to these Groups."[516] For this reason in 1940 he prepared a new version (Fig. 3.7a, Fig. 3.7b, Fig. 3.7c, Fig. 3.7d) valid for all the blood groups and stating that "[t]here is no difference in value or quality between the four different blood groups".[517]

The Emergency Transfusion Services could not guarantee the constant engagement with the donor that had been a feature of the British Red Cross service, but still they had to nurture trust and collaboration due to the voluntary nature of the service and the cards they gave to donors should contribute to this mission, as hospital reports and reply cards had contributed to the success of the Red Cross Blood Transfusion Service before the war. In 1943 even the Emergency Transfusion Services began to distribute cardboard certificates with insets for thanking the volunteers.[518]

In the area of its competence, London and the Home Counties, the Medical Research Council pursued a standardisation of the information technologies related to blood transfusion, in

[514] These index cards survived by chance. As in the case of the lists mentioned above these cards were discarded because the donor withdrawn from the service and the cardboard was re-used for storing copies of the outgoing mail.

[515] In the official version adopted by the MRC the cards for the blood groups A, B and AB were to be sent as postcards, while the card for the O donors, the ones more commonly used, was to be sent with an envelope. (Letter from the MRC to the Medical Officer of Health [unknown area], 24th August 1939, TNA, FD/1/5751)

[516] Letter from H. F. Brewer to A. N. Drury, 6th November 1940, TNA, FD/1/5923. Quotation reproduced with permission from the Medical Research Council.

[517] Luton Depot donor cards 1940 (TNA, FD/1/5923). Quotation reproduced with permission from the Medical Research Council.

[518] Letter from A. N. Drury to P. N. Panton, 23rd September 1943, WL, GC/107/1.

particular for the enrolment forms and the cards sent to donors stating their blood group.[519] Standardisation was not fully achieved because adjustments *ad hoc* were introduced, as seen for the donor cards issued by the Luton Depot, but there was no significant distortion in the management of the donor records. Even the clerical organisation of the regional transfusion services closely followed the choices of the Medical Research Council for the enrolment forms and the cards sent them stating their blood groups.

In the system of internal administration of the donors, however, not all the services followed the colour code borrowed from the Red Cross. Such is the case of the Regional Transfusion Service of Manchester and Salford. The service adopted a system of index cards that were sorted not by colour, but through the punches on their edges.[520] The Manchester cards were technically more sophisticated, but not qualitatively different from the others employed by the wartime transfusion organisations. They provided the same type of information ensuring, if not the perseverance of the material standard, the one of the information model it embodied (Fig. 3.8a, Fig. 3.8b).

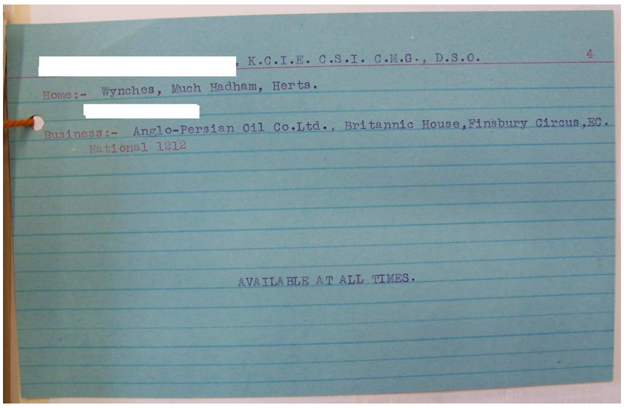
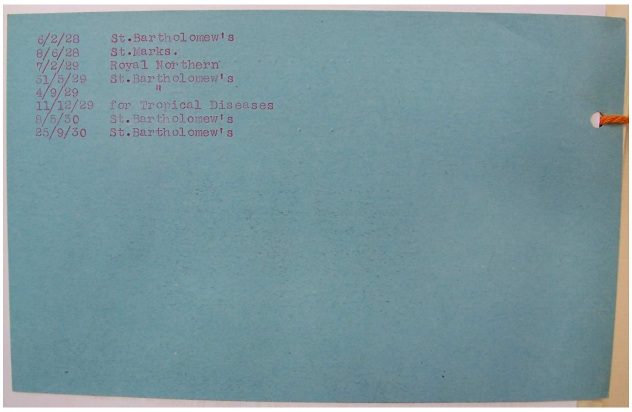


Fig. 3.1a British Red Cross Blood Transfusion Service, index card (recto) for a blood group O donor (late 1920s). The number 4 on the top right is the number correspondent to the O group in the Moss classification.
Credits: Reproduced with permission from the Medical Research Council. Document deposited at The National Archives, Ref. FD/1/3245.

---

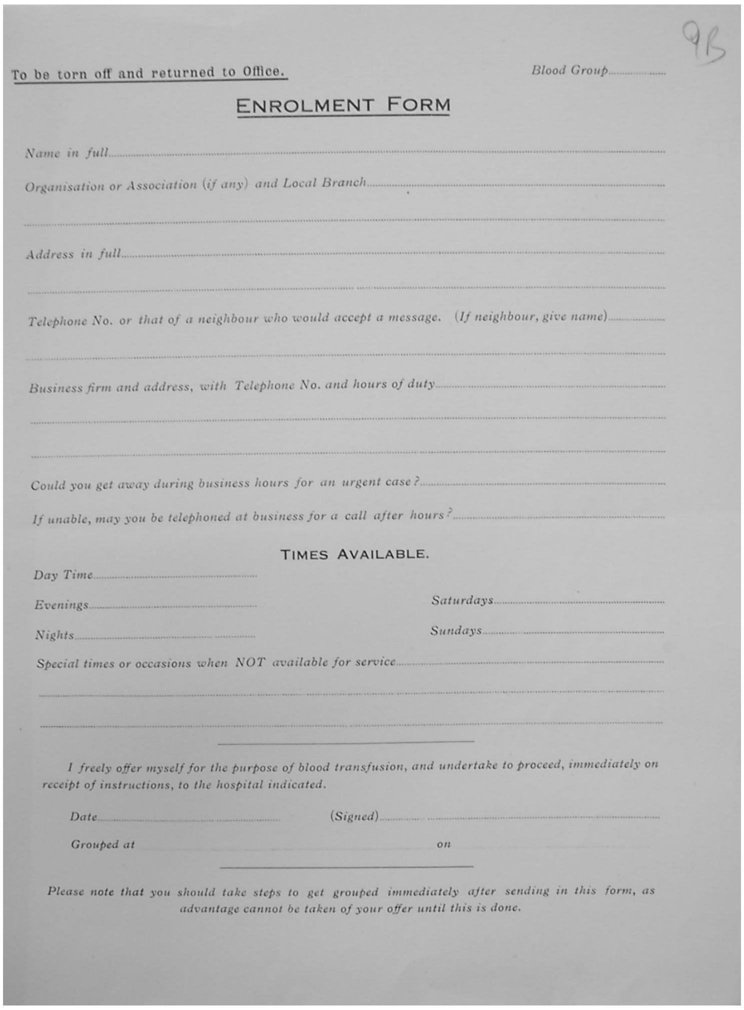
519 "We have ordered a series of blue, white, yellow and pink cards for use at the Sector empanelling stations, and [...] we have had an additional quantity of each card prepared for the Depots. [Follow explanation of the colour code] These cards are to be ready in the desired quantities on the 3rd or 4th July, and the printers, George Pulman & Sons, Ltd., have been instructed to deliver them direct to the Depot and Sector empanelling stations." (Letter from the MRC to H. F. Brewer, 24th June 1939, TNA, FD/1/5854. Quotation reproduced with permission from the Medical Research Council)

520 For the description of the sorting of these cards see J. Robson (1929), pp. 282-283.



6/2/28 St.Bartholomew's
8/6/28 St.Marks.
7/2/29 Royal Northern
31/5/29 St.Bartholomew's
4/9/29 "
11/12/29 for Tropical Diseases
8/5/30 St.Bartholomew's
25/9/30 St.Bartholomew's

Fig. 3.1b British Red Cross Blood Transfusion Service, index card (verso) for a blood group O donor (late 1920s).
Credits: Reproduced with permission from the Medical Research Council. Document deposited at The National Archives, Ref. FD/1/3245.



9B

To be torn off and returned to Office. *Blood Group*

ENROLMENT FORM

Name in full

Organisation or Association (if any) and Local Branch

Address in full

Telephone No. or that of a neighbour who would accept a message. (If neighbour, give name)

Business firm and address, with Telephone No. and hours of duty

Could you get away during business hours for an urgent case?

If unable, may you be telephoned at business for a call after hours?

TIMES AVAILABLE.

Day Time

Evenings *Saturdays*

Nights *Sundays*

Special times or occasions when NOT available for service

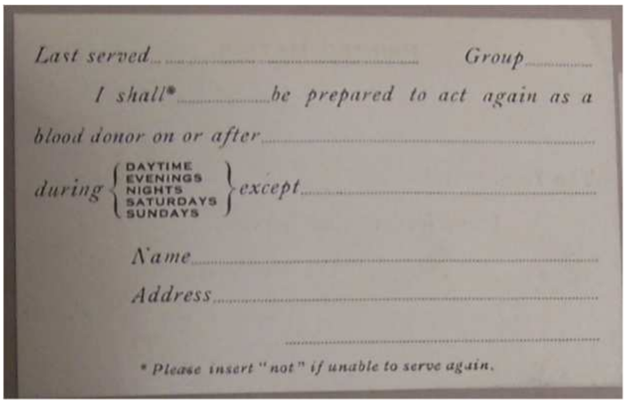
I freely offer myself for the purpose of blood transfusion, and undertake to proceed, immediately on receipt of instructions, to the hospital indicated.

Date *(Signed)*

Grouped at *on*

Please note that you should take steps to get grouped immediately after sending in this form, as advantage cannot be taken of your offer until this is done.

Fig. 3.2 British Red Cross Blood Transfusion Service, enrolment form.
Credits: Reproduced with permission from the Medical Research Council. Document deposited at The National Archives, Ref. FD/1/3245.



Last served.......... Group..........

I shall*..........be prepared to act again as a blood donor on or after..........

during { DAYTIME / EVENINGS / NIGHTS / SATURDAYS / SUNDAYS } except..........

Name..........

Address..........

....

* Please insert "not" if unable to serve again.

Fig. 3.3 British Red Cross Blood Transfusion Service, donor reply card.
Credits: British Red Cross Museum and Archives, Ref. Acc/562/69.

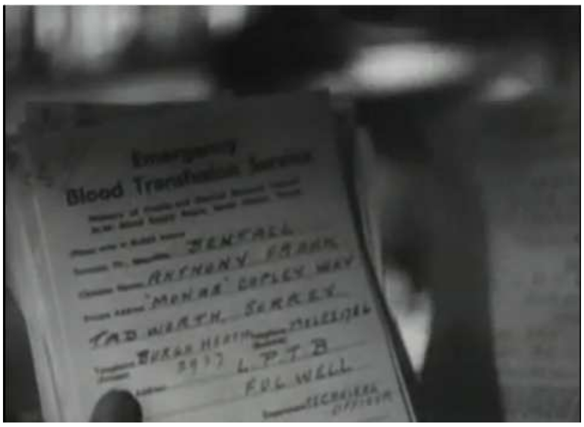


Fig. 3.4 Emergency Blood Transfusion Services, enrolment forms, Slough Depot. Still image obtained from the film "Blood Transfusion Service", UK Ministry of Information, 1941.
Credits: Wellcome Library, 2009. System No. .b16758651. The film is made with material conserved at the BFI archive.

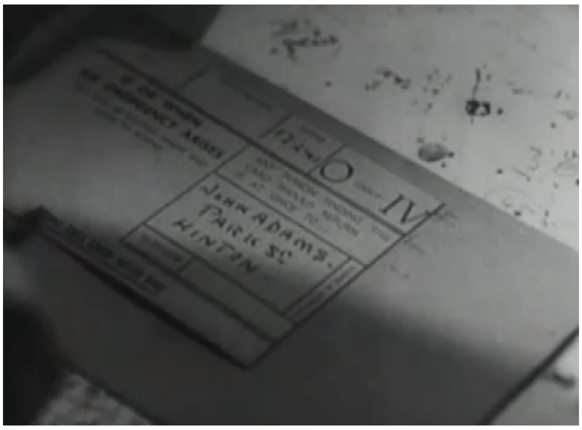
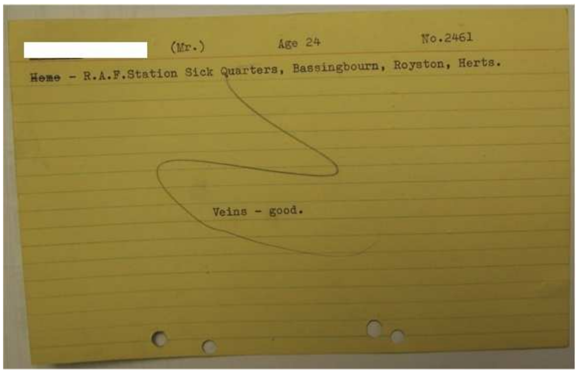


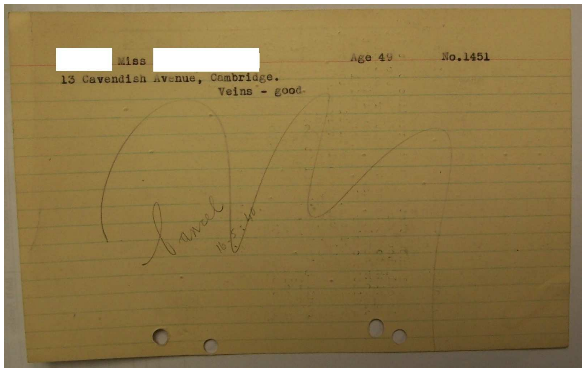
Fig. 3.5 Emergency Blood Transfusion Services, donor index card, Slough Depot. Still image obtained from the film "Blood Transfusion Service", UK Ministry of Information, 1941.
Credits: Wellcome Library, 2009. System No. .b16758651. The film is made with material conserved at the BFI archive.



[redacted] (Mr.) Age 24 No.2461

Home - R.A.F.Station Sick Quarters, Bassingbourn, Royston, Herts.

Veins - good.



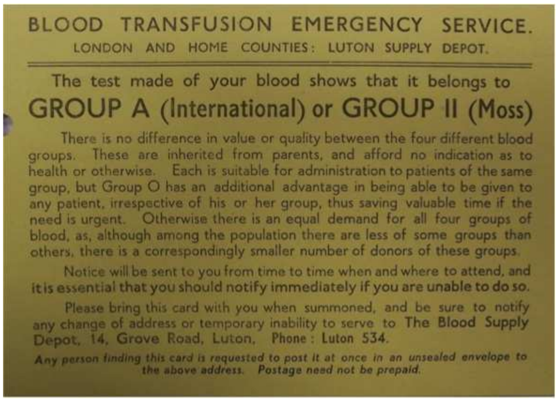
[redacted] Miss [redacted] Age 49 No.1451

13 Cavendish Avenue, Cambridge.

Veins - good.

Fig. 3.6a, 3.6b Galton Laboratory Serum Unit at the department of pathology, Cambridge University, donor index card, blood group A (yellow) and blood group AB (white). The Galton Laboratory Serum Unit in Cambridge relied on both soldiers of the Royal Air Force and local Cambridge inhabitants, mainly women, for the preparation of testing serum.

Credits: Reproduced with permission from the Medical Research Council. Document deposited at The National Archives, Ref. FD/8/3.



BLOOD TRANSFUSION EMERGENCY SERVICE.

LONDON AND HOME COUNTIES: LUTON SUPPLY DEPOT.

The test made of your blood shows that it belongs to

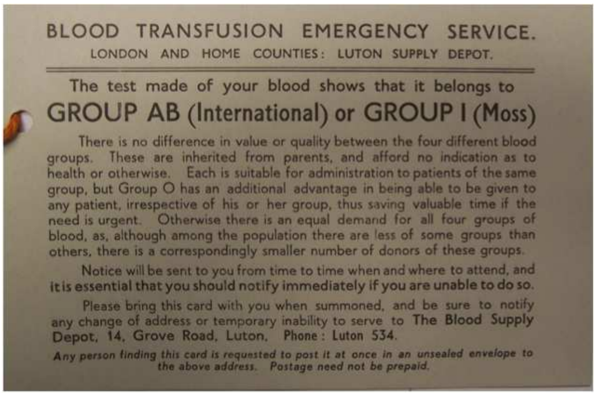
GROUP A (International) or GROUP II (Moss)

There is no difference in value or quality between the four different blood groups. These are inherited from parents, and afford no indication as to health or otherwise. Each is suitable for administration to patients of the same group, but Group O has an additional advantage in being able to be given to any patient, irrespective of his or her group, thus saving valuable time if the need is urgent. Otherwise there is an equal demand for all four groups of blood, as, although among the population there are less of some groups than others, there is a correspondingly smaller number of donors of these groups.

Notice will be sent to you from time to time when and where to attend, and **it is essential that you should notify immediately if you are unable to do so.**

Please bring this card with you when summoned, and be sure to notify any change of address or temporary inability to serve to **The Blood Supply Depot, 14, Grove Road, Luton. Phone: Luton 534.**

Any person finding this card is requested to post it at once in an unsealed envelope to the above address. Postage need not be prepaid.



BLOOD TRANSFUSION EMERGENCY SERVICE.

LONDON AND HOME COUNTIES: LUTON SUPPLY DEPOT.

The test made of your blood shows that it belongs to

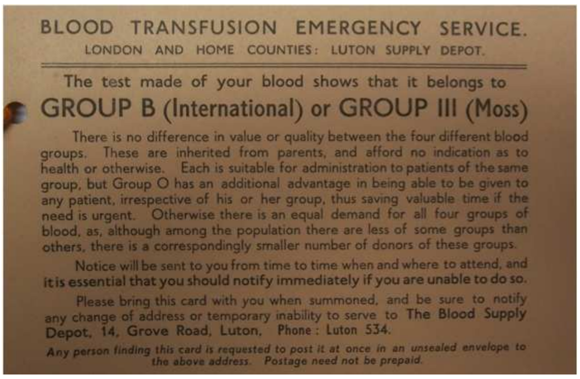
GROUP AB (International) or GROUP I (Moss)

There is no difference in value or quality between the four different blood groups. These are inherited from parents, and afford no indication as to health or otherwise. Each is suitable for administration to patients of the same group, but Group O has an additional advantage in being able to be given to any patient, irrespective of his or her group, thus saving valuable time if the need is urgent. Otherwise there is an equal demand for all four groups of blood, as, although among the population there are less of some groups than others, there is a correspondingly smaller number of donors of these groups.

Notice will be sent to you from time to time when and where to attend, and **it is essential that you should notify immediately if you are unable to do so.**

Please bring this card with you when summoned, and be sure to notify any change of address or temporary inability to serve to **The Blood Supply Depot, 14, Grove Road, Luton. Phone: Luton 534.**

Any person finding this card is requested to post it at once in an unsealed envelope to the above address. Postage need not be prepaid.



BLOOD TRANSFUSION EMERGENCY SERVICE.

LONDON AND HOME COUNTIES: LUTON SUPPLY DEPOT.

The test made of your blood shows that it belongs to

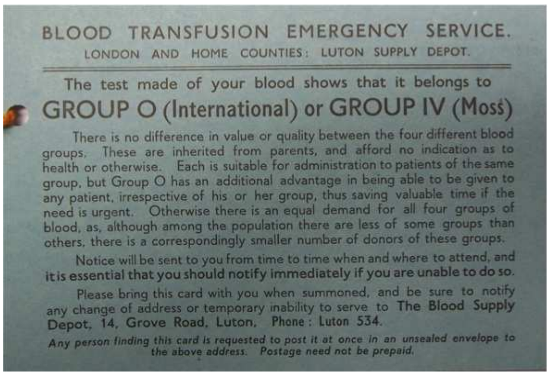
GROUP B (International) or GROUP III (Moss)

There is no difference in value or quality between the four different blood groups. These are inherited from parents, and afford no indication as to health or otherwise. Each is suitable for administration to patients of the same group, but Group O has an additional advantage in being able to be given to any patient, irrespective of his or her group, thus saving valuable time if the need is urgent. Otherwise there is an equal demand for all four groups of blood, as, although among the population there are less of some groups than others, there is a correspondingly smaller number of donors of these groups.

Notice will be sent to you from time to time when and where to attend, and **it is essential that you should notify immediately if you are unable to do so.**

Please bring this card with you when summoned, and be sure to notify any change of address or temporary inability to serve to **The Blood Supply Depot, 14, Grove Road, Luton. Phone: Luton 534.**

Any person finding this card is requested to post it at once in an unsealed envelope to the above address. Postage need not be prepaid.



BLOOD TRANSFUSION EMERGENCY SERVICE.

LONDON AND HOME COUNTIES: LUTON SUPPLY DEPOT.

The test made of your blood shows that it belongs to

GROUP O (International) or GROUP IV (Moss)

There is no difference in value or quality between the four different blood groups. These are inherited from parents, and afford no indication as to health or otherwise. Each is suitable for administration to patients of the same group, but Group O has an additional advantage in being able to be given to any patient, irrespective of his or her group, thus saving valuable time if the need is urgent. Otherwise there is an equal demand for all four groups of blood, as, although among the population there are less of some groups than others, there is a correspondingly smaller number of donors of these groups.

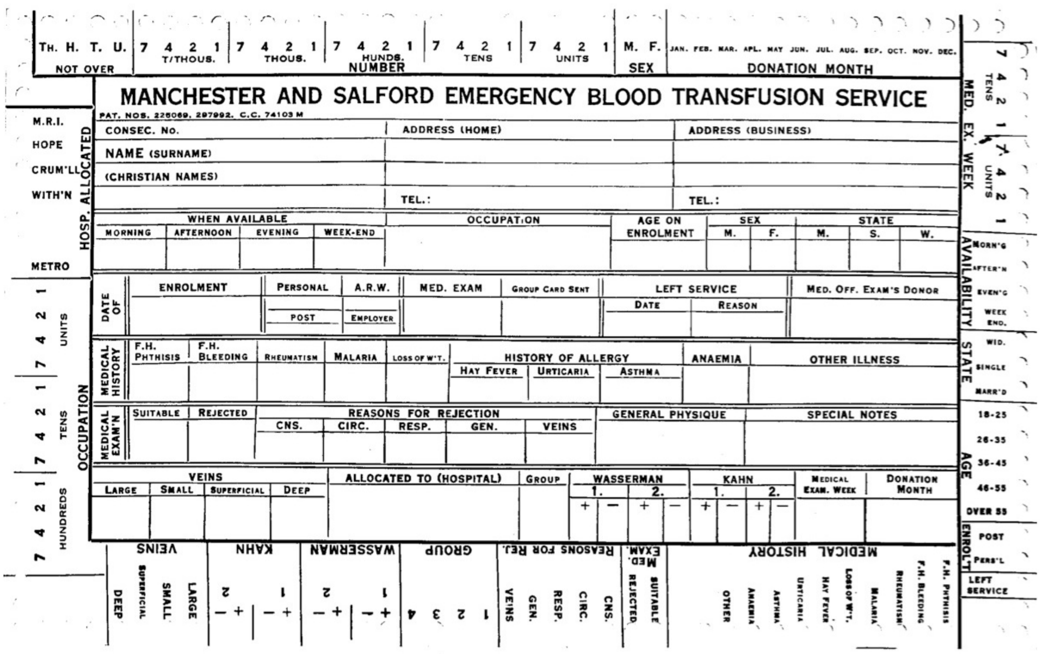
Notice will be sent to you from time to time when and where to attend, and **it is essential that you should notify immediately if you are unable to do so.**

Please bring this card with you when summoned, and be sure to notify any change of address or temporary inability to serve to **The Blood Supply Depot, 14, Grove Road, Luton. Phone: Luton 534.**

Any person finding this card is requested to post it at once in an unsealed envelope to the above address. Postage need not be prepaid.

Fig. 3.7a, 3.7b, 3.7c, 3.7d Emergency Blood Transfusion Services, cards sent to donors stating their blood group, November 1940, Luton Depot.

Credits: Reproduced with permission from the Medical Research Council. Document deposited at The National Archives, Ref. FD/1/5923.



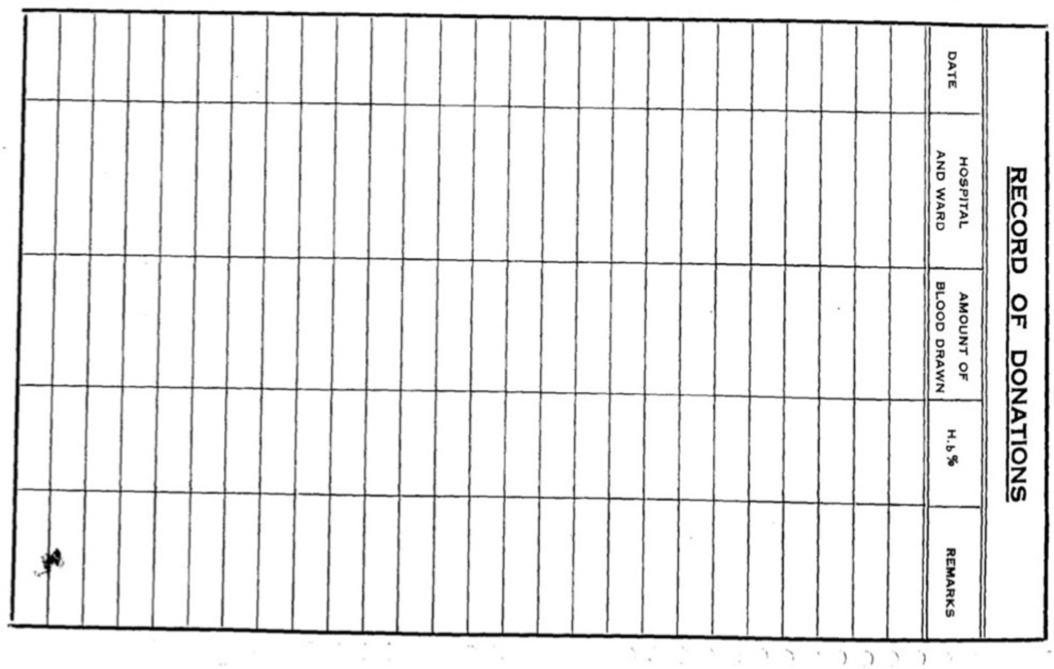


Fig. 3.8a, 3.8b Emergency Blood Transfusion Services, Manchester and Salford Blood Transfusion Service, donor index card, November 1941. The quality of the image is poor, but all around the edges of the card there are circular holes. In order to fill in the information on the edge of the card it is sufficient to cut out in the right way the corresponding hole. For example if the donor is male a u-shaped portion of the cardboard will be cut out eliminating the corresponding hole. The sorting of these cards is very simple. A rod is passed through the circle with the relevant information. In the batch, all the cards that satisfy the condition will not be captured by the rod and when the rod is lifted they will simply drop away.

Credits: Document deposited in the Fisher Papers at the Barr Smith Library, The University of Adelaide.

3.5.3 *Computing equipment for the blood group survey*

Alongside the information technologies adopted by the Emergency Transfusion Services for the management of the donor records, instrumental in the making of the ABO blood group survey were the computing facilities of the Galton Laboratory. The statistical analysis of the donor data collected at Rothamsted, in fact, required a considerable amount of number crunching and Fisher explicitly offered to the transfusion officers his computing services in return for their cooperation in the survey.[521]

The equipment that moved with Fisher and his staff to Rothamsted Experimental Station in 1939 was much more advanced than the hand-operated Brunsviga machines popular in the laboratory under Pearson.[522] In 1935 the Galton Laboratory received a grant of three hundred pounds for renovating its calculating machines and the following year a further amount of eighty-five pounds. In 1937-38 a contribution of thirty pounds was used to purchase a second-hand Millionaire calculator, Fisher's favourite computing machine since his time at Rothamsted.[523] Already in 1935 the computing room of the Galton Laboratory was satisfactorily equipped according to Warren Weaver, mathematician and influential director of the Natural Sciences Division of the Rockefeller Foundation, who visited the laboratory during the grant application for the serological unit.[524]

In Fisher's time, the computing facilities of the Galton Laboratory were increased also by the use of punched-card equipment. The accounting book of the laboratory records several payments to the British Tabulating Machine Company, the main British corporation in the punched-card business. The company sorted and tabulated cards for the Galton Laboratory, sold it punched-cards and possibly rented it also tabulating equipment.[525]

Since 1936 Fisher's laboratory housed also calculating machines purchased by the British Association Mathematical Tables Committee, the main organisation for table making in Britain. The computing work of the British Association was not related to the laboratory's activity, but Fisher, a long serving member of the Committee, provided workspace and constant advice to the

---

521 Letter from R. A. Fisher to J. R. Copland, 13th September 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

522 For the computing equipment at the Galton Laboratory under Pearson see D. A. Grier (2007), p. 117.

523 Galton Laboratory Ledger, 1934-1939, UCL Record Office, AR 181, Folder 4.

524 "The computing laboratory is now satisfactorily equipped, - Pearson did all his work on an antiquated hand machine." (Warren Weaver Diary, 16th May 1935, RA, Box 16, Folder 220)

525 "BTM [the British Tabulating Machine Company] ran punched card service bureaux all over Britain. These were punched card installations which carried out tabulations on a contract basis for anyone who wished to use them." (M. Croarken (1990), p. 46). The 1930s, in particular the years 1936-1939, represented the heyday of the punched-card machine industry in Britain and scientific computation was one of the fields that benefited from this expansion (see M. Campbell-Kelly (1989), pp. 72-172). For an account of punched-card equipment in scientific computation in Britain see M. Croarken (1990), chapters 3-4.

British Association staff.[526] The contiguity in which the two efforts took place is likely to have been a source of further strength for the table making work of the Galton Laboratory that in the same years was engaged, along with the Rothamsted statistics department, in the preparation of the *Statistical Tables for Biological, Agricultural and Medical Research* (**chapter 2**).[527]

The computing work required by the blood group survey was mainly undertaken by Fisher alone and not by his staff as a computing unit, because the personnel of the laboratory was quickly dispersed after the move to Rothamsted. Early in 1940 Mary Karn, the computer of the laboratory, found mathematical work elsewhere and Horace Norton, one of Fisher's assistant statisticians, returned to the U. S. In April 1940 also Rose Tysser, the computer of the British Association, resigned her position. Early in 1941 left Wilfred Stevens, Fisher's other statistical assistant, and Alexandre Fabergé his genetical assistant. Of the original staff remained only Fisher's secretary Barbara Simpson and the genetical assistant Sarah North.[528] When Fisher left as well, taking up his appointment at Cambridge in 1943, the computing equipment of the Galton Laboratory remained at Rothamsted, and was used by the staff of the local statistics department until the end of WWII, when the calculating machines were returned to University College.[529]

For the computing work related to the blood group survey Fisher probably used his Millionaire or, possibly, the more advanced electric Monroe calculators available in the computing equipment of the laboratory. If the number crunching for the survey was certainly aided by calculators, no mechanisation of the information processing related to the donor records can be suggested. Fisher and his secretary must have sorted and counted enrolment forms and index cards by hand. The sheets of data surviving in the Blood Group Survey correspondence, in fact, are just columns of figures, handwritten or typewritten, for male and female donors for each phenotype and each geographic area. Sometimes the patient hand that sorted the transfusion records used also the tick method for counting the donors, strengthening the idea that the process did not benefit from any mechanical aid.[530]

---

[526] For the computing machines of the British Association at the Galton Laboratory see M. Croarken (2003).
[527] R. A. Fisher and F. Yates (1963).
[528] J. Fisher Box (1978), p. 375.
[529] "We are needing additional machines because of increased staff and because for the last two years we have been able to use machines belonging to the Galton Laboratory which will shortly have to be returned to their owners." (Letter from D. J. Finney to W. R. Black, 7th July 1945, TNA, MAF/33/333)
[530] Data sheets for the Slough Depot donors, undated, R. A. Fisher Papers, BSL, The University of Adelaide.

3.6 *The ABO blood group survey: data analysis and results*

3.6.1 *Statistical methods for the calculation of gene frequencies*

"[T]he immediate result of the examination of the blood groups of a population is a series of numbers each expressing the total of individuals falling into a given serological class. Each of such classes [...] constitutes a phenotype", states Arthur Mourant in *The Distribution of the Human Blood Groups*.[531]

Counting forms and cards of the Emergency Transfusion Services gave Fisher precisely the frequencies of the ABO phenotypes among the donors, but for him as a geneticist mattered the frequencies of the three allelic genes O, A and B. The gene frequencies offered a more direct way of comparing the blood group content of the population and they could be of immediate use in understanding whether a batch of data contained a reasonable distribution of the ABO blood groups and thus the serological technique was reliable. As previously mentioned, the genotype frequencies could not be derived from serological tests, but they could be extracted from the phenotype data with statistical methods.

In the mid-1920s Felix Bernstein had established a first set of formulae for the computation of the gene frequencies. In a new in-depth contribution to the statistical treatment of blood groups published in 1930 Bernstein proposed an improved estimate of the gene frequencies and explicitly adopted the chi-square test that Fisher had discussed in his 1925 textbook, *Statistical Methods for Research Workers*, for testing the agreement between observations and hypothesis. As pointed out by Fisher the usefulness of the chi-square test regarded "the comparison of the numbers actually observed to fall into any number of classes with the numbers which upon some hypothesis are expected."[532] That sounded perfectly fit for the blood group case because the aim of the test of significance was precisely to determine whether the phenotype figures were in agreement or not with the expectations for the gene frequencies. Notably, Fisher's *Statistical Methods for Research Workers* is the only mathematical reference – apart from Bernstein's previous papers – mentioned in the bibliography of Bernstein's 1930 article.

When the Galton Laboratory Serum Unit began to publish the results of its serological researches in 1938-1939, Bernstein's contributions represented a necessary reference.[533] However, before WWII the unit was interested in the investigation of the system OA_1A_2B, an extended

[531] A. E. Mourant (1954), p. 4.
[532] F. Bernstein (1930). Quotation taken from R. A. Fisher (1946), p. 78.
[533] R. A. Fisher and F. Bernstein were correspondents since 1926. (Correspondence with F. Bernstein, May 1926 – July 1954, R. A. Fisher Papers (digitized), BSL, The University of Adelaide)

version of the ABO system described by Bernstein, and the mathematical formulae for extracting the gene frequencies should be modified accordingly.[534]

Furthermore, in 1938 Wilfred Stevens presented a refinement of Bernstein's method with formulae based on Fisher's principle of maximum likelihood.[535] Through Fisher's principle it was possible to derive an efficient statistic able to give a better accuracy, but at the cost of more complicate computations, as it is evident confronting the mathematics in Bernstein's and William's papers.[536] While in the former just a bit of straightforward algebra was required, in the latter partial derivatives should be obtained and a matrix compiled. Unsurprisingly, even the serologists at the Galton Laboratory were more familiar with Bernstein's methods rather than with the more elaborate formulations developed by their statistical colleagues and left to Stevens the examination of their figures with the maximum likelihood method.[537]

A further adjustment in the statistical analysis of the ABO blood groups was required by the wartime survey. Examining the transfusion records, Fisher was confronted by the systematic errors in the determination of the group AB. Unlike the population studies done before WWII by highly qualified personnel on samples of more manageable size, the wartime donor records were the product of mass collections often performed by staff hastily trained and thus liable to make clerical or technical mistakes. For dealing with this issue, Bernstein's method for the calculation of the chi-square was not of much help, remarked Fisher.

> What influences me is that, as there is only one degree of freedom for the Bernstein criterion, every sort of disturbance, either in grouping any of the four wrongly, or due to race mixtures, etc, will either increase or decrease the deficiency of AB, and all we have to observe is the balance of a number of slight disturbing influences.[538]

In order to overcome this limitation Fisher proposed a new method for testing the agreement between observed and expected data (Appendix 3.4). Only the phenotype figures for the O, A and B groups were used to calculate the gene frequencies and to determine the expected value of the group AB. Such figure was then used to evaluate the chi-square to describe the fit of the data

---

534 Even during the blood group survey George Taylor did further tests on blood samples, especially from Scotland, to improve the examination of the ratio A_1:A_2. (Letter from G. L. Taylor to R. A. Fisher, 5th April 1940, R. A. Fisher Papers, BSL, The University of Adelaide)

535 W. L. Stevens (1938). The application of the principle of maximum likelihood requires preliminary estimates of the gene frequencies on which the corrections worked out with this method are then applied. Thus, easy methods of calculating the gene frequencies, like Bernstein's, are in any case requested. Before WWII, Fisher also suggested a statistical method for tackling the family material collected by the Galton Laboratory Serum Unit (G. L. Taylor and A. M. Prior (1939a), p. 19). More information on W. L. Stevens can be found in his obituary written by F. Yates (F. Yates (1959)).

536 Fisher defined the efficiency of a statistic as "the ratio which its intrinsic accuracy bears to that of the most efficient statistics possible. It expresses the proportion of the total available relevant information of which that statistic makes use". (R. A. Fisher (1922a), pp. 309-310)

537 G. L. Taylor and A. M. Prior (1938b). W. L. Stevens (1938).

538 Letter from R. A. Fisher to G. L. Taylor, 27th April 1942, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

with expectations. The new method was used throughout the wartime, not only by Fisher, but also by his co-workers, Edwin Hart and John A. Fraser Roberts, and it was published in 1946 in the paper that summarised the blood grouping work done by the Galton Laboratory Serum Unit in Cambridge during WWII.[539]

3.6.2 *From medical to genetical records*

In the warfare there was a natural tendency to select donor records because not all the blood groups could be employed to the same extent. AB's were extremely rare and B's rare, while A's and O's accounted for over eighty per cent of the British population. In some cases the transfusion services kept records only of the O donors that were of universal use and so better suited to the wartime situation.

Moreover not all the people who filled in an enrolment form could become donors. People with unsuitable veins, low haemoglobin, positive Wasserman reaction, i.e. affected by syphilis, were automatically excluded by the donor panel.[540] Constantly Fisher checked with the transfusion officers that all the data, even of the people grouped but unsuitable as donors, were sent to him.[541] The selection in the donor panel represented a potential risk for the survey because the sample could not be fully representative of the whole population. Thus Fisher always reminded to his correspondents that "[i]n a population sample, one wants people chosen at random, at least in the sense that nothing is known in respect of their blood group or relationship with others of known group."[542] Beyond intended selection, in some cases records could also be mislaid during the preparation of indexes by filing clerks, a further problem that Fisher had to confront.[543]

Besides the selection of records, another relevant issue was the duplication of the data. In order to guarantee the success of his survey Fisher had to contact more than one person in the same geographic area with possible misunderstandings.[544] When the donor numbers were relevant and the records could be mailed to Rothamsted only bit by bit, the risk of duplication came from the batches of records sent twice. In such cases Fisher constantly warned his

[539] A. M. Dobson and E. W. Ikin (1946), p. 221; E. W. Hart (1943), p. 93; J. A. Fraser Roberts (1947), p. 109.
[540] Letter from D. F. Cappell to R. A. Fisher, 25th October 1939, R. A. Fisher Papers, BSL, The University of Adelaide.
[541] Letter from R. A. Fisher to D. F. Cappell, 23rd October 1939, R. A. Fisher Papers, BSL, The University of Adelaide.
[542] Letter from R. A. Fisher to P. L. Oliver, 14th December 1939, R. A. Fisher Papers, BSL, The University of Adelaide.
[543] Letter from J. Vaughan to R. A. Fisher, 10th November 1939, R. A. Fisher Papers, BSL, The University of Adelaide.
[544] Letter from R. A. Fisher to A. C. Scott, 15th January 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

correspondents that "[a]lmost all conclusions to be drawn from the data may be seriously affected if duplicate countings of the same material cannot be absolutely avoided".[545] And to prevent mistakes, since the beginning of the survey he recommended to George Taylor to send "further data to the Galton Laboratory at Rothamsted, as I am always afraid of getting somebody's incompletely entered up".[546]

The first crucial contribution that Fisher gave to the blood group survey was his statistical outlook in ensuring that no duplication or selection had taken place among the donor records. This control came out of a patient work, with the help of his secretary, of sorting and counting forms and, less often, index cards working even batches of five thousand records per week.[547] In some cases the transfusion services sent to Rothamsted already lists of figures for the donors in their panel. In these cases Fisher was even more careful in scrutinise them, knowing that a "serious accident [...] may occur, and [...] we are entirely unable to check, when we merely receive totals, as from some of the sectors."[548] Incomplete records were to represent a limitation also for the data from the South-West of England examined by John Fraser Roberts. The omission of the age in the enrolment forms of the area, especially for Cheltenham and surroundings, reduced of one-third the overall number of records available for comparison by age.[549]

Fisher had no direct control on the discrepancies determined by the misclassification of donors – the director of the Maidstone depot, Montague Maizel, estimated a two per cent error in the initial groupings due to clerical mistakes or actual problems in the serological technique – and constantly asked George Taylor's advice on this point.[550] Sometimes even the transfusion officers warned Fisher about the uneven quality of their groupings. Janet Vaughan, for example, pointed out to Fisher that not all the batches of data from her depot, Slough, were of the same standard in relation to the serological technique

---

545 Letter from R. A. Fisher to W. I. Stanbury, undated, R. A. Fisher Papers, BSL, The University of Adelaide.

546 Letter from R. A. Fisher to G. L. Taylor, 23rd November 1939, R. A. Fisher Papers, BSL, The University of Adelaide. See also Letter from R. A. Fisher to G. L. Taylor, 24th January 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

547 "we have found the original forms most convenient for the purpose, although there is a trifling risk of occasionally counting the same person twice from different places, which risk might conceivably be eliminated in the secondary indexes which many people have prepared. Our only experience of index cards was, in fact, extremely troublesome; they were actually very badly done. Consequently, if Manchester has ages, it would suit us very well if they sent us packets of, say, 5000 original forms" (Letter from R. A. Fisher to G. L. Taylor, 14th November 1941, R. A. Fisher Papers (digitized), BSL, The University of Adelaide)

548 Letter from R. A. Fisher to G. L. Taylor, 15th November 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

549 J. A. Fraser Roberts (1947), p. 114.

550 Letter from M. Maizel to R. A. Fisher, 20th October 1939, R. A. Fisher Papers, BSL, The University of Adelaide. For Fisher's requests of advice on technical problems in serology see Letter from G. L. Taylor to R. A. Fisher, 2nd November 1939; letter from G. L. Taylor to R. A. Fisher, 12th December 1939. Both letters are in the R. A. Fisher Papers, BSL, The University of Adelaide. See also Letter from R. A. Fisher to R. R. Race, 6th November 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

> I send you another big batch of forms. [...] I think [they] are probably not as well grouped as the rest you have had from here recently. They were done by the Sutton Slide Technique which is more liable to error. Owing to heavy pressure of work we had regretfully to allow this to be done knowing they would be rechecked at the time of bleeding.[551]

Dealing with the variable quality of the blood grouping methods employed by the transfusion centres and extracting from the transfusion records gene frequencies was the second crucial contribution that Fisher gave to the blood group survey and derived from a combination of his expertise in statistics and genetics.

As I have mentioned, Fisher developed a statistical method *ad hoc* for limiting the overall effect of the errors in the determination of the AB group and so improved the value of the transfusion records for his genetical analysis. Once the gene frequencies had been extracted, the medical information gathered from the donor records had been completely transformed in genetical information and the ABO blood groups, from medical tools for blood transfusion, had become an instrument for studying human genetic diversity. For this reason forms and index cards of the Emergency Transfusion Services can be considered boundary objects, that is

> objects that both inhabit several communities of practice and satisfy the informational requirements of each of them. In working practice, they are objects that are able both to travel across borders and maintain some sort of constant identity. They can be tailored to meet the needs of any one community [...]. At the same time, they have common identities across settings.[552]

The transfusion records proved to have such plasticity and endurance: their materiality was unaffected crossing the border of the Emergency Transfusion Services to reach the temporary accommodation of the Galton Laboratory at Rothamsted Experimental Station, but the different elaboration of their information and Fisher's attitude towards such records guaranteed a second life for forms and index cards that became tools for the investigation of British ethnology.

3.6.3 *Exploring human biological diversity by numbers*

The "racial, namely ethnic, plurality in British history – of Saxons, Celts, Normans, Irish, Welsh, Scots and English, with numerous subdivisions" was well accepted in British anthropology at the beginning of the twentieth century.[553] Blood groups, however, had never been systematically employed to explore such ethnic inhomogeneity and before WWII only a few

---

[551] Letter from J. Vaughan to R. A. Fisher, 24th July 1940, R. A. Fisher Papers, BSL, The University of Adelaide.
[552] The idea of boundary objects was originally presented by J. R. Griesemer and S. Leigh Star (1989). The quotation, instead, is taken from the later book by G. C. Bowker and S. Leigh Star (2000), p. 16, in which boundary objects are discussed in the broader context of classification in the sciences.
[553] E. Barkan (1992), p. 23.

studies, involving a limited number of people, had been published on the subject giving only sparse figures on the distribution of the ABO blood groups in Britain (Appendix 3.4).[554]

Fisher's survey filled this gap using the extensive records of the Emergency Transfusion Services, but the analysis that Fisher made of the data collected is open to disputes. As pointed out by Lisa Gannett and James Griesemer, the use of blood groups for uncovering anthropological and ethnological relations is not an objective procedure. Even though blood groups are independent from the environment and their transmission is known, subjective judgements are instead "implicated at each link of the chain of reference that connects blood samples in the field to published tables and maps of blood group frequencies".[555]

Racial groups have to be construed theoretically in order to make sense of the gene frequencies extracted through the fieldwork. One interesting example in this sense is discussed by Veronika Lipphardt and concerns the study of the Jewish community in Rome conducted in the 1950s by the serologist Leslie Dunn and his son, the sociologist Stephen Dunn. Lipphardt argues that "biologists used historical, social and administrative data to construct what they considered to be an isolated population that would subsequently serve the purpose of studying evolutionary and genetic processes", that is bio-historical narratives guided sampling procedures in serological research.[556]

In Fisher's survey the basis for interpretation was constituted by conjectures already formulated by ethnologists, such as the Scandinavian influence in Scotland, or by linguistic instruments as the use of surnames. This second criterion inspired Fisher's first publication related to the ABO blood group survey, co-authored with Janet Vaughan. In the area administered by Vaughan's Depot, Slough, blood donors with Welsh surnames had a significantly lower proportion of A's than the remaining part of the panel.[557] In the preparation of this study Fisher was aided by the (amateur) genealogist Byrom S. Bramwell, a member of the Eugenics Society. Bramwell suggested the eight Welsh surnames for the study on account of Guppy's book, *The Homes of Family Names*, where frequencies of the surnames were presented divided by geographic area. The use of family names in the analysis of the blood groups data was endorsed also by Fisher's co-workers in the survey, the geneticist John Fraser Roberts that applied it to the

[554] For the limited development of blood group research in Britain before WWII see W. H. Schneider (1995) and W. H. Schneider (1996b).
[555] L. Gannett and J. R. Griesemer (2004), p. 148.
[556] V. Lipphardt (2010), p. 307.
[557] R. A. Fisher and J. Vaughan (1939).

analysis of data from Wales and the Army Captain, Edwin Hart, for the study of the figures collected by the transfusion services of the Army in Northern Ireland.[558]

Fisher tested even the association of blood groups frequencies with sex, but in this case the results were ambiguous. If the data provided by the donors in the South West of England suggested an excess of A's among men and an excess of O's among women, data from other regions only mildly confirmed this result or even disproved it.[559] Also the investigation of a relation between blood groups and age did not bring to any definite result. Fisher justified this inconclusiveness with the heterogeneity of the British population for the proportion of A donors and with the internal migrations "so that at any one place, such as Slough, donors in the fifties may not have been born in the same part of the country as donors in the twenties."[560]

The data gathered by Fisher and his co-workers showed an overall north-south gradient in the distribution of the O and A blood groups with an increase in O's and a decrease in A's going from the South of England to Scotland, where the phenotypic ratio A/(A + O) resulted very low, equal to 39.7. Scotland was an area in which Scandinavian influences were anticipated, but in the 1940s the ratio A/(A + O) for Norway, Sweden and Denmark, was respectively 58.0, 58.6 and 50.0, contradicting thus the widespread belief in a Scandinavian influence on Scottish ethnology.

The low ratio for Scotland, instead, was compatible with the ratio for Iceland, 36.6, an area that had been colonized by Scandinavian people many centuries before. Therefore, in their letter to *Nature* in April 1940 Fisher and Taylor, in order to rescue a Scandinavian descent for Scottish people, had to point out that they were referring to proto-Scandinavian influences – evident among Iceland inhabitants – and that "we must distinguish it sharply from the modern Scandinavian peoples, which have evidently changed greatly by infiltration from central or eastern Europe, since the Viking period."[561]

According to Elazar Barkan, Fisher's interpretation of Scottish ethnology was linked to his British ego and his racist, namely eugenics, attitude that "confused biological with historical time-scales".[562] Certainly Fisher was a convinced eugenicist, but I want to argue that his interpretation

---

[558] J. A. Fraser Roberts (1942a, 1942b), E. W. Hart (1943). Actually, Fisher was rather disappointed by Hart's paper. "I should have preferred a more extensive summary, or perhaps better a discussion underlining the ethnographical importance of the English surnames being still distinguishable, for I think that is the most remarkable result of the enquiry and there is literally no discussion." (Letter from R. A. Fisher to G. L. Taylor, 12th March 1943, R. A. Fisher Papers (digitized), BSL, The University of Adelaide)

[559] R. A. Fisher and J. A. Fraser Roberts (1943).

[560] Letter from R. A. Fisher to W. C. Boyd, 8th May 1953, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

[561] R. A. Fisher and G. L. Taylor (1940). The suggestion of a letter to *Nature* was made by G. L. Taylor: "What about a short preliminary note, say a letter to Nature, about the changes in the blood-group percentages as we go from the South to Scotland?" (Letter from G. L. Taylor to R. A. Fisher, 24th February 1940, R. A. Fisher Papers, BSL, The University of Adelaide)

[562] E. Barkan (1992), p. 226-227. Barkan reminds his readers that at the beginning of the twentieth century, race "had

of Scottish ethnology was more influenced by his limited possibility to explore human diversity through the blood group data, rather than an expression of his eugenic concerns.[563]

Fisher was a statistician and geneticist and the co-author of the letter, Taylor, a serologist who explicitly declared "to know nothing about much of the ethnology" Fisher mentioned in the communication to *Nature*.[564] Even Fisher who drafted the letter had just a basic grasp of ethnology. The only instrument on which he could rely were the figures: on the one hand the gene frequencies extracted from the records of the Emergency Transfusion Services and on the other the previous collections of data prepared by serologists and anthropologists that reported such frequencies around the world. In this sense Fisher literally explored human biological diversity by numbers.

The preparation of the letter on Scottish ethnology required, in fact, the use of the blood group figures collected before WWII by William Boyd and, for Iceland, by Alexander Wiener, and a comparison with the about eleven thousand data available on Scotland. The Iceland figures were closer to the Scottish ones and it was the proximity of the figures – "the only foreign sample we know of comparable to the new Scottish data is from Iceland" – and the lack of other possible ethnological explanations that inspired Fisher's conclusion. The interpretation of the proto-Scandinavian influence seems an attempt to rescue an established knowledge, apparently contradicted by the data, rather than a willing choice to gratify Fisher's British ego.

In this exploration of human diversity using only numbers as a guiding principle lies the main limitation of Fisher's wartime survey. Fisher's statistical and genetical knowledge was sound, but the lack of anthropological and ethnological expertise greatly reduced the interpretative possibilities and constrained the formulation of new theoretical frameworks for understanding the results gathered from the ABO gene frequencies. As a matter of fact, Fisher never published during or after WWII an extended article on the results of the survey, but entrusted partial results to brief communications appeared usually as letters in *Nature*.

The collection of about two hundred thousand blood group data made available at the conclusion of the war by the Galton Laboratory Serum Unit was probably the main contribution of Fisher and his co-workers to the study of blood groups in Britain. Considering that before WWII, about 1,355,000 people had been involved in blood grouping tests worldwide and among them less than six thousand were in Britain, the new data available at the end of the war were a

a far wider meaning than at present, being used to refer to any geographical, religious, class-based or color-based grouping". (E. Barkan, 1992, p. 2) Due to this broader definition the eugenicist R. A. Fisher can be considered a 'racist'.

563 Some of the limitations intrinsic in Fisher's work are also addressed by J. Marks (1996) in his essay on the reception of serological studies in American anthropology.

564 Letter from G. L. Taylor to R. A. Fisher, 13th March 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

consistent step ahead.[565] The figures were gathered mainly from the population in Cambridge and from soldiers of the Royal Air Force who had offered their blood to the Galton Unit in the preparation of testing serum for the empanelling centres. As the unit had grouped the volunteers in its panel examining both their cells and serum, the results were much more reliable than the ones gained by the transfusion centres. The military figures, Fisher suggested, could be considered a trustworthy average of the blood group distribution in the British population.[566]

3.6.4 ***Improving the efficiency of the Emergency Transfusion Services***

The investigation of British ethnology was only one of the potential outcomes in the examination of the donor records. In the standard letter sent to the Scottish transfusion centres, beyond the genetical relevance of the data that could be gathered from the records, Fisher explicitly mentioned that the statistical examination of the ABO figures obtained by each centre could be

> important in another respect in that comparison of the frequencies recorded from different laboratories working in the same town, provide a good first check on the accuracy of the grouping, and this, as is well known, may fail either through the inexperience of some of the workers, or through unsuspected lack of potency in the sera used.[567]

Serum of poor quality, laboratory mistakes, clerical errors all compromised the reliability of the donors' grouping. Moreover, according to Fisher's analysis, the misgroupings did not err "on the safe side from the point of view of transfusion, as one might possibly have expected if workers were concerned to test large numbers as rapidly as possible, with a view to finding a number of reliable O donors".[568] In some cases the error was so evident to call for urgent intervention, like in the case of a bunch of data from Glasgow with an alarming proportion of B's.[569]

Even the faith in his colleagues of the serologist George Taylor was shaken by the mixed results obtained by the transfusion centres.

> Before our experience of results sent in by all sorts of workers, I should have been inclined to agree that most hospital pathologists could be entrusted to do the simple ABO groupings; they can do no such thing. They would be saved from all sorts of mischief if they would, every time they used them, see that their reagents are what they think them to be, by using controls – we find it necessary to do so.[570]

[565] Global blood group data are taken from W. H. Scheneider (1966b), p. 287.
[566] The figures gathered from the soldiers of the Royal Air Force were presented as an average of the British population in R. R. Race and R. Sanger (1954), p. 20.
[567] Letter from R. A. Fisher to J. R. Copland, 13th September 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.
[568] Letter from R. A. Fisher to G. L. Taylor, 26th January 1940, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.
[569] Letter from R. A. Fisher to G. L. Taylor, 11th November 1939, WL, SA/BGU/F1/1/1.
[570] Letter from G. L. Taylor to R. A. Fisher, 4th March 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

The transfusion officers were aware of these pitfalls of serology and paid attention to the discrepancies that Fisher found in his calculations. Sometimes, they even asked explicitly to him and to his co-worker, Taylor, about the statistical examination of their figures in order to monitor possible inconsistences. Both Fisher and Taylor were very willing to cooperate and never refused their help. When the Glasgow physician Gilbert Millar wrote Taylor asking advice,

> Lastly, to close this already overlong letter, I wonder if Professor Fisher could be bothered to give me some guidance in a statistical problem. I think, in point of fact, that I have tackled it in the appropriate manner, but, not being at all confident in my statistical insight and as the methods applicable are those he has made his special province, I felt I would like to sound him on the matter. But before doing so it struck me that you would be in a position to tell me whether he would be likely to view such an approach with an unduly jaundiced eye!

Taylor immediately invited him to contact Fisher, who thanked his co-worker "for answering Dr Millar exactly in the sense I should like."[571]

The statistical control of the consistency of the donor records was strategic in the re-definition of Fisher's research programme in serology during WWII and in the establishment of his collaboration with the Emergency Transfusion Services for the ABO blood group survey. However, as of September 1939, Fisher considered only an "administrative convenience" this accessory opportunity of the survey and offered a statistical feedback to the transfusion officers, just as a lure and compensation for their collaboration in the data collection.[572]

By July 1941, when it was clear that no work of national importance would be given to the Galton Laboratory, Fisher's evaluation of the survey had changed.[573] In controversy with the authorities of University College London that wanted to dismiss, Barbara Simpson, Fisher's secretary with whom he shared the task of sorting and counting the donor records, the statistician asked to Arthur Landsborough Thomson, secretary of the Medical Research Council, to defend Simpson's position on account of her contribution to the survey, a 'war work' that improved the efficiency of the transfusion services.[574]

Thus, Fisher eventually singled out the survey not only as a precious form of data collection and as a contribution to the study of British ethnology, but as a war work in which his expertise

[571] Miller's quotation is in the Letter from G. L. Taylor to R. A. Fisher, 8th November 1940, R. A. Fisher Papers, BSL, The University of Adelaide. Fisher's reply is the letter from R. A. Fisher to G. L. Taylor, 11th November 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

[572] Letter from R. A. Fisher to J. R. Copland, 13th September 1939, R. A. Fisher Papers (digitized), BSL, The University of Adelaide.

[573] A further evidence of the gradual recognition of the national value of Fisher's survey is in the pamphlet *Science in War*, published in 1940. The book was written by a group of scientists close to the operational researcher Solly Zuckerman. Fisher's survey is mentioned in the book – possibly at the suggestion of Frank Yates – for its value in the study of human heredity, while its relevance for increasing the efficiency of the Emergency Transfusion Services, more relevant in the general perspective of the book, is not stated yet. Another year was to pass before Fisher began to present his interaction with the Emergency Transfusion Services on both perspectives.

[574] Letter from R. A. Fisher to A. Landsborough Thomson, 11th July 1941, TNA, FD/1/5845; Letter from R. A. Fisher to A. Landsborough Thomson, 23rd July 1941, TNA, FD/1/5845.

in statistics and genetics could contribute to improve the activity of the wartime organisation. In this perspective, the blood group survey can be assessed also as another case in which statistical expertise was deployed during WWII to gain efficiency, a goal very much alike to the one that featured in operational research.

3.7 ***Ronald Fisher's lifetime collaboration with serologists***

Since the opening of the Galton Laboratory Serum Unit Ronald Fisher became a consultant of serologists, helping his co-workers in the application of statistical methods to the blood group data. Most of the papers published by the Galton Laboratory Serum Unit before WWII acknowledge the contribution of Fisher or his statistical assistants for dealing with the data analysis.[575]

As discussed above, the collaboration strengthened during the warfare, when inexperienced pathologists and hospital physiologists were called to do blood groupings on a large scale. Through the survey a new class of research workers became aware of the potentialities offered by the analytical tools of statistics, as a simple chi-square test was enough to understand whether a set of donors should be regrouped or not.

George Taylor contributed to disseminate Fisher's statistical methods for testing the consistency of blood groups data among his fellow serologists. Even before the war, he began to do calculations on behalf of his colleagues – "almost the last serious work I did at the Galton Laboratory was to send off Friedenreich [the Danish serologist who had contributed to Taylor's laboratory training] 'touristing', whilst I did the χ^2 analysis of his [Danish] figures".[576] And during the war he went on promoting the use of such statistical methods in the publications that presented the work of his unit.[577] But even Taylor, despite Fisher's close advice, struggled to

575 "Our thanks are due to Prof. R. A. Fisher and Mr W. L. Stevens for much help in the treatment of data" (G. L. Taylor and A. M. Prior (1938b), p. 360); "A new statistical method suggested by R. A. Fisher of dealing with data on the extended *ABO* system has been used and explained" (G. L. Taylor and A. M. Prior (1939a); "We are grateful to W. L. Stevens and H. W. Norton for advice on the statistical aspect of this paper" (G. L. Taylor and A. M. Prior (1939b)); "Thanks are due to W. L. Stevens for making the heavy computations of the maximum likelihood treatment of the A_1A_2BO data" (E. W. Ikin et al. (1939), p. 411).

576 Letter from G. L. Taylor to R. A. Fisher, 13th March 1940, R. A. Fisher Papers, BSL, The University of Adelaide.

577 "The agreement between the results observed and expected can be used to indicate the reliability of a worker's methods and the agreement can be tested for significance by the well known χ^2 test. Fisher (1941) [*Statistical Methods for Research Workers*] has constructed tables of χ^2 values corresponding to various levels of probability and different numbers of degrees of freedom. By calculating the χ^2 value for a batch of results and by consulting the tables the limits between which the probability of the observed results lies can be seen at once." (G. L. Taylor et al., 1942, pp. 81-82)

make sense of statistical methods applied to serology and admitted to make use of them without really understanding the point:

> I take it that you agree that the test of significance which you proposed has been performed by me properly. I must say that I was able to do it only cookery-book-wise from your recipe and I have spent several hours trying to get at the meaning of what I did. [...] If you can conveniently explain this to me I shall be grateful. I am sorry to be so stupid but I really have tried at home with your book and Mather's two books.[578]

Taylor prematurely died in 1945, but Fisher carried on in the aftermath of WWII his collaboration with the former members of the Galton Laboratory Serum Unit. For over twenty years he kept in touch with Robert Race, the young serologist that Taylor trained, and went on advising him on statistical matters.[579]

If Taylor had tried to understand statistics, Race, who was to become an authority on blood groups, did use it, but did not struggle to make sense of it. "I wish I could appreciate the beauty of this maximum likelihood method. I suppose it is an intellectual triumph", he wrote to Fisher, but was not eager to make further enquiries.[580] Troubles came with the new co-worker, Ruth Sanger, who joined Race after the war and was less shy of mathematics. To Race's dismay, she began to ask questions related to calculations for the Rh system: "Miss Sanger has asked some awkward questions about the meaning of the 'matrix of sampling covariance of these estimates'. I said I was too busy to explain. I will be able to keep this up for a day or two, then perhaps I had better have 'flu, and after that she may have forgotten about it."[581] To the relief of the anguished Race, Fisher patiently explained to him how the matrix was made, what it represented and how 'Miss Sanger' could make the best use of it.[582]

Fisher's attitude towards Race was always kind, almost paternal, but from time to time he could not refrain from poking fun at Race's clumsy way of handling mathematics – "Fancy your sum coming out correct to within a fiftieth of a child. I wonder which part you have amputated" – and hinted "[y]ou really ought to tell Drury that you want a nice tame mathematical assistant to do this sort of thing for you".[583]

But if Fisher was amused by the little troubles of Race with mathematics, Race's colleagues looked instead at the serologist as a promoter of the application of statistics to the study of blood

---

578 Letter from G. L. Taylor to R. A. Fisher, 26th March 1943, R. A. Fisher Papers, BSL, The University of Adelaide. "Your book" is likely to be Fisher's *Statistical Methods for Research Workers*, while the other author mentioned is the geneticist Kenneth Mather, former assistant of Fisher at the Galton Laboratory.

579 After WWII Fisher was unable to regain control of his unit that moved from Cambridge to the Lister Institute in London. After Taylor's death, Robert Race was in charge of the unit.

580 Letter from R. R. Race to R. A. Fisher, 1st July 1947, WL, SA/BGU/F4/3/1.

581 Letter from R. R. Race to R. A. Fisher, 17th March 1947, WL, SA/BGU/E7/1.

582 Letter from R. A. Fisher to R. R. Race, 18th March 1947, WL, SA/BGU/F1/1/1.

583 Letter from R. A. Fisher to R. R. Race, 1st February 1949, WL, SA/BGU/E7/1. Alan Drury was the director of the Lister Institute.

groups and Race's obituary reminded that "[h]is early applications of statistical techniques, while protesting that he really didn't understand any mathematics, showed the way to the solutions of many serological and genetic problems arising from blood banking."[584]

Thus, even the serologists that collaborated closely with Fisher struggled to grasp statistical methods, but in the same time they contributed to spread these mathematical tools among their colleagues using tests of significance in their publications and illustrating methods for the analytical treatment of serological data. How much they were interested in understanding the potentialities and limitations of such statistical methods or rather in having mathematical recipes ready to use it is open to debate.[585] Serology textbooks, like *Blood Groups in Man* written after the war by the before-mentioned Robert Race and Ruth Sanger, or collections of blood groups data, like Mourant's *The Distribution of the Human Blood Groups*, just enunciated these statistical methods and gave some hints for their practical use, but did not go any further in their discussion, and carefully skipped the more technical formulations like the ones derived from Fisher's maximum likelihood principle. Statistics was integrated in the cultural baggage of serology, but never became familiar to its practitioners as their laboratory paraphernalia.

On the other hand, a mutual exchange of competences was necessary to make statistical methods useful for serology. In order to formulate mathematical tools for dealing with blood group data Fisher had to collaborate closely with his colleagues engaged in the laboratory work. For example, the shortage of AB's in the blood group survey was widely discussed in the wartime correspondence between Fisher and George Taylor, with Fisher consulting Taylor on the technical reasons that could account for the discrepancy and Taylor volunteering to retest some blood samples in order to shed light on the question.[586]

Thus, the integration of serological and statistical expertise was not straightforward, but certainly it was the winning route, as gene frequencies could be extracted from the phenotype data only through statistical methods, which were also powerful instruments for an immediate feedback on the quality of the laboratory work.

---

[584] E. Hackel (1984), p. 1.

[585] See also the comments on the reception of the *Statistical Tables for Biological, Agricultural and Medical Research* among research workers discussed in chapter 2 and H. M. Marks (2000) for the reception of statistics among physicians (pp. 129-163).

[586] R. A. Fisher Papers (digitized), BSL, The University of Adelaide: Letter from R. A. Fisher to G. L. Taylor, 7th December 1939; Letter from R. A. Fisher to G. L. Taylor, 26th January 1940. R. A. Fisher Papers, BSL, The University of Adelaide: Letter from G. L. Taylor to R. A. Fisher, 13th December 1939; Letter from G. L. Taylor to R. A. Fisher, 19th February 1940; Letter from R. A. Fisher to G. L. Taylor, 21st March 1940.

3.8 *Conclusion*

The widespread use of blood groups in medicine, genetics, anthropology during the first half of the twentieth century required an instrumental alliance between serology and statistics. In this chapter I have examined such alliance using as a case study the collaboration between the statistician and geneticist Ronald Fisher and the staff of the serum unit that he set up at the Galton Laboratory in the 1930s. Among the contributions to serology of Fisher and his co-workers, I have singled out the wartime survey of the ABO blood group distribution in Britain as an ideal context for discussing the complex relation between data collection, statistical methods for the analysis of blood group data and computing, which featured in blood group research.

Fisher's survey benefited from the material nature of enrolment forms and index cards, the information technologies adopted by the Emergency Transfusion Services. These records were humble paper tools that could be mailed to Rothamsted where Fisher and his secretary patiently sorted and counted them, with the aim of preserving and recording data valuable for human genetics. This physical journey corresponded to the immaterial travel of the information collected in the donor records that, I have argued, should be considered boundary objects able to move from medicine to genetics.

The ABO blood group survey was a contingency of WWII, but it was also heavily indebted with the idea of blood management that had started in Britain at least twenty years before. The information technologies adopted during the warfare, in fact, were borrowed from the office organisation of the Red Cross Blood Transfusion Service, the charitable initiative that had been active in Britain before the war in providing blood donors. The Red Cross service was organised with an accounting philosophy in mind: for each donor it should be possible to tell at once when and where he or she had given his or her blood, who had received it and the outcome of the transfusion. The Emergency Transfusion Services borrowed this attitude and adopted a systematic and reliable information management that was a pre-condition for the study of the blood group distribution using the transfusion records.

The data collection, however, was only the first step. The phenotype figures obtained from the records did not give by themselves useful information for the study of human biological diversity in Britain. What mattered were the gene frequencies that could be extracted from the phenotype data using statistical methods. Mathematical formulae for extracting such frequencies had been established since the mid-1920s by the mathematician Felix Bernstein and refined during the 1930s by Fisher and his co-workers, but the data analysis in the blood group survey required a new method for the calculation of the gene frequencies, because the transfusion records were not reliable for the determination of the rarest blood group, AB. In targeting the statistical methods to the quality of the data available, Fisher gave a crucial contribution for transforming the medical information of the donor records in genetic information useful for understanding human biological diversity in Britain.

The interpretation of the gene frequencies gathered in the ABO blood group survey, however, was constrained by Fisher's limited possibility to set interpretative schemes, as argued in the investigation of the Scandinavian descent in Scotland. Despite this intrinsic limitation, the survey can be considered successful for the methods that Fisher introduced in the study of British ethnology. While blood groups had been little used in the study of the population before WWII, after the conflict blood transfusion records were routinely employed by British geneticists and anthropologists and the Royal Anthropological Society began to promote the use of blood groups data.[587]

The survey, aside from its contribution to the study of British ethnology, can also be framed in the wider scenario of the deployment of statistical expertise in WWII. Fisher had hoped for the Galton Laboratory an involvement in computing work of national importance, similar to the contribution given by the laboratory under the guidance of Fisher's predecessor, Karl Pearson, during WWI. But in 1939 war work was not only a matter of number crunching. Scientific expertise was instead concerned with planning of operations and resources, and loaded with ideological values and Fisher was excluded from it. Thus, he gradually valued his collaboration with the Emergency Transfusion Services as his job of national importance, because his statistical analysis of the donor records could give an immediate feedback on the quality of the serological work done by the transfusion centres. As a tool for increasing the efficiency of the Emergency Transfusion Services, Fisher's survey can be seen as a further contribution of statistics to WWII.

Fisher's engagement with the co-workers of the Galton Laboratory Serum Unit lasted for thirty years, well beyond the period in which he was officially associated with the unit. The serologists that closely collaborated with Fisher, at first George Taylor and later Robert Race, contributed to familiarise their fellow colleagues with the statistical tools – in particular tests of significance – useful for dealing with blood group data. Serologists never managed statistics with the same confidence of their laboratory tools, but their technical competence was a necessary complement to Fisher's statistical expertise in dealing with blood group data and in the analysis of problematic issues. Serology, from this point of view, is but another example that suggests how the development of statistical methods for experimental research required the constant interaction between statisticians and research workers.

---

[587] For the use of transfusion records and blood groups data in Britain after WWII see, for instance, and I. Morgan Watkin (1964), Anonymous (1951).

A concluding remark on the historiography of blood groups. The relevance of blood groups in human genetics during the first half of the twentieth century can be compared to the role of DNA in the second half of the century. However, if the historical accounts of DNA in human genetics have shown a growing interest for the computational and mathematical tools employed in the field and for the strategies of information management they required, in the case of blood group research these aspects have been much less appreciated.[588] The paper tools adopted for managing the donor records or the statistical formulae employed for extracting gene frequencies might seem naive when compared to the digitization of genetic information, but still they are crucial elements for the historical account and deserve more attention than the they have so far received.

[588] See for instance S. de Chadarevian (2002), M. Garcia-Sancho (2012), J. November (2012).

APPENDIX 3.1: Inheritance of the ABO blood groups.

Phen.	Genotype (Von Dungern and Hirszfeld)	Total probability (Von Dungern and Hirszfeld) p = probability of gene A $\underline{p} = (1 - p)$ = probability of gene a q = probability of gene B $\underline{q} = (1 - q)$ = probability of gene b	Genotype (Bernstein)	Total probability (Bernstein) p = probability of gene A q = probability of gene B r = probability of gene R $p + q + r = 1$
O	aabb	$\underline{p}^2 \cdot \underline{q}^2$	RR	r^2
A	AAbb Aabb	$(1 - \underline{p}^2) \cdot \underline{q}^2$	AA AR	$p^2 + 2pr$
B	aaBB aaBb	$\underline{p}^2 \cdot (1 - \underline{q}^2)$	BB BR	$q^2 + 2qr$
AB	AABB AaBB AABb AaBb	$(1 - \underline{p}^2) \cdot (1 - \underline{q}^2)$	AB	$2pq$
	$(\underline{A} + \underline{AB})\ (\underline{B} + \underline{AB}) = \underline{AB}$ $(\underline{A} + \underline{AB}) = [(1 - \underline{p}^2) \cdot \underline{q}^2][(1 - \underline{p}^2) \cdot (1 - \underline{q}^2)] =$ $[(1 - \underline{p}^2)\ (\underline{q}^2 + 1 - \underline{q}^2)] = (1 - \underline{p}^2)$ $(\underline{B} + \underline{AB}) = [\underline{p}^2 \cdot (1 - \underline{q}^2) + (1 - \underline{p}^2) \cdot (1 - \underline{q}^2)] =$ $[(\underline{p}^2 + 1 - \underline{p}^2)\ (1 - \underline{q}^2)] = (1 - \underline{q}^2)$ $(\underline{A} + \underline{AB})\ (\underline{B} + \underline{AB}) = (1 - \underline{p}^2)\ (1 - \underline{q}^2) = \underline{AB}$ **EXPECTATION NOT CONFIRMED BY POPULATION STUDIES**		**$p + q + r = 1$** **GOOD AGREEMENT WITH POPULATION STUDIES**	

(The table is extracted from F. Bernstein, 1966b)

APPENDIX 3.2: ABO blood group survey.

a. Progress of the data collection in the ABO blood group survey

Up to...	Data Collected	Source
October, 1939	Over 58,000 (in the London area)	R. A. Fisher and G. L. Taylor, *BMJ*, October 1939.
April, 1940	106,477 for Southern England 8,716 for Northern England 10,969 for Scotland	R. A. Fisher and G. L. Taylor, *Nature*, April 1940.
August, 1941	About 150,000 (in Britain and Northern Ireland)	G. L. Taylor, R. R. Race and R. A. Fisher, *BMJ*, August 1941.
1941-1942	120,874 in South-Western England	J. A. Fraser Roberts, *Annals of Eugenics*, 1947.
1942	2,550 in North Wales	J. A. Fraser Roberts, *Annals of Eugenics*, 1942.
January, 1943	10,784 in Northern Ireland	E. W. Hart, *Annals of Eugenics*, 1943.
June, 1943	51,135 in Yorkshire 33,042 in the Bristol Area	R. A. Fisher and J. A. Fraser Roberts, *Nature*, June 1943.
1946	190,177 Royal Air Force members 9,044 Cambridge volunteers	A. M. Dobson and E. W. Ikin, *The Journal of Pathology and Bacteriology*, 1946.

b. R. A. Fisher's collaboration with the Emergency Transfusion Services

Area	Town/County	Institution	Representative	Transfusion Records
England	Luton	Blood Depot	(Dr) H. F. Brewer, Medical Officer (Mr) P. L. Oliver, Clerical Assistant	(nearly) 25,000
England	Slough	Blood Depot	(Dr) J. Vaughan, Medical Officer	(nearly) 30,000
England	Sutton	Blood Depot	(Dr) J. O. Oliver, Medical Officer	----------
England	Maidstone	Blood Depot	(Dr) M. Maizel, Medical Officer	10,000
England	London	St. George's Hospital	J. A. Boycott, Blood Supply Officer Sector 7	3,539 (incomplete data)
England	London	West London Hospital	Richard Pearce, Royal Hospital, Richmond, Surrey	1,800
England	London	Royal Northern	Dr. Ellison	1,293

		Hospital, London		
England	Enfield	Emergency Blood Transfusion Service Sector II	(Dr) J. F. Loutit, Blood Supply Officer Sector II, Royal Northern Hospital, Winchmore Hill	3,304
England	London	British Red Cross Blood Transfusion Service	P. L. Oliver, service organiser	6,000-7,000 (not reliable for the sex classification)
England	Cambridge	Galton Laboratory Serum Unit, Department of Pathology, Cambridge	G. L. Taylor, chief serologist Galton Laboratory Serum Unit	nearly 200,000
England	Finchampstead, Berkshire	Voluntary panel	(Dr) E. Billing, physician Finchampstead	116
England	Winchester	National Emergency Blood Transfusion Service at Royal Hampshire County Hospital	(Dr) C. H. Wrigley, Transfusion Officer	------
England	Reading	Royal Berkshire Hospital	J. Mills, Pathologist	-------
South-Western England	Bristol Area [Gloucestershire, Wiltshire, Dorsetshire, Somersetshire, Devonshire, Cornwall]	Army Blood Transfusion Service	L. Whitby, Pathologist	120,874 Data examined by J. A. Fraser Roberts, published in Fraser Roberts (1947).
England	Oxfordshire	Emergency Blood Transfusion Service Southern Region	A. G. Sanders, Regional Transfusion Officer Oxfordshire	--------
England	Wanstead	Emergency Hospital Organisation, Sector I	(Dr) D. B. Byrom, Transfusion Officer	2,951 (Fords employees, Dagenham)
England	Leeds, West Yorkshire	Leeds Blood Transfusion Service	(Dr) G. M. Bonser, Seacroft Emergency Hospital (Dr) W. I. Stanbury, Regional Blood Transfusion Officer	70,000
England	Manchester, Midlands	Regional Blood Transfusion Service, North –Western	John F. Wilkinson, Regional Transfusion Officer	---------

		Region, Royal Infirmary Manchester		
Wales	Caernarvonshire, Denbighshire, Flintshire	Emergency Transfusion Services, Regional Transfusion Centre in the Department of Pathology, University of Liverpool	T. B. Davie, Regional Blood Transfusion Officer	2,550 Data examined by J. A. Fraser Roberts, published in Fraser Roberts (1942).
Scotland	Glasgow	Glasgow Blood Transfusion Service	(Mr) Scott, Town Clerk Depute (Dr) W. Gray, Transfusion Officer (Dr) A. Marshall	2,415
Scotland	Edinburgh	Edinburgh Blood Transfusion Service	J. R. Copland, Organiser of the Edinburgh Blood Transfusion Service (Dr) C. P. Stewart, Clinical Laboratory, Royal Infirmary Edinburgh	5,205
Scotland	Dundee	Blood Transfusion Service Dundee	D. F Cappell, Professor of Pathology, Dundee Royal Infirmary	(nearly) 2,000
Scotland	Aberdeen	Aberdeen Public Health Office	Dr Harry Rae, Medical Officer of Health	------
Northern Ireland	Holywood	Army Blood Transfusion Service, Northern Ireland	Dr John O. Oliver, Major R.A.M.C. Edward W. Hart, Captain R.A.M.C.	10,784 Data examined by E. W. Hart, published in Hart (1943).

*Data collected from the Fisher Papers, Barr Smith Library, The University of Adelaide and from the published papers by R. A. Fisher and co-workers, 1939-1947.

APPENDIX 3.3: Formulae for extracting gene frequencies.

a. F. Bernstein's formulae for extracting gene frequencies

F. Bernstein, 1925	**F. Bernstein, 1930**
$\underline{O}$, $\underline{A}$, $\underline{B}$ = frequencies of the phenotypes O, A and B in the population $r = \sqrt{\underline{O}}$ $q = 1 - \sqrt{(\underline{O} + \underline{A})}$ $p = 1 - \sqrt{(\underline{O} + \underline{B})}$	$\underline{O}$, $\underline{A}$, $\underline{B}$ = frequencies of the phenotypes O, A and B in the population $D = \sqrt{(\underline{O} + \underline{A})} + \sqrt{(\underline{O} + \underline{B})} + \sqrt{\underline{O}} - 1$ $r = (1 + ½ D) (\sqrt{\underline{O}} + ½ D)$ $q = (1 + ½ D) (1 - \sqrt{(\underline{O} + \underline{A})})$ $p = (1 + ½ D) (1 - \sqrt{(\underline{O} + \underline{B})})$ $(1/N)\chi^2 = [(\underline{O} - r^2)^2/r^2] + [(\underline{A} - (p^2 + 2pr))^2/(p^2 + 2pr)]$ + $[(\underline{B} - (q^2 + 2qr))^2/(q^2 + 2qr)] + [(\underline{AB} - 2pq)^2/2pq]$ 1 degree of freedom

b. R. A. Fisher's formulae for extracting gene frequencies in the blood group survey

O = total group *O* in the sample; A = total group *A* in the sample; B = total group *B* in the sample $s = \sqrt{O}$ $t = \sqrt{(O + A)}$ $u = \sqrt{(O + B)}$ $v = t + u - s$	 $p = [(t - s)/v]$ $q = [(u - s)/v]$ $r = s/v$
$w = v^2$ x = expected number of AB bloods y = observed number of AB bloods z = AB (expected – observed) = x – y	$\chi^2 = [tuz^2/wx]$ 1 degree of freedom

APPENDIX 3.4: Data on the ABO blood group system in Britain before WWII.

Investigator(s)	Year	Town/Area	Total No.	Phenotype			
				O	A	B	AB
L. & H. Hirschfeld	1919	All Britain	500	232	217	36	15
J. R. Learmonth	1920	Glasgow	80	38	32	4	6
J. A. Buchanan & E. T. Higley	1921	U.S.A.	218	132	61	21	4
W. Alexander	1921	Dundee	225	98	76	38	13
S. C. Dyke & C. H. Budge	1923	London	194	105	70	17	2
A. H. Tebbutt	1923	Australia (Sydney)	1176	619	433	88	36
A. R. Jones & E. E. Glynn	1926	Liverpool	40	18	12	7	3
G. K. Kirwan-Taylor	1930	London	500	202	234	48	16
M. Penrose & L. S. Penrose	1933	East Counties of England	1000	432	477	64	27
E. A. Shipton	1935, 1936	Australia (Sydney)	220	98	96	20	6
D. Matta	1937	Glasgow	400	195	144	40	21
W. C. Boyd & L. G. Boyd	1937	Wales	192	92	63	31	6
		Ireland	399	220	124	48	7
G. L. Taylor & A. M. Prior	1938	London	422	202	179	35	6
	------	------	**5566**	------	------	------	------

(The table has been compiled from G. L. Taylor and A. M. Prior (1938b) and W. C. Boyd (1939))

Archive materials

<u>Barr Smith Library, The University of Adelaide</u>

1) R. A. Fisher Papers (digitized)

- Correspondence with F. Bernstein, May 1926 – July 1954 [accessed 24th March 2012, http://digital.library.adelaide.edu.au/dspace/handle/2440/67580]
- Letter from R. A. Fisher to K. Pearson, 5th March 1934 [accessed 29th January 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67912/8/1934-03-05.pdf]
- Letter from R. A. Fisher to J. R. Copland, 13th September 1939 [accessed 23rd March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67586/13/1939-09-13.pdf
- Letter from R. A. Fisher to E. J. Russell, 27th September 1939. [accessed 1st February 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67969/8/1939-09-27.pdf
- Letter from R. A. Fisher to D. F. Cappell, 18th October 1939. [accessed 25th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67586/12/1939-10-18.pdf]
- Letter from R. A. Fisher to R. R. Race, 6th November 1939. [accessed 30th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/67/1939-11-06.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 15th November 1939. [accessed 29th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/66/1939-11-15.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 7th December 1939. [accessed 29th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/63/1939-12-07.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 26th January 1940. [accessed 29th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/62/1940-01-26.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 3rd February 1941. [accessed 23rd March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/44/1941-02-03.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 14th November 1941. [accessed 30th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/31/1941-11-14.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 27th April 1942 [accessed 24th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/25/1942-04-27.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 12th March 1943. [accessed 30th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/12/1943-03-12.pdf]
- Letter from R. A. Fisher to G. L. Taylor, 27th March 1943. [accessed 23rd March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68047/9/1943-03-27.pdf]
- Letter from R. A. Fisher to W. C. Boyd, 8th May 1953. [accessed 29th March 2012, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67592/1/1953-05-08.pdf]

2) R. A. Fisher Papers

a) Blood Group Survey Correspondence

- Letter from M. Maizel to R. A. Fisher, 20th October 1939.
- Letter from R. A. Fisher to D. F. Cappell, 23rd October 1939.
- Letter from D. F. Cappell to R. A. Fisher, 25th October 1939.
- Letter from J. Vaughan to R. A. Fisher, 10th November 1939.
- Letter from R. A. Fisher to J. Vaughan, 5th December 1939.
- Letter from R. A. Fisher to P. L. Oliver, 14th December 1939.

- Letter from R. A. Fisher to A. C. Scott, 15th January 1940.
- Letter from J. Vaughan to R. A. Fisher, 24th July 1940.
- Letter from R. A. Fisher to W. I. Stanbury, undated.
- Data sheets for the Slough Depot donors, undated.

b) Correspondence with G. L. Taylor

- Letter from G. L. Taylor to R. A. Fisher, 28th August 1937.
- Letter from G. L. Taylor to R. A. Fisher, 5th October 1939.
- Letter from G. L. Taylor to R. A. Fisher, 2nd November 1939.
- Letter from R. A. Fisher to G. L. Taylor, 23rd November 1939.
- Letter from G. L. Taylor to R. A. Fisher, 12th December 1939.
- Letter from R. A. Fisher to G. L. Taylor, 13th November 1939.
- Letter from R. A. Fisher to G. L. Taylor, 24th January 1940.
- Letter from G. L. Taylor to R. A. Fisher, 19th February 1940.
- Letter from G. L. Taylor to R. A. Fisher, 24th February 1940.
- Letter from G. L. Taylor to R. A. Fisher, 4th March 1940.
- Letter from G. L. Taylor to R. A. Fisher, 13th March 1940.
- Letter from R. A. Fisher to G. L. Taylor, 21st March 1940.
- Letter from G. L. Taylor to R. A. Fisher, 5th April 1940.
- Letter from G. L. Taylor to R. A. Fisher, 8th November 1940.
- Letter from R. A. Fisher to G. L. Taylor, 11th November 1940.
- Letter from G. L. Taylor to R. A. Fisher, 26th March 1943.

<u>The National Archives of the UK</u>

1) Medical Research Committee and Medical Research Council: Files, Ref. FD/1.

- Letter from G. L. Taylor to (Sir) E. Mellanby, 28th September 1938, FD/1/5845.
- Letter from A. Landsborough Thomson to G. L. Taylor, 25th July 1939, FD/1/5845.
- Letter from R. A. Fisher to A. Landsborough Thomson, 11th July 1941, FD/1/5845.
- Letter from R. A. Fisher to A. Landsborough Thomson, 23rd July 1941, FD/1/5845.
- Letter from the MRC to the Medical Officer of Health [unknown area], 24th August 1939, FD/1/5751.
- Letter from the MRC to H. F. Brewer, 24th June 1939, FD/1/5854.
- Letter from H. F. Brewer to A. N. Drury, 6th November 1940, FD/1/5923.
- Luton Depot donor cards, 1940, FD/1/5923.

2) Medical Research Council: Blood Group Unit: Reports and Papers, Ref. FD/8.

Donor list, Regional Blood Transfusion Service, Sheffield, 30th December 1942, FD/8/6.

3) Agriculture, Fisheries and Food Departments: National Agricultural Advisory Service and Predecessors, Ref. MAF/33.

Letter from D. J. Finney to W. R. Black, 7th July 1945, MAF/33/333.

<u>Rockefeller Archive Center</u>

- Letter from K. Pearson to C. M. Gayley, June 1925, Collection RF, Record Group 1.1, Series 401 England, Box 16, Folder 218, Galton Laboratory 1928-1930.

- Letter from A. Mawer to D. P. O'Brien, 26th October 1934, Collection RF, Record Group 1.1, Series 401 England, Box 16, Folder 219, Galton Laboratory 1932-1934.
- Prospectus of the Eugenics Department, 1934, Collection RF, Record Group 1.1, Series 401 England, Box 16, Folder 219, Galton Laboratory 1932-1934.
- Letter from R. A. Fisher to D. P. O'Brien, 18th July 1934, Collection RF, Record Group 1.1. Series 401 England, Box 16, Folder 219, Galton Laboratory 1932-1934.
- Cost of the New Scheme in Serological Genetics (written by R. A. Fisher), 1934, Collection RF, Record Group 1.1. Series 401 England, Box 16, Folder 219.
- New Scheme of Research in Serological Genetics (written by R. A. Fisher), 1934, Collection RF, Record Group 1.1. Series 401 England, Box 16, Folder 220, Galton Laboratory 1935.
- Letter from H. R. Dean to R. A. Fisher, 4th October 1934, Collection RF, Record Group 6.1 Field Office Paris Series 1.1 MS England, Box 16, Folder 17, Fisher Serological Lab.
- Memorandum from D. P. O'Brien to A. Gregg, 1st March 1935, Collection RF, Record Group 1.1. Series 401 England, Box 16, Folder 220, Galton Laboratory 1935.
- Memorandum from W. E. Tisdale to W. Weaver, 5th March 1935, Collection RF, Record Group 1.1. Series 401 England, Box 16, Folder 220, Galton Laboratory 1935.
- Memorandum from D. P. O' Brien to A. Gregg, 6th March 1935, Collection RF, Record Group 1.1. Series 401 England, Box 16, Folder 220, Galton Laboratory 1935.
- Warren Weaver Diary, 16th May 1935, Collection RF, Record Group 1.1, Series 401 England, Box 16, Folder 220, Galton Laboratory 1935.
- Letter from R. A. Fisher to D. P. O'Brien, 18th May 1936, Collection RF, Record Group 1.1, Series 401 England, Box 16, Folder 221.
- Galton Laboratory Serum Unit Financial Prospect, 7th December 1938, Collection RF, Record Group 1.1, Series 401 England, Box 16, Folder 221, Galton Laboratory 1936-1947.

<u>Rothamsted Research, Library and Archive</u>

1) Staff Council Minutes, 11th July 1927, Vol. 1, STA 2.1.
2) Correspondence with W. L. Stevens and F. C. R. Jourdain on British Birds, Ref. STATS 6.4.
Letter from R. A. Fisher to Mr F. C. R. Jourdain, 21st February 1928, STATS 6.4.

<u>University College London</u>

1) Special Collections
Letter from J. B. S. Haldane to J. D. Bernal, 10th January 1944, Haldane Papers, Correspondence Box 4.
2) Record Office
- Appeal Fund, 1907-1920, ARC/2007/342, AR 179, Folder 4.
- Galton Laboratory Ledger, 1934-1939, ARC/2007/344, AR 181, Folder 4.

<u>Wellcome Library</u>

1) Blood Transfusion Service (1943-1954), Ref. GC/107.
- Letter from A. N. Drury to P. N. Panton, 23rd September 1943, GC/107/1.
- Blood Transfusion Service: correspondence, circular, forms, reports, statistical returns, notes on research, GC/107/2.
2) Dame Janet Vaughan Papers, Ref. GC/186.

- Minutes of the Meeting of the Sub-Committee on Blood Transfusion of the MRC, Emergency Pathological Services, 17th April 1939, GC/186/1.

3) Medical Research Council Blood Group Unit (1935-1995), Ref. SA/BGU.

- Correspondence between G. L. Taylor and V. Friedenreich, 1936-1937, SA/BGU/A1/2.
- Letter from R. A. Fisher to G. L. Taylor, 11th November 1939, SA/BGU/F1/1/1.
- Letter from A. E. Mourant to R. A. Fisher, 8th January 1946, SA/BGU/E7/1.
- Letter from R. A. Fisher to A. E. Mourant, 9th January 1946, SA/BGU/E7/1.
- Letter from A. E. Mourant to R. A. Fisher, 11th February 1946, SA/BGU/E7/1.
- Letter from R. R. Race to R. A. Fisher, 17th March 1947, SA/BGU/E7/1.
- Letter from R. A. Fisher to R. R. Race, 18th March 1947, SA/BGU/F1/1/1.
- Letter from R. R. Race to R. A. Fisher, 1st July 1947, SA/BGU/F4/3/1.
- Letter from R. A. Fisher to R. R. Race, 1st February 1949, SA/BGU/E7/1.

4) Percy Lane Oliver Papers, Uncatalogued.

- "Report on the Work of the Blood Transfusion Service", written by P. L. Oliver, about 1937, Box 8.
- "Red Cross Blood Transfusion Service Telephone Duties", anonymous report, undated, Box 8.
- Minutes of the Meeting of the Committee of the London Blood Transfusion Service, 9th April 1930, Box 2.
- Minutes of the Meeting of the British Red Cross Society Blood Transfusion Committee, 27th June 1932, Box 2.
- Minutes of the Meeting of the British Red Cross Society Blood Transfusion Committee, 13th April 1939, Box 2.

5) Film *Blood Transfusion Service* (1941), Part 3.

A Ministry of Information film. Made for the Ministry of Health with the cooperation of the Medical Research Council and the Blood Transfusion Units of the Fighting Services. Directed and edited by H.M. Nieter. A Paul Rotha production. http://catalogue.wellcome.ac.uk/record=b1675865~S3.

Chapter 4

THE COMPUTERIZATION OF THE ROTHAMSTED STATISTICS DEPARTMENT

4.1 *Introduction*

In the second half of the twentieth century statistics applied to agriculture and biology has increasingly become a computational science: algorithms for implementing statistical methods on computers have been designed by statisticians and computer scientists, and journals and associations dedicated to statistical computing have appeared. New hybrid disciplines, such as bioinformatics and systems biology, which mix statistics and computing, have also been born. The diffusion of off-the-shelf statistical software – from the basic applications made possible by spreadsheets to the more sophisticated software packages – has made statistics easily available to an increasing number of users, albeit at the expense of the user's awareness of the statistical methods underpinning the software applications, and the personal computer has become a key tool in the teaching of statistics.

To account for this transformation, since the 1960s it has been suggested that computers have promoted a revolutionary change in statistics.[589] However, the thesis seems too simplistic to be true, because it attributes agency to the technology alone and disregards the complex interactions between tools and practices in science and the constraints represented by institutional and disciplinary boundaries in the "domestication" of digital computers in statistics.[590] Historical studies of computerization in actuarial statistics challenge as well this thesis. JoAnne Yates, in fact, has pointed out that in the adoption of digital computers in life insurance, business needs and technology mutually shaped each other and that the use of pre-computer tools, such as punched-card equipment, influenced the following adoption of mainframes.[591]

A detailed account of computerization in present day statistics for agriculture and biology is beyond the scope of my thesis, but in the present chapter I will use the acquisition of the first mainframe in the Rothamsted statistics department as a case study to examine the role of computers in statistics during the 1950s and 1960s and to point out the problems raised by the introduction of the new technology. Computerization at Rothamsted started in 1954, when a prototype mainframe, named Elliott 401, was leased to the station by the National Research

[589] See for instance F. Yates (1966), L. Råde (1990).
[590] On the domestication of technologies in science see N. Rasmussen (1996).
[591] J. Yates (2008).

Development Corporation (NRDC), a government body which promoted the industrial development of British inventions, and paid for by the Agricultural Research Council (ARC), the main funding institution for agricultural research in Britain.

During WWII the department was engaged in operational research along with its head, Frank Yates, and these wartime experiences prompted the expansion of the Rothamsted statistics department in post-war Britain and the acquisition of up-to-date computing tools. Scientific reasons, however, cannot account by themselves for this case study of early computerization. In the acquisition of the Elliott 401 the alliances established in operational research by Frank Yates were, in fact, decisive. The key figures in the ARC and NRDC who advised favourably on the acquisition of the mainframe, were all former acquaintances of Yates in operational research and their support was necessary to gain the digital computer, a piece of technology still costly and rare in the early 1950s.

In the four sections which constitute the chapter I am going to investigate the intertwining of scientific goals and wartime experiences combined in the acquisition of a digital computer for the Rothamsted statistics department and re-assess the role of the new technology in the working life of the department. I will begin accounting the involvement of the Rothamsted statisticians and, in particular their chief, Frank Yates, in operational research during WWII and the expansion of the department after 1945. I will then examine the claims made in support of computerization at Rothamsted and how the engagement of Frank Yates in operational research during WWII was instrumental for the acquisition of the Elliott 401. I will also discuss in detail the agreement signed by Rothamsted Experimental Station for renting the mainframe.

In the third section, after a brief introduction on the post-war developments of the British computer industry and the making of the Elliott 401, I am going to discuss the role acquired by the mainframe in Frank Yates' department and how the new technology interfaced with the computing equipment already available. The Elliott 401 successfully settled into the research activity of the Rothamsted statistics department side by side by the desk calculators and the punched-card equipment already in use, and even though the mainframe had been acquired for research purposes, it became a key component as well of the routine work required by the analysis of agricultural experiments and surveys, which were the main activities of the Rothamsted statistics department.

The last section of the chapter will be devoted to some of the scientific applications for which the Elliott 401 was employed. I will describe the programs written for the analysis of replicated experiments and the general survey program prepared for the mainframe. The latter application will be an opportunity to discuss the relevance of autocodes in the preparation of statistical software and how these early attempts at generalization paved the way for statistical packages like GenStat. The last scientific application of the Elliott 401 I will consider is the program for linkage analysis written for the examination of the pedigree data collected by Sylvia Lawler and James Renwick, two researchers at the Galton Laboratory. This is just an example of the advisory work done by the Rothamsted staff for exploring the use of mainframes in the solution of statistical problems in genetics, an area in which digital computers have gained throughout the decades a key role in data analysis. Further information on the scientific activity of the Rothamsted statistics department can be found in the appendices (Appendix 4.2 and 4.3) of the chapter, in which I have reported the first hand accounts of a statistician, John C. Gower, and a member of the assistant staff, Vera Wiltshire, who worked in the department during the 1950s and 1960s.

Frank Yates has probably been the first to claim a computer revolution in statistics, but in conclusion I will argue that the computerization of his own department was all but a revolutionary event. The acquisition of the Elliott 401 was deeply rooted in the Rothamsted tradition that conceived computing technologies as research tools in statistics and needed the alliances built by Frank Yates among his former colleagues in operational research. Moreover, computerization was effectively realized only when programming skills had been acquired and a library of statistical routines built. This process did not happen overnight, but required several years and a new generation of statisticians able to engage with mainframes and to rephrase statistics in the language of computer algorithms.

4.2 *The Rothamsted statistics department after WWII*

4.2.1 *Warfare, operational research and the post-war expansion of statistics in agriculture and biology*

Britain's involvement in WWII has been accounted as a deployment of military technologies, science and invention, experts and technocrats.[592] An outcome of the wartime mobilisation of British science was operational research, that is the use of statistical and probabilistic methods for approaching complex problems.[593] The new discipline emerged in connection with the development of radar technologies and military strategies and its practitioners – a mixed crowd of experts in fields such as physics and statistics, genetics and anatomy – were enrolled in the warfare as scientific advisors.[594]

[592] D. Edgerton (2011).
[593] E. P. Rau (2001), E. P. Rau (2005); J. Agar (2003), pp. 248-251.
[594] The term 'operational research' was already employed before WWII to describe the joint effort of military officers

Due to their different background, operational researchers resulted an inhomogeneous expert group, having in common only the use of statistical methods for tackling complex problems.[595] They chose to associate their discipline with statistics, as statistics offered means "far more accurate and rapid than mere common sense for sizing-up any situation, however complicated, and for measuring the actual effect of any steps taken to deal with it".[596]

The applications of operational research were not limited to military problems and operational research did not disappear at the conclusion of the war, but significantly contributed to the diffusion of statistical and system thinking in the second half of the twentieth century.[597] In post-war Britain the expansion of statistics promoted by operational research met with the global increase of funding available for scientific research.[598] Statistics benefited from this climate with a general expansion of the discipline in terms of people engaged, publications and field of applications for statistical methods.

The Rothamsted statistics department was at the forefront of this development. Since the 1920s the department had been a well-established centre for the design and analysis of agricultural and biological experiments, mainly for the experimental station and its associated centres, but also for outside institutions interested in a more scientific approach to agricultural and biological research. However the staff had always been scarce – two or three statisticians, three or four human computers, a few temporary workers – and its funding dependant from the overall success of the experimental station in gaining public grants (**chapter 1**).

The development of operational research during the warfare and the involvement of the Rothamsted statisticians in it – an involvement prompted by the personal and professional liaisons of Frank Yates, the head of the department – brought new opportunities for research and

---

and scientists engaged in the set up of radar technologies (J. Agar (2003), p. 248; E. P. Rau (2001), pp. 219-225). On the mobilisation of British scientists during WWII and their involvement in operational research see M. Fortun and S. S. Schweber (1993), S. de Chadarevian (2002), pp. 20-28, and E. P. Rau (2001), pp. 225-242).

595 For the use of statistics as a unifying element in operational research see J. Agar (2003), p. 250. Agar considers the mere reliance on statistical techniques a shaky foundation for operational research and suggests that the new speciality, due to its vagueness, but not only, struggled to find recognition in post-war Britain (J. Agar (2003), p. 250; J. Agar and B. Balmer (1998)). E. P. Rau, instead, remarks that in the case of British operational research such inhomogeneity was vital as "[a]uthority was conferred and a new professional identity formed through a working relationship between scientists and military officers within a framework of patronage" (E. P. Rau (2001), p. 216). He agrees, however, with Agar that "OR's [operational research] early entry in civilian government [...] was rather checkered, and its adoption in industry was gradual", with a consequent shift of the discipline from Britain to the US, where the Cold War offered a more promising context (E. P. Rau (2001), p. 243).

596 *Science in War* (1940), p. 9.

597 E. P. Rau (2001). For instance, A. Kaijser and J. Tiberg offer an overview of the post-war intertwining between operational research and systems thinking in their home country, Sweden (A. Kaijser and J. Tiberg (2000)).

598 De Chadarevian offers an interesting account of the expansion of British science after WWII giving facts and figures (S. de Chadarevian (2002), pp. 33-43). The budget of the ARC, for instance, increased from £0.3 million in 1945-46 to £5.6 million in 1960-61 (S. de Chadarevian (2002), p. 36).

fresh funding during WWII and afterwards.[599] The staff in the Rothamsted statistics department was only marginally involved in military operations, but actively engaged in several statistical investigations useful for assessing availability and rational distribution of agricultural commodities.[600] The Rothamsted statisticians were called as consultants in the set-up of several wartime surveys. "The first call I had – Frank Yates remembers – was from the Forestry Commission within a week of declaring war" and his task was to estimate the availability of timber for the whole country.[601] In 1941-1942 the Rothamsted statistics department took part also in the qualitative survey of English and Welsh farms, sponsored by the Ministry of Agriculture.[602]

"A piece of operational research in the agricultural field", started at Rothamsted during WWII and successfully carried on in the following decades, was concerned with the fertilizer policy.[603] In 1940 the Rothamsted statistics department undertook, in collaboration with the chemistry department of the station, a collection and meta-analysis of the fertilizers trials conducted in Britain since the beginning of the twentieth century.[604] The data analysis prompted a sampling survey on fertilizer practice and by 1945 about forty counties in England and Wales had been partially or fully surveyed giving a better understanding of the real necessities of fertilizers for agriculture.[605]

The wartime experience changed in depth the Rothamsted statistics department: its work expanded well beyond the experimental station and its staff more than doubled if compared to the pre-war levels (Appendix 4.1). At the end of the conflict it became thus urgent a reorganisation in terms of funding and equipment. On the basis of a memorandum submitted by Frank Yates in April 1945, the Agricultural Research Council in association with the Agricultural Improvement Council (AIC) and the Scottish Agricultural Advisory Council discussed a comprehensive plan for the development of statistical facilities in agriculture.[606] Yates' memorandum "drew attention to the urgent need for better advice and assistance, and for increased facilities for training, in the use of statistical methods in agricultural research" and

---

599 Frank Yates and his fellow statisticians did not refrain from adding the term operational research in the reviews of the department activity during the 1950s (RES (1952), p. 71; RES (1953), p. 130; RES (1954), p. 133).
600 RES (1946), pp. 113-115.
601 Sir David Cox interviews Frank Yates, 24th April 1986, RSS. See also RES (1946), p. 106 and F. Yates (1943), pp. 4-5.
602 RES (1946), pp. 116-117.
603 F. Yates (1950), p. 221; F. Yates and D. A. Boyd (1965).
604 RES (1946), pp. 262-265. E. M. Crowther and F. Yates (1941).
605 Sir David Cox interviews Frank Yates, 24th April 1986, RSS; F. Yates and D. A. Boyd (1965), pp. 204.
606 The ARC was the main funding body for agricultural research in England and Wales since the early 1930s (for an history of the council see G. W. Cooke, 1931); the AIC for England and Wales was created in 1941 to facilitate the application of best practices suggested by scientific research to agricultural science and to point out agricultural problems worth of scientific investigation. The Scottish Agricultural Advisory Council was the corresponding organisation for agricultural research in Scotland.

suggested "to increase the number of statistical units at research centres and, in addition, to set up one or more larger units at appropriate centres".[607]

The first scheme drafted in 1946 envisioned three main centres of agricultural statistics to be implemented at Rothamsted Experimental Station (possibly transforming the statistics department in an independent unit), Cambridge University (specialised in veterinary research, genetics and agricultural meteorology and attached to the department of genetics headed by Ronald Fisher) and Edinburgh University (under the statistician Alexander C. Aitken working in the university department of mathematics).[608]

According to Yates' proposal the tasks of the Rothamsted statistics department should include research in statistical methods for agriculture and biology; assistance to Rothamsted and other agricultural institutions in the planning and analysis of field experiments; promotion of sampling surveys in agricultural research; statistical consultancy for the British colonies; training of postgraduate students in statistics for agricultural research; computing services offered to research institutions that could rely just on basic apparatus. Due to the increased workload, provisions were made for hiring new staff and Yates aimed at a department of about sixty people.[609]

The actual development of the three centres had a much slower progress than expected in the initial plans. The Rothamsted statistics department remained part of the experimental station, Ronald Fisher never supervised a unit in Cambridge sponsored by the ARC, and a statistical centre in Scotland, headed by a former co-worker of both Ronald Fisher and Frank Yates, David Finney, was founded only in 1954 at Aberdeen University.[610] Nevertheless, in 1947 the Rothamsted statistics department became a general statistical service for agriculture and biology and gave support to several research workers in the country and abroad. In the early 1960s, the Elliott 401 was still unrivalled for statistical computations by the computers working at London and Cambridge University.[611]

---

[607] Joint Committee of the AIC (England and Wales), the ARC and the Scottish Agricultural Advisory Council, J.C./P.65, TNA, MAF 117/88.

[608] Joint Committee of the AIC (England and Wales), the ARC and the Scottish Agricultural Advisory Council, J.C./P.65, TNA, MAF/117/88. A short biography of A. C. Aitken is in chapter 2, Appendix 2.1.

[609] Joint Committee of the AIC (England and Wales), the ARC and the Scottish Agricultural Advisory Council, J.C./P.65, TNA, MAF 117/88; Statistical department accommodation, summary of memorandum compiled by F. Yates for W. Ogg, 16th May 1950, RR Library and Archive, SITE 1.5.2.

[610] See G. W. Cooke (1981), pp. 53-54 for the statistics units sponsored by the ARC. D. J. Finney worked at the Galton Laboratory under R. A. Fisher in the late 1930s and from 1939 to 1945 was a staff member of the Rothamsted statistics department.

[611] F. Yates (1962), p. 275.

4.2.2 ***Frank Yates as statistician, computer and operational researcher***

Born in Didsbury, Manchester, in 1902, Frank Yates studied mathematics at St. John's College, Cambridge, where he graduated with first class honours in 1924. Yates then worked as mathematical schoolmaster in a public school, Malverne College, before being appointed to the Gold Coast (Ghana) Geodetic Survey in 1927. In 1931 he started his career at Rothamsted as assistant of Ronald Fisher (**chapter 1**) and, only two years later, he became head of the department, when Fisher left the experimental station.[612]

Before WWII the influence of Fisher on his former department and on his new – and inexperienced – head was decisive. Yates concentrated mainly on experimental design and the theory and practice of sampling, two strands of research started at Rothamsted in Fisher's time.[613] Like Fisher, Yates believed that statistics had to "satisfy the needs of the biologist and agriculturist rather than pursuing theory for theory's sake" and for this reason he gained the election to the Royal Society in 1948.[614] A dedicated computer since his experience as surveyor, Yates collaborated with Fisher authoring and constantly revising the *Statistical Tables for Biological, Agricultural and Medical Research* (**chapter 2**).[615]

The scientific collaboration between Fisher and Yates during the 1930s was further strengthened by their common interest for genetics, Fisher's second career (**chapter 3**). Yates "learnt quite a lot of genetics before the war with Fisher", who introduced him to several British geneticists such as Cyril Darlington and Kenneth Mather, promoted Yates' entrance in the Genetical Society and involved him in experiments on animal and plant genetics.[616]

Yates "would have kept up with genetics had it not been for the war" that fully engaged him in operational research, alongside the anatomist and civil servant Solly Zuckerman.[617] Yates entered in the *Tots and Quots*, Zuckerman's dining club, when it was revived at the outbreak of the

---

612 For a scientific biography of Frank Yates see his obituary as fellow of the Royal Society (D. J. Finney (1995)). Other biographical recollections of Frank Yates' life written by his former colleagues are M. J. R. Healy (1995) and G. Dyke (1995). The Lawes Agricultural Trust, the governing body of Rothamsted Experimental Station, appointed Yates as head of the statistics department in November 1933 (Rothamsted Staff Council Minutes, 9th November 1933, Volume 2, RR Library and Archive, STA 2.1).

613 The scientific correspondence of Fisher and Yates (RR Library and Archive STATS 7.5 and the Fisher Papers held at the Barr Smith Library, the University of Adelaide) offers information on their collaboration. At Yates' retirement from Rothamsted, his former co-workers W. G. Cochran, D. J. Finney and M. J. R. Healy sponsored the reprinting in book form of Yates' main contributions to experimental design (F. Yates (1970)).

614 Letter from F. Yates to E. J. Russell, 19th March 1948, E. J. Russell Correspondence, MERL, Reading, FR HERT 11/1/1.

615 R. A. Fisher and F. Yates (1963).

616 Sir David Cox interviews Frank Yates, 24th April 1986, Royal Statistical Society. Yates acted as secretary of the section on "Statistical Genetics" held at the 1939 Congress of Genetics in Edinburgh; recorder of the session was R. A. Fisher. C. Darlington was director of the John Innes Horticultural Institution and later professor of botany at Oxford University; K. Mather was assistant geneticist of R. A. Fisher at the Galton Laboratory in the 1930s, then moved to the John Innes Institute and later to the chair of genetics at Birmingham University.

617 Sir David Cox interviews Frank Yates, 24th April 1986, RSS.

conflict.[618] Several members of the club were scientists of left-wing and liberal sympathies directly engaged in operational research and the club promoted the mobilization of scientists for the warfare.[619] The pamphlet *Science in War* (1940) emerged from the dining club discussions, and Yates contributed to the book with the sections related to agriculture.[620]

Along with Zuckerman, Yates was involved in the Bombing Survey Unit, the Bombing Analysis Unit, and the Allied Expeditionary Air Force.[621] He took part in investigations on casualties during air raids, the efficacy of the Anderson shelters to protect civilians, the determination of effective bombing strategies. He spent several months abroad in Sicily in 1943 and in France, Belgium and Germany in 1944-1945.[622] Besides his collaboration to military operations, Yates contributed to the warfare with the agricultural surveys entrusted to his Rothamsted department.

After WWII Yates' career path drifted away from the one of Ronald Fisher and genetics was not to be resumed, while research on sampling techniques and the planning and analysis of censuses and surveys became a relevant component of Yates' work, adding to his traditional involvement in the design and analysis of agricultural and biological experiments. Yates was a member of the United Nations sub-commission on statistical sampling (1947-1952) and his book *Sampling Methods for Censuses and Surveys* was written, under the auspices of the commission, "to assist in the execution of the projected 1950 World Census of Agriculture, and the 1950 World Census of Population".[623]

With the expansion of his department after WWII both the statistics and computing work increased and Frank Yates, constantly interested in efficient computing solutions, requested funding not only for hiring more staff, but also for the acquisition of up-to-date tools, at first punched-card equipment and later electronic computers. To Yates the computing equipment was not just an artefact, but an integral part of the research activity in the department

---

618 S. Zuckerman (1988), p. 112. For an history of the *Tots and Quots* see S. Zuckerman (1988), pp. 109-112; pp. 393-404.

619 The campaign for the mobilization of scientists in the warfare was mainly driven by left-wing intellectuals (D. Edgerton, 2011, pp. 140-142). It is likely that even Frank Yates' political views inclined to the left. His second wife, married in 1939, Pauline Shoubersky, was Russian by birth and the head of the science division at the London headquarters of the Society for Cultural Relations with the USSR. From 1940 to 1949, Pauline and Frank Yates were controlled by the secret service as suspected members of the Communist party (The Security services personal files, Frank Yates alias Francis Gray, TNA, KV 2/3082).

620 Letter from B. Schafer (on behalf of C. Darlington) to S. Zuckerman, 21st September 1940, S. Zuckerman Papers, UEA, SZ/TQ/2/2/6. Further information about *Science in War* in S. de Chadarevian (2002), pp. 23-24.

621 Letter from S. Zuckerman to R. S. Capon, 24th June 1945, S. Zuckerman Papers, UEA, SZ/BBSU/1/22. On Zuckerman's wartime involvement in operational research see also E. P. Rau (2001), pp. 238-242.

622 RES (1946), pp. 106-107; S. Zuckerman (1988), pp. 207-210, pp. 242-243, pp. 251-253. Further information on the collaboration between Zuckerman and Yates is offered by the archive materials held in the Zuckerman papers at the University of East Anglia.

623 F. Yates (1953), p. *v*.

our Hollerith installation should be regarded as a piece of research apparatus which, in addition to its routine uses for the particular statistical analyses that we are carrying out, will enable new methods of analysis to be developed. Our policy should therefore be to keep in touch with modern developments of computing machinery, especially modern electronic and relay computers so as to be able to adapt them to biological and agricultural needs as need arises.[624]

Yates' vision gained for Rothamsted a pioneering role in the acquisition of digital computers for scientific research. With the arrival of the Elliott 401 in the department in 1954, Rothamsted was among the first five government institutions provided of a stored program electronic computer.[625] As a pioneer user of digital computers, when the British Computer Society was established in 1957, Yates entered in its council and in 1960-61 served as its president.[626]

Throughout Frank Yates' career at Rothamsted the statistics department was in charge of both statistics and computing and only at Yates' retirement in 1968 two distinct units were created, one concerned with the preparation of software suitable for statistical analysis and the other responsible for the centralised computing services sponsored by the ARC.[627] Yates conserved an office in the Rothamsted computing department and remained associated with Rothamsted Experimental Station for over sixty years.[628]

Fig. 4.1 Frank Yates (undated).
Credits: Copyright Rothamsted Research Ltd.

4.2.3 *Staff and tasks of the Rothamsted statistics department after WWII*

Since 1945 the main addition to the traditional work of the Rothamsted statistics department was represented by the development of operational research for agriculture, in particular the

[624] Review of Hollerith equipment at Rothamsted Experimental Station, F. Yates, 27th December 1950, TNA, MAF 117/88. For an history of punched-card equipment see L. Heide (2009). M. Campbell-Kelly (1989), chapters 1-7, gives a detailed description of Hollerith equipment and its manufacturers in Britain.
[625] J. Agar (2003), pp. 299-300; p. 504.
[626] RES (1958), p. 182.
[627] RES (1970), pp. 264-280.
[628] D. J. Finney (1995), pp. 571-572.

sampling surveys, which became a speciality of Frank Yates and his co-workers. The department was consulted in relation to surveys by British institutions – agricultural research stations, the Ministry of Agriculture and Fisheries, the ARC, the AIC – and international bodies, such as the United Nations and their Food and Agriculture Organisation.[629]

The development of the survey work did not hinder the advisory services in the design and analysis of agricultural and biological experiments traditionally offered by the Rothamsted statistics department. In particular, after WWII Frank Yates' department devoted his forces also to the planning and analysis of animal experiments, previously just a small component of its work.[630] The growth in the design and analysis of agricultural experiments was mainly influenced by the statistical support given to the National Agricultural Advisory Service (NAAS), instituted in 1946 "to give, free of charge, technical advice and instruction, whether practical or scientific, on agricultural matters".[631] Within its consultancy duties towards local farmers and breeders, the NAAS was also expected to carry out experiments in crop and fruit production and animal husbandry and to support private experimentation.[632] The Rothamsted statisticians regularly attended the meetings of the NAAS provincial committees advising "on points of design, layout and on the analysis of results" and promoted the use of statistics in agricultural experimentation also lecturing and preparing pamphlets for the NAAS.[633]

In December 1947 the Rothamsted statistics department became a general research statistical service, as agreed by the Ministry of Agriculture and Fisheries, the ARC and the AIC. It was in this capacity that Yates' department gave statistical advice to several British institutions in the design and analysis of agricultural and biological experiments, the planning of field experiments programmes, the analysis of scientific surveys and the critical examination of large bodies of experimental data.[634] Statistical advice on experimental design and analysis and on sampling surveys was also given to the colonies of the British Empire and in 1950 a post of colonial statistician was created in Yates' department for the related advisory work.[635]

With the acquisition of the Elliott 401 in 1954 a new relevant activity in the Rothamsted statistics department became programming. Off-the-shelf statistical software did not exist and the

[629] A few examples of the work done of behalf of the mentioned institutions are: survey of maincrop potatoes for the AIC (RES (1949), p. 47); survey of marginal land in England and Wales for the ARC (RES (1950), p. 55); the schemes for sampling roads vehicles prepared by F. Yates for the United Nations statistical office (RES (1952), p. 72); survey of diseases of dairy cattle for the Veterinary Laboratory, Weybridge (RES (1954), p. 134); sampling analysis on farms for the Ministry of Agriculture and Fisheries (RES (1963), p. 178).
[630] RES (1952), p. 70. Roughly the number of experiments examined increased four times before and after the war, comparing the figures for 1934 (115) and 1951 (437).
[631] R. Rae (1955), p. 260. The quotation is taken from the NAAS charter.
[632] R. Rae (1955), p. 265.
[633] Quotation from R. Rae (1955), p. 265. See also RES (1952), p. 70.
[634] RES (1948b), pp. 44-45.
[635] RES (1951), p. 65.

library of routines for the analysis of experiments and surveys with the Elliott 401 had to be built from scratch, as no program had been written for the computer other than input and output routines, when it arrived at the experimental station.[636] During the 1950s only a few of the Rothamsted statisticians were directly involved in the programming work, as new skills, beyond statistical knowledge, were needed for the task. Before the set up of the mainframe, Michael Healy and Douglas H. Rees attended a course on programming methods held at the Mathematical Laboratory Cambridge, one of the pioneering institution interested in digital computers in Britain.[637] In the 1950s the chief programmers, apart from Healy and Rees, were Frank Yates, Steve Lipton, John Gower, Howard Simpson, Brian Leech, Neil Gilbert – a statistician hired by the John Innes Institute, but seconded to Rothamsted for a few years – joined in the 1960s by Gavin Ross and Alan Martin.[638] All the programmers at Rothamsted were statisticians, because Frank Yates was convinced that only a combined expertise in statistics and computing could lead to adequate programs for statistical analysis.[639]

The programs written in Frank Yates' department were published in both computing and statistics journals, like the journals of the Royal Statistical Society and the *Computer Journal*, the official publication of the British Computer Society. In this way they could reach two distinct publics: on the one hand the statisticians who did not read computer journals and on the other the people engaged in the development of digital computers, but not strictly interested in statistics. Frank Yates and his colleagues had to negotiate their role within the raising discipline of computer science and learn how to rephrase in algorithmic language the statistical ideas traditionally presented through the tools of mathematical analysis.[640]

With the development of programming work and the computerization of routine tasks, such as the analysis of experiments and surveys, the role of the assistant staff in the Rothamsted statistics department changed as well. As in the 1920s and 1930s (**chapter 1**), again after WWII the human computers were almost all women, without university degrees and engaged in routine work on the desk calculators. With the increased sophistication of the computing tools in the

636 The electronic computer at Rothamsted, Progress Report, 1954-1956, NAHC, Manchester, NAHC/NRD C10/8.
637 RES (1954), p. 132. The head of the Mathematical Laboratory Cambridge, Maurice Wilkes, co-authored one of the first programming manuals available in the 1950s (M. V. Wilkes et al. (1951)).
638 G. Ross, personal memories.
639 F. Yates (1966), p. 250. However, not all the statisticians at Rothamsted were programmers (F. Yates (1962), p 278). Apart from F. Yates, the senior scientific staff in his department showed little interest for the digital computer (Interview with J. C. Gower, September 2010).
640 The dichotomy in the public of computer programs for statistics remained for decades and in the 1970s the task force on computing of the American Statistical Society still pointed out the inadequacies of the statistical journals in refereeing paper on computer algorithms remarking that "statisticians do not read computing journals, and potential authors of high quality papers on computing and statistics have good reason to want to address their market of users directly" (Report of the ASA task force on statistical computing 1970-1971, American Statistical Association Collection, Special Collections Department, Iowa State University Library, Box 56 Folder 7).

department their task was also to deal with the Hollerith equipment, punch the data tape for the Elliott 401 and handle the output of the mainframe.[641] Although the human computers interacted with more advanced technologies, they progressively lost control on the calculations performed in the department and, with the arrival of the digital computer, they became just scientific clerks who did not even need to know the "complexities of routine computation, such as the analysis of variance".[642]

Fig. 4.2 Statistics and field experiments departments (1958). In the list below I have underlined the members of the statistics department distinguishing between the statisticians (S) and the assistant staff (A).

Back Row (left to right): H. D. Patterson (S), I. Canton, M. H. Westmacott (S), C. R. L. Scowen, J. H. A. Dunwoody (S), G. F. Jarvis, S. Lipton (S), F. B. Leech (S), J. C. Gower (S).

Middle row: (left to right): Mary D. Rainbow (A), Betty Sparkes (A), Joan E. Coiley (A), Evelyn M. Hawkins (A), Mary G. Hills (S), Myrtle Hughes, Doreen Mitchell (A), Constance Hunt (A), Florence V. W. Jordan (A), Edith E. Challenger (A), Sheila Lawrence (A), Verona A. Roberts (A), Margaret P. Whittingham (A), O. B. Chedzoy (S), H. R. Simpson (S), D. H. Rees (S).

Front row (left to right): Ann Newman, Patricia R. Walters, Betty E. Artiss (A), M. J. Allen, H. V. Garner, F. Yates (S), M. J. R. Healy (S), D. A. Boyd (S), Ruth T. Hunt (A), Marjory G. Morris (S), Muriel E. Davis (S), Doreen Bispham (A), Edith Spetch (S).

Credits: Copyright Rothamsted Research Ltd.

---

641 For more information on the routine work of the assistant staff in the Rothamsted statistics department during the 1960s see the interview with Mrs (Lucille) Vera Wiltsher, who worked there as data processor on the Elliott 401 and the Orion computer. (Appendix 4.3)

642 F. Yates (1960b), p. 210.

4.3 ***Alliances in the computerization of the Rothamsted statistics department***

4.3.1 ***The claims for a digital computer***

In the preparation of a comprehensive plan for the post-war reorganisation of agricultural science, the ARC relied on visiting groups formed by scientists and council officers who reviewed the research activity of the grant-aided institutions.[643] The departments of Rothamsted Experimental Station were visited between May and November 1953 by several groups that drew preliminary reports on the current activity of the station departments, their planned expansion in the following five years, concerns about facilities and staff.[644] In October 1953 the physicist Patrick M. S. Blackett and the mathematician and geneticist J. B. S. Haldane, accompanied by an ARC officer, reviewed the work of the Rothamsted statistics department and were asked to report in detail on its request to build a mainframe.[645]

Blackett and Haldane commended Yates' department, considered "well staffed and organized", "productive both of valuable practical and theoretical results", "in short, very efficient and a most valuable national asset".[646] Blackett was well-disposed towards the acquisition of a digital computer at Rothamsted, but did not support the idea of a home-made machine.[647] In the final report he and Haldane suggested that "a commercially built machine of proved design" should rather be purchased to carry out "more complex statistical work, for instance multivariate analysis and the analysis of more extensive surveys and experiments".[648] Blackett and Haldane suggested also that the Rothamsted statisticians, tinkering with a mainframe, could "study the general problems arising in the handling of agricultural and biometric data by electronic machines".[649]

Thus, in the acquisition of a digital computer for the Rothamsted statistics department the main argument was not to pursue an economy of time and labour, but to reach new goals in the application of statistical methods and to explore to what extent mainframes could be useful in

---

643 For the post-war expansion of agricultural science sponsored by the ARC see G. W. Cooke (1981), pp. 49-67.
644 Rothamsted Experimental Station, Report of the Visiting Groups, ARC 18/54, TNA, MAF 200/21.
645 For a scientific biography of Blackett and Haldane see their obituaries as fellows of the Royal Society (B. Lovell (1975), N. W. Pirie (1966)) and the respective entries in the *Oxford Dictionary of National Biography* (M. J. Nye (2004), V. M. Quirke (2004)). On Haldane see also chapter 3 (3.4). On the intention to build a computer for the Rothamsted statistics department see F. Yates (1966), p. 234.
646 Report on statistics department, ARC 18/54 (Appendix J), TNA, MAF 200/21 (Quotation reproduced with permission from the BBSRC).
647 F. Yates (1966), p. 234.
648 Report on statistics department, ARC 18/54 (Appendix J), TNA, MAF 200/21 (Quotation reproduced with permission from the BBSRC).
649 Report on statistics department, ARC 18/54 (Appendix J), TNA, MAF 200/21 (Quotation reproduced with permission from the BBSRC). The report contained also the suggestion, turned down, to constitute at Rothamsted an operational research section on agricultural policy.

agriculture and biology.[650] The perspective suggested by Blackett and Haldane's report met – or possibly was recommended by – the vision of Frank Yates, who considered computing tools in statistics as elements belonging to the research activity itself.[651] In fact, Yates, in his repeated requests to the ARC for a mainframe, pointed out the difficulty, almost impossibility, to tackle certain statistical problems, such as multivariate analysis and multivariate regression, with the current computing tools of his department, and advocated the digital computer as the only opportunity for removing this stumbling block.[652]

In 1954 the ARC officially examined Yates' request on account of the positive report received by the visiting group. The council considered which digital computers – already available on the market or soon to be commercialized in Britain – could be suitable for Rothamsted. Four alternatives were examined: a Ferranti Mark I, considered, however, too expensive (ninety thousand pounds) and "much larger" than needed; a DEUCE produced by English Electric, "cheaper than the Ferranti"; a smallish computer produced by the British Tabulating Machine Company costing fifteen thousand pounds; a prototype machine, the Elliott 401 (estimated cost twenty-five pounds) manufactured by Elliott Brothers for the NRDC, already completed, but under improvement work at the Mathematical Laboratory at Cambridge University.[653] Eventually Rothamsted acquired the Elliott 401 on lease from the NRDC. The digital computer arrived at the station in April 1954.[654]

The computerization of the Rothamsted statistics department presents several affinities with the cases examined by Jon Agar in his comparative study on the acquisition of digital computers in scientific research during the 1950s. Agar points out that computerization took place "where there *already existed* material and theoretical computational practices and technologies" and that it "was usually first proposed when the existing practices and technologies were still capable of the

650 See also the statistics department, Rothamsted: proposal to acquire electronic calculating machine, office note, ARC 15/54, TNA, MAF 200/19: "The justification for having such a machine is not so much that it would enable existing types of calculations to be done more expeditiously" (Quotation reproduced with permission from the BBSRC).
651 In Blackett and Haldane's report Yates particularly liked the remark on the potentialities offered by tinkering with the machine. (F. Yates (1966), p. 234)
652 Statistics department, Rothamsted: proposal to acquire electronic calculating machine, office note, ARC 15/54, TNA, MAF 200/19. Actually the first attempts at multivariate analysis in the Rothamsted statistics department were made with the Hollerith equipment, but the results were not satisfactory (RES (1953), p. 132).
653 Statistics department, Rothamsted: proposal to acquire electronic calculating machine, office note, ARC 15/54, TNA, MAF 200/19 (Quotations reproduced with permission from the BBSRC). For an overview of the first digital computers marketed in Britain during the 1950s see S. Lavington (2011), pp. 217-220. Since November 1953, however, it was clear that the ARC could not afford a Ferranti computer and that the more likely solution was the Elliott 401. For the estimated cost of the Elliott 401 see the letter from W. K. Slater to Lord Halsbury, 2nd November 1953, NAHC, Manchester, NAHC/NRD C10/8.
654 RES (1955), p. 137.

computational task at hand, but a plausible claim could be sustained that they were straining at their limits".[655]

Both these aspects can be found in the acquisition of the Elliott 401. The Rothamsted statistics department had a long tradition in scientific computation and it had already started the mechanization of computing work in the 1940s with the acquisition of punched-card equipment. Yates and his co-workers had certainly plausible claims for an electronic computer: digitization could improve the analysis of agricultural experiments and survey allowing to study more than one variable at a time and the development of multivariate analysis in biology would benefit as well from the availability of a digital computer.

Thus, at Rothamsted certainly "computerization coincided with qualitative shifts in what scientists were able to do".[656] Among the statistical problems in agriculture and biology that could be handled with the electronic computer there were the analysis of replicated experiments, probit analysis, sampling problems and experimental errors, a wide array of genetic problems ranging from the estimation of linkage to the fit of gene frequencies by maximum likelihood, the mathematical modelling of insects populations, models of cattle metabolism, survey analysis, the computation of mathematical tables and the opportunity to build up expertise in multivariate analysis.[657] Frank Yates baptized mainframes the second revolution in statistics – the first being the introduction of desk calculators – and forecast "that the much better computational facilities provided by electronic computers will stimulate the rapid development of new branches of statistical theory and methodology".[658]

However, as Agar underlines, revolutionary claims in the historical accounts of computerization during the 1950s are often rhetorical strategies that deserve investigation, not unquestioning acceptance.[659] For instance, the computer revolution declared by Yates clashes with the co-existence of mechanical desk-calculators and punched-card equipment in his department throughout his whole career and their constant use despite the availability of the digital computer. It is important to point out, also, that all the statistical software for the Rothamsted mainframe had to be developed by the staff of the department, and years were spent in acquiring programming skills and building up a library of routines for statistical work. Besides, the software was hardware-dependant in the 1950s and thus the efforts of Yates and his co-workers were basically valuable only for their prototype computer and the acquisition of a new

[655] J. Agar (2006), p. 872; p. 873.
[656] J. Agar (2006), p. 899.
[657] The electronic computer at Rothamsted, progress report, 1954-1956, NAHC, Manchester, NAHC/NRD C10/8.
[658] F. Yates (1966), p. 235.
[659] J. Agar (2006), p. 872.

mainframe in the 1960s required again a consistent programming work because the routines prepared for the Elliott 401 had to be adapted to the new machine.

Moreover, digital computers were not within the reach of every researcher or institution in post-war Britain. Again Agar reminds us that "the availability of computers, especially in the 1950s and 1960s, depended on contacts with providers of the scarce machines. [...] Computerization therefore depended on one's position within labour economies, patronage ties, and moral economies of science."[660] For the Rothamsted statistics department it was Yates' engagement in operational research during WWII that offered the connections able to prompt the computerization.

4.3.2 *The legacy of operational research*

The thesis that operational research influenced computerization in the sciences, in particular the life sciences, after WWII is not novel. Soraya de Chadarevian has argued that the protein crystallographer John C. Kendrew turned to digital computers for solving the structure of myoglobin due to his wartime experience in operational research. De Chadarevian claims that Kendrew approached the analysis of protein structure in "operational terms", and this attitude drove him towards digital computers as effective systems to handle and retrieve large amounts of data.[661]

Following de Chadarevian's argument, Joseph November has examined the career of Robert S. Ledley, licensed dentist, MS in physics and mathematics at Columbia University, operational researcher and computer enthusiast.[662] In Ledley's vision the mathematization of biomedicine depended on the coupling of operational research methods, which provided a way to systematically reduce a complex phenomenon in simple components, with the computational power of digital computers, which could solve equations in numerical form.[663]

In de Chadarevian's and November's accounts operational research promoted computerization in as much as it offered a model of reasoning suitable for tackling complex biological and medical problems and for reducing them to a format manageable with digital computers. In my case study, instead, operational research counted for the personal connections it created between Frank Yates and the people who granted him access to the Elliott 401.

---

[660] J. Agar (2006), p. 873.

[661] S. de Chadarevian (2002), pp. 118-120. The case of J. C. Kendrew is considered also in the comparative analysis of J. Agar mentioned above (J. Agar (2006), pp. 888-892).

[662] J. November (2012), pp. 43-44.

[663] J. November (2012), pp. 43-56.

Blackett and Haldane, the members of the visiting group that reported on the electronic computer, were long-term acquaintances of Frank Yates. Both had been members of the *Tots* and *Quots* and actively engaged in operational research during WWII.[664] Besides the ARC scientific advisors, also key members of the ARC had been engaged in operational research along with Frank Yates. First of all, the ARC secretary, William Slater, who had been a member of the *Tots* and *Quots* and advisor on agricultural policy, since 1941 in the AIC, and from 1949 in the ARC.[665] Besides Slater, also Solly Zuckerman served in the ARC. His influence on the acquisition of the Rothamsted mainframe cannot be assessed from the surviving archival and published sources, however it is reasonable to assume that his vote went in favour of the request made by Yates, as Zuckerman had already received statistical advice from Yates' department and with his co-workers at Birmingham University directly benefited from the Elliott 401 for the comparative analysis of the teeth of humans and great apes.[666]

According to Yates' later account, it was "[Patrick Blackett] and the then Secretary of the ARC, Sir William Slater, who were mainly responsible for enabling us to get started with a computer at Rothamsted".[667] In particular, Blackett had a key role in the computerization of the Rothamsted statistics department, because he did not only contribute to a positive report along with Haldane, but was also a member of the NRDC council and recommended the corporation to lease the Elliott 401 to the ARC.[668] Yates' posthumous account is confirmed by the correspondence exchanged since 1953 between the secretary of the ARC and the president of the NRDC.

> I have had a word with Lord Rothschild [chairman of the ARC], who has seen Blackett about the electronic calculator for Rothamsted. I also took an opportunity of having a very short discussion with Yates. We all feel that it would be very advantageous if it were possible for the NRDC to agree to the loan of the Elliot machine which is in Cambridge for a number of years, in order that Yates might experiment with its use in connection with agricultural problems and operational research surveys.[669]

As evident in Blackett's case and, in part, also for Zuckerman, Yates' acolytes in operational research shared his interest for digital computers and thus it was not too difficult for the head of the Rothamsted statistics department to seek their collaboration in his attempt to introduce digital computers in statistics for agriculture and biology. The acquisition of the Elliott 401 at

[664] S. Zuckerman (1988), p. 402, p. 109; M. Fortun and S. S. Schweber (1993). Blackett is considered a founding father of British operational research.

[665] For more information on W. Slater see his obituary as fellow of the Royal Society, H. D. Kay (1971).

[666] F. Yates and M. J. H. Healy (1951); RES (1952), p. 32; RES (1957), p. 172; E. H. Ashton et al. (1957).

[667] F. Yates (1966), p. 233.

[668] F. Yates (1966), p. 234. Yates acknowledged Blackett's key role in the computerisation of the Rothamsted statistics department even congratulating Blackett upon his nomination as Companion of Honour: "Your work for us on the computer front continues to bear fruit. For this we are eternally grateful". (Letter from F. Yates to P. M. S. Blackett, 15th June 1965, RSCHS, Blackett Papers, PB/1/33/3/8)

[669] Letter from W. K. Slater to Lord Halsbury, 2nd November 1953, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

Rothamsted Experimental Station, thus, can be added to the cases of computerization in the sciences in which strong claims for the relevance of operational research can be made.[670]

In my case study, however, the emphasis should be placed on operational research as an opportunity for personal contacts among people favourable to digital computers rather than on operational research as *modus operandi*. In fact, the Rothamsted statisticians already used to interpret experimental problems in agriculture and biology as liable to analytical and computational solutions. They did not need a justification for their choice to digitize statistical problems, but rather the opportunity to test how the acquisition of a mainframe could improve and broaden their established effort in the mathematization of agriculture and biology.[671]

4.3.3 *The agreements with the National Research Development Corporation*

The NRDC was created in 1948 "to promote the adoption by industry of new products and processes invented in government laboratories, universities and elsewhere, advancing money where necessary to bring these to commercially profitable stage".[672] Throughout the 1950s and 1960s the corporation played a major role in the development of the British computer industry offering research and manufacturing contracts to companies considered able to enter the new market. The Elliott 401 was the result of one such contract placed with the company Elliott Brothers.

The first contacts between the NRDC and the ARC on the acquisition of a mainframe for the Rothamsted statistics department began in November 1953.[673] At that stage, however, the NRDC could not assure the availability of its Elliott 401, but granted to the ARC the first refusal in case of sale, suggesting twenty-five thousand pounds as the right price for the transaction.[674]

In the following months the idea that the ARC should buy the digital computer outright from the corporation was left aside in favour of "a collaborative effort in operating the machine at

[670] It has been argued that more generally digital computers became the tools of the trade within a systems approach – including operational research, systems engineering, systems analysis, system dynamics – and that the computer mystique fuelled the work of experts in this field (T. P. Hughes and A. C. Hughes (2000), pp. 1-24).

[671] The connection between statistical laboratories engaged in the analysis of agricultural data and digital computers is not a peculiarity of the Rothamsted statistics department. For instance, D. A. Grier has argued that they were a "fertile ground for the development of electronic computers", considering the Iowa State College Statistical Laboratory, as a context for the development of the ABC computer, a forerunner of the digital computer. (D. A. Grier, 2000)

[672] S. T. Keith (1981), p. 92. On the history of the NRDC see also J. Hendry (1990). A first-hand account of the NRDC's involvement in British computing is H. J. Crawley (1993). Crawley was the computer projects manager for NRDC.

[673] Letter from W. K. Slater to Lord Halsbury, 2nd November 1953, NAHC, Manchester, NAHC/NRD C10/8.

[674] ARC's interest in computing machines, visit by Sir William Slater on 16th December 1953, NRDC internal memo, 18th December 1953, NAHC, Manchester, NAHC/NRD C10/8.

Rothamsted".[675] The first rental agreement for the Elliott 401 was concluded in May 1954 by the NRDC, the ARC, and the Lawes Agricultural Trust, the body governing Rothamsted Experimental station. It was "a formal but not too formal agreement" lasting one year and stating the move of the computer to Rothamsted and the transfer of its routine maintenance from the NRDC to the staff of the statistics department.[676] The ARC paid two thousand pounds as a rent for the mainframe and guaranteed that the computer remained in working conditions, while the corporation retained the right to access the machine and "as a working principle it was agreed that roughly 50% of the time should be available to Rothamsted and 50% to the NRDC".[677]

Yates' judgement on the computer that arrived at Rothamsted is trenchant. The computer was "in a very imperfect state. Only one-third of the store was utilisable and the machine was none too reliable in operation".[678] Moreover, the software for the computer was almost non-existent. Among the staff of the Rothamsted statistics department it was Douglas Rees who closely followed the Elliott 401 since its arrival at the station and in association with the NRDC technicians, in particular Harry C. Carpenter, did the preliminary work on the computer and put it in working conditions.[679] As of October 1954, Rees was the only person at Rothamsted "fully conversant with both engineering and programming aspects of the machine", and he was in charge of "the greater part of the burden of supervising programs, advice on checking out of programs, maintenance of the machine, and training of other staff in its use".[680] For the routine maintenance of the Elliott 401 at Rothamsted also a technician, D. J. Knight, was hired and trained.[681]

---

675 15th Sub-Committee Report, 26th January 1954, NAHC, Manchester, NAHC/NRD C10/8; Letter from D. Hennessey to W. Slater, 1st March 1954, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

676 Letter from D. Hennessey to F. Yates, 24th February 1954, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

677 Note of the meeting between the NRDC and Rothamsted statistics department, 12th February 1954, NAHC, Manchester, NAHC/NRD C10/8 (quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester); 16th Sub-Committee Report, 23rd February 1954, NAHC, Manchester, NAHC/NRD C10/8; 55th Board, 28th April 1954, NAHC, Manchester, NAHC/NRD C10/8.

678 The electronic computer at Rothamsted, progress report, 1954-1956, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

679 Probably Rees spent time also in Cambridge with the NRDC staff, before the transfer of the Elliott 401 to Rothamsted. (Letter from H. G. Carpenter to D. H. Rees, 16th February 1954, NAHC, Manchester, NAHC/NRD C10/8)

680 Letter from H. G. Carpenter to D. H. Marlow, 28th October 1954, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester. D. H. Rees, along with F. Yates, was also involved in the set up of the British Computer Society (RES (1958), pp. 182-183), and became the head of the computer department created in 1968, after Yates' retirement.

681 The technician was appointed by the NRDC at a salary of £575 per annum and his salary then reimbursed by the ARC to the NRDC (Letter from H. G. Carpenter to D. J. Knight, 21st April 1954, NAHC, Manchester, NAHC/NRD C39/1; Letter from R. A. E. Walker to F. Yates, 4th May 1954, NAHC, Manchester, NAHC/NRD C10/8). Knight was transferred from the NRDC to the Rothamsted staff in 1956. For Knight's training see Letter from F. S. Ellis to Messrs. Creed & Co., 14th June 1954, NAHC, Manchester, NAHC/NRD C10/8.

Despite the unfavourable initial impression, the overall performance of the computer during the first year was satisfactory and the ARC and Rothamsted decided to extend the lease of the Elliott 401 for other five years at the same annual rent.[682] Again the Rothamsted staff and the NRDC had "the right to use the 401 each for one half of the time in which the 401 is in condition for operation".[683] The NRDC rented computer time to external users at a cost of twenty-five pounds per hour of 'useful time', where the useful time was calculated considering the amount of work done by the Elliott 401, when no technical problem arose, while the time spent in checking the program or running it for doubtful values was not charged.[684] Among the external users of the Elliott 401 there were private companies as well as research institutions. To the former group belonged the De Havilland Aircraft Company, which solved a flutter problem with the NRDC mainframe, and the Distillers Company, interested in the use of electronic computers for statistical analysis, while among the research institutions there were the European Organization for Nuclear Research (CERN), which performed some computations for the design of a cyclotron; the National Institute of Industrial Psychology, interested in analysing its extensive collection of experimental results; the Electricity Authority, willing to determine characteristics functions for transmission lines; the Directorate of Colonial Office for a geodetic survey.[685]

Frank Yates and his colleagues, as host of the Elliott 401, were involved in the management of the external users and contributed to earn "some more cash for the kitty", assisting with programming advice and sometimes punching tapes on their behalf.[686] Throughout the years, however, the NRDC did not fully employ its share of the computer time giving more freedom to the Rothamsted statisticians to experiment with the machine.[687]

With the new agreement Rothamsted acquired the right to make "reasonable minor modifications to the 401", as to improve its performance, ease programming and add input and output facilities able to integrate the computer in the working routine of the statistics department.[688] The agreement between the ARC and the NRDC for the Elliott 401 was further

---

682 For an account of the first four years of the Elliott 401 in the Rothamsted statistics department see F. Yates and D. H. Rees (1958). More generally on the development over time of statistical computing at Rothamsted see J. C. Gower (1986).

683 Agreement for hire of the 401 electronic digital computer, 22nd November 1955, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

684 Letter from F. Yates to H. H. Brazier, 29th April 1955, NAHC, Manchester, NAHC/NRD C15/2.

685 Computers, Elliott 401, user enquiries: correspondence 1952-56, NAHC, Manchester, NAHC/NRD C15/2.

686 Letter from D. H. Rees to F. S. Ellis, 7th June 1957, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

687 The electronic computer at Rothamsted, progress report, 1954-1956, NAHC, Manchester, NAHC/NRD C10/8.

688 Agreement for hire of the 401 electronic digital computer, 22nd November 1955, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

renewed and lasted until 1965, when the computer was officially switched off and the corporation donated it to the Science Museum.

4.4 *The Elliott 401 in the Rothamsted statistics department*

4.4.1 *Brief history of the Elliott 401 and its role in the British computer industry*

In the aftermath of WWII the development of computers in Britain was mainly a prerogative of three centres: Manchester University, where a prototype machine was in use since 1948 and a full scale mainframe since 1949; Cambridge University, where the team of Maurice Wilkes, head of the Mathematical Laboratory, completed in 1949 its own mainframe, the EDSAC; the National Physical Laboratory at Teddington with the Pilot ACE, a reduced version of the computer originally designed by Alan Turing, available since 1950.[689]

In 1949 the NRDC "inherited a portfolio of patents resulting from the pioneering work of various research teams including those at the Universities of Cambridge and Manchester, the National Physical Laboratory and the General Post Office" and went on subsidizing the development of British computers.[690] By 1955 in Britain there were about thirty computers in operation, all manufactured in the country.[691]

Elliott Brothers, the company that developed the Elliott 401, had a tradition in instrument making and several connections with the Admiralty for classified work.[692] In 1950 the NRDC approached the company with a proposal for collaboration, remembers W. S. (Bill) Elliott, the manager in charge of the computer division of Elliott Brothers.

> They had heard of the work we were doing in the development of techniques for the construction of electronic digital computers for special purposes and they wanted to see if they could come to some arrangement with us, whereby we would develop, under their auspices, general purpose machines making use of these techniques.[693]

The NRDC offered a contract for the development of a prototype computer, later called Elliott 401 after the accounting number given to the project by Elliott Brothers.[694] The Elliott 401 was built between September 1952 and March 1953 at the Borehamwood laboratories of Elliott

---

[689] M. Campbell-Kelly (1989), pp. 163-165. Additional sources on the early days of digital computers in Britain are S. Lavington (2011), pp. 650-670; M. Croarken (1990), chapter 10; J. Agar (2003), pp. 266-274. On the EDSAC, which was the computer used by J. C. Kendrew for his crystallographic calculations, see also S. de Chadarevian (2002), pp. 107-111.
[690] S. T. Keith (1981), p. 121.
[691] S. Lavington (2011), p. 661.
[692] For more information on the history of Elliott Brothers see S. Lavington (2011) chapter 1. For the making of the Elliott 401 and the interaction between NRDC and Elliott Brothers see S. Lavington (2011), pp. 147-174.
[693] Letter from W. S. Elliott to J. C. P. Miller, 10th March 1954, SM London T/1965-445. On W. S. Elliott see S. Lavington (2011), p. 38.
[694] S. Lavington (2011), p. 152.

Brothers. The logic scheme of the computer was entrusted to Andrew St. Johnston, the project leader, who was a graduate in electrical engineering with a wartime experience as radar officer.[695]

The Elliott 401 was a “minimal machine built around a small magnetic disc store and employing single word magnetostrictive lines for the [...] arithmetic registers”.[696] It used about five hundred thermionic valves.[697] Its innovative aspect was the use of plug-in packaged circuits, that is the mainframe was built with a limited set of standardised units, which should reduce on the one hand the overall cost of manufacturing the computer, and on the other, in case of fault, could be immediately changed by the operator.[698]

The Elliott 401 was the starting point of the Elliott 400 series, which included three commercial models (402, 403, 405), all produced using packaged circuits. The Elliott 402 was the production version of the prototype 401 and in 1961 the company Elliott Brothers gave to Rothamsted an exemplar of 402, which was previously employed at the Elliott Computing Service Bureau. In Yates’ department the newly acquired Elliott 402 was fitted with input/output devices for punched-cards and used for the analysis of surveys.[699]

Although produced by Elliott Brothers, the Elliott 401 was owned by the NRDC, which had paid for its development and manufacturing. In April 1953 the computer was demonstrated at the Physical Society Exhibition in London and from June 1953 to March 1954 the mainframe was hosted by the Mathematical Laboratory at Cambridge University, where the NRDC invested in further engineering and logical development.[700] Despite the new effort, when the mainframe moved to Rothamsted in 1954, it needed further adjustments before becoming fully operative.[701]

At Rothamsted the computer remained in active service for eleven years. During this period changes were made to the hardware as to facilitate programming, increase the memory storage and provide new facilities for the input and output of data.[702] During its time at Rothamsted the Elliott 401 had periods of satisfactory performance alternated to “bad spells”.[703] However, the

[695] For biographical information on A. St. Johnston see S. Lavington (2001), p. 174, for his involvement in the Elliott 401 project see Letter from W. S. Elliott to J. C. P. Miller, 10th March 1954, SM London T/1965-445.
[696] The Elliott-NRDC Computer 401, SM London T/1965-445.
[697] S. Lavington (2011), p. 161.
[698] The Elliott-NRDC Computer 401, SM London T/1965-445.
[699] RES (1962), p. 164.
[700] Letter from F. S. Ellis to the Secretary, Ministry of Supply, 22nd November 1954, NAHC, Manchester, NAHC/NRD C10/8.
[701] Letter from D. Hennessey to P. M. S. Blackett, 19th February 1957, NAHC, Manchester, NAHC/NRD C10/8.
[702] S. Lavington (2011), p. 539.
[703] Summary of the performance of the 401 at Rothamsted during May 1955, NAHC, Manchester, NAHC/NRD C10/8. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester.

only major breakdown was the failure of the magnetic disc store in 1960, replaced with a more modern magnetic drum store.[704]

Overall, the Elliott 401 gave a reliable service at Rothamsted and Frank Yates did not regret the choice to keep the NRDC mainframe after the first year.[705] Only in 1960 it was decided to replace the Elliott 401 with an Orion produced by the computer company Ferranti.[706] The Orion was delivered at Rothamsted in December 1963 and in the meantime the agreement with the NRDC was renewed and the first mainframe remained operative in the department, while suitable software was prepared for the new computer.[707]

The working life of the Elliott 401 ended in July 1965, when it was switched off during an official ceremony held at Rothamsted Experimental Station at the presence of representatives of the NRDC, Elliott Brothers and the experimental station. In that occasion the Science Museum took charge of the machine as a valuable output of the early British computer industry.[708]

During the closing down ceremony, the NRDC manager H. John Crawley who had closely followed the development of the computer gave a petty accounting of the Elliott 401 project for the corporation balance. The initial expense for the development and maintenance of the computer in its early years had been about sixty-two thousand pounds, only partially covered by the revenues (about twenty-one thousand pounds) paid by Elliott Brothers to the NRDC for the mainframes of the Elliott 400 series and by the rental agreement with the ARC, which gave approximately twenty-two thousand pounds.[709] However, the relevance of the Elliott 401 for the NRDC and the British computer industry can be assessed in a more generous way in terms of the new potential market for mainframes that the NRDC had contributed to open. According to Simon Lavington, in fact,

---

[704] RES (1961), pp. 176-177.

[705] The electronic computer at Rothamsted, progress report, 1954-1956, NAHC, Manchester, NAHC/NRD C10/8. For the performances of the Elliott 401 at Rothamsted and its impact on the work of the department see RES (1957), p. 168; RES (1958), p. 176; RES (1959), p. 156; RES (1960), pp. 151-152; RES (1961), p. 177; RES (1962), pp. 163-164; RES (1963), p. 174. Since the late 1950s, in order to cope with the workload, the mainframe was also run 'overtime', i.e. out of the office hours.

[706] RES (1961), p. 178.

[707] RES (1964), p. 169.

[708] "I have come to the conclusion that the machine ought to be preserved at the Science Museum and that it will be possible, with a little help from Elliot's, to make it into an interesting and instructive historic exhibit." (Letter from H. R. Calvert to H. J. Crawley, 2nd March 1965, NAHC, Manchester, NAHC/NRD C11/3. Quotation reproduced by courtesy of the University Librarian and Director, UML, The University of Manchester). According to the original plans, the computer should have been on permanent display in the Science Museum since 1966 (Letter from D. H. Rees to D. Hennessey, 21st July 1965, NAHC, Manchester, NAHC/NRD C11/3), but the archives of the institution do not record any public appearance of the mainframe (email communication with Rachel Boon). The Elliott 401 is currently hosted at Blythe House, the museum's store, and volunteers of the Computer Conservation Society (http://www.computerconservationsociety.org/) are restoring it.

[709] S. Lavington (2011), pp. 173-174. The computing time on the Elliot 401 sold to outside customers never gave appreciable revenues.

> For NRDC, the significance of the 401 [...] was that it introduced a second player to the UK marketplace. Ferranti Ltd. was already producing limited numbers of large-scale computers for sale, at a price of about £90 K each. Now the hope was that Elliott Brothers, or some derivative of Elliott Brothers, could produce small- or medium-scale machines in commercially realistic quantities at a price in the region of £20 K each.[710]

If the figures quoted by Crawley cannot be a satisfactory assessment of the role of the Elliott 401 in the business history of computers, least of all they can recapture the role of the mainframe as a living piece of technology. In his official discourse of acceptance, the Science Museum representative, H. R. Calvert, in fact, brilliantly pointed out that the

> NRDC commissioned the computer and Messrs Elliott made it, but it was still until it was used, a collection of components in metal cupboards. It was Dr Yates and Mr Rees of the Rothamsted Experimental Station who gave the machine a personality – or it should be a computerality – by setting it to work on important statistical problems.[711]

The library of programmes written by Yates and his colleagues certainly gave to the Elliott 401 a statistical soul – the Rothamsted mainframe was likely to be the only computer fully devoted to statistical research in Britain during the 1950s – and, neglecting the minor impact of the NRDC external users, shaped the applications feasible with the Elliott 401, because no other software for the mainframe existed. Moreover, the Rothamsted statisticians had also specific requests on the computer hardware. For instance, they considered a priority to provide the Elliott 401 with punched-card input and output facilities, as to connect it with the Hollerith equipment already available in the department.

Thus, it is interesting to ask alongside the question 'What difference did computers make to science?', which underpins Agar's comparative study of computerization, also the reverse one 'What difference did the sciences [in which digital computers were applied] make to computers?' in terms of hardware improvement and software development.[712] Adopting this approach, Joseph November in his analysis of biomedical computing in the United States has argued, for instance, that the requests of the life scientists consistently influenced the agenda of the computer scientists.[713] To some extent, I will attempt an answer for the Rothamsted mainframe examining below how the Elliott 401 was integrated in the scientific activity of the statistics department.

[710] S. Lavington (2011), p. 163.
[711] Speech of Dr H. R. Calvert, Closing down ceremony of the Elliott/NRDC 401, 30th July 1965, SM London T/1965-445.
[712] J. Agar (2006).
[713] J. November (2012), p. 9. In particular see chapters 3 and 5.

Fig. 4.3 Elliott 401 at Rothamsted Experimental Station (undated). One of the cupboards is left open to show the packaged circuit units. The main cabinet contained the ventilating equipment and power supplies, memory disc, amplifiers and control circuits. Next to the main body of the computer there is the monitor console containing two cathode-ray oscilloscopes and three number generators and the electric typewriter for printing the computer output. The technical features of the computer when it arrived at Rothamsted in 1954 were as follows "the 401 is a serial computer with a word length of 32 binary digits and a word time of 100 μs. The main store is a magnetic disc of 23 tracks, each holding 128 words (orders or numbers); 7 of the tracks are always immediately available to the computer, but only 1 of the remaining 16 is accessible at any given time, switching between these tracks being by means of high-speed relays. There are 5 single-word immediate-access registers; one is the accumulator, which can be coupled with a second register for double length working; the other 3 registers can each be used to modify orders (B modification) as well as for temporary storage of numbers. Input is by five-hole punched tape and output by either electrical typewriter or teleprinter-punch. The machine has a two-address code to allow for optimum programming." (RES (1955), p. 138)
Credits: Copyright Rothamsted Research Ltd.

Fig. 4.4 Closing down ceremony for the Elliott 401, Rothamsted Experimental Station, 30th July 1965. F. Yates is in the foreground on the left. The ceremony had press coverage, as evident from the journalist and photographer clearly recognizable in the image.
Credits: Copyright Rothamsted Research Ltd.

Fig. 4.5 Orion computer (section) (1965).
Credits: Copyright Rothamsted Research Ltd.

4.4.2 *Old and new computing technologies at Rothamsted in the 1950s*

An handwritten inventory compiled in August 1945 recorded the calculating machines available in the Rothamsted statistics department in the aftermath of WWII. The list included: one new Monroe (not yet arrived); two fully automatic Monroes (one out of order most of the time); three ordinary Monroes; three Millionaires; two hand machines; two adders; three Friden calculating machines; one Monroe and one Marchant calculator belonging to the Galton Laboratory and soon to be returned to University College London; one Brunsviga calculator of the British Association Mathematical Tables Committee.[714]

Apparently the computing equipment had not changed much from the pre-war days, when statisticians and human computers working at Rothamsted could rely only on mechanical and electro-mechanical calculators (**chapter 1**), but that would be a false impression. WWII had brought in the department a technological upgrade with the direct availability of Hollerith punched-card equipment. For example the returns of the wartime farm survey had given the opportunity for a systematic use of punched-card equipment and Oscar Kempthorne, the statistician in charge of the investigation, had published a detailed account of the use of this equipment in the survey.[715]

---

[714] Calculating machines inventory, 29th August 1945, RR Library and Archive, SITE 1/5/2. For the technical features of desk calculating machines in the Rothamsted statistics department see chapter 1. For an update on the desk calculating machines available in the 1940s see Office Machines Research Inc. (1940). On the availability of calculating machines belonging to the Galton Laboratory and to the British Association Mathematical Tables Committee see respectively chapter 3 and chapter 2.
[715] O. Kempthorne (1946).

In 1948 a sorter-counter was available for the Rothamsted statistics department, replaced in 1949 by a Junior Rolling Total Tabulator and a sorter.[716] The tabulator had a feeding mechanism for reading the cards, counters where the results of the calculations done from the cards were stored, and an interchangeable plugboard that allowed the machine to perform several tasks, such as listing, adding, subtracting, according to the way in which the wires were connected by the operator. It was even possible to print out the results of the operations done by the tabulator. In 1950 Rothamsted acquired also a reproducer-summary punch to duplicate the data on a card and to punch the results of the calculations done on the tabulator not on paper, but again on punched-cards.[717]

The reports of the statistics department constantly described the Hollerith equipment as a research tool that needed to be "on the spot" under the direct control of the statisticians to maximize its potential for research.[718] Indeed, the Rothamsted staff suggested alterations to the punched-card equipment that could facilitate their daily work. For instance, the tabulator was modified as to control the distribution between counters of the numbers read from the cards in more than one control column.[719] In that way it was possible to separate at once the different treatments under consideration in the analysis of replicated factorial experiments.[720] The tabulator was also employed in the analysis of surveys and offered an opportunity to approach multivariate analysis in biological research and to solve mathematical problems.[721]

The use of the Hollerith equipment represented the first attempt at the mechanisation of the work in the Rothamsted statistics department and therefore a first step towards the later computerization. It suggested strategies for condensing statistical information in new formats and procedures for splitting routine work in simple operations that could be performed with the punched-card equipment, but it did not entail the dismissal of desk calculators nor made redundant the human computers, as the methods available for dealing with punched-card equipment in the analysis of agricultural and biological data were still unsatisfactory, when compared with the performance of desk calculators.[722] The Hollerith equipment, thus, did not radically change the computing practices of the department, but on the one hand aided in the

716 RES (1949), pp. 49-50. On the rolling total tabulator see M. Campbell-Kelly (1989), pp. 90-94. For a description of the use of punched-card equipment in statistics see H. O. Hartley (1976). F. Yates (1960a) also gives a detailed description of the punched-card equipment used in surveys (pp. 112-127).
717 RES (1952), p. 74.
718 RES (1950), p. 57.
719 RES (1951), p. 67.
720 Factorial experiments are experimental set-ups in which more elements (for instance several combinations of fertilizers) are tested at once.
721 RES (1951), pp.67-68.
722 F. Yates and D. H. Rees (1958), p. 49. There was Hollerith equipment in the Rothamsted statistics department at least until the late 1960s (RR Library and Archive, SITE 1/5/2).

execution of computations, especially when large amounts of data were involved, and on the other offered the opportunity to tackle problems which were rather difficult to approach on desk calculators, such as the analysis of co-variance.

When the Elliott 401 arrived, the Rothamsted statistics department wanted to combine the mainframe with the Hollerith equipment. Among the priority modifications listed by the Rothamsted staff for the computer, in fact, there was the acquisition of input/output equipment for connecting the mainframe to the Hollerith card system. Despite the high cost – the estimated expense was about one thousand pounds, half the annual rent of the mainframe – the Rothamsted statistics department was ready to pay as "many simple arithmetical operations require to be performed on data which can then be efficiently analysed by punched-card machines, and [...] because a great deal of data which can best be analysed on electronic machines at present exists on punched cards."[723]

In order to secure an input and output from punched-cards for the Elliott 401 negotiations began since June 1954 with the British Tabulating Machine Company and the NRDC.[724] The first results of the collaboration were visible in 1956 when a punched-card reader was connected to the mainframe.[725] The tape to card converter, instead, was delivered only in January 1958.[726]

The Rothamsted statisticians evaluated their experience with the Elliott 401 in a similar way as the one with the Hollerith equipment, "where we found that it was only by having equipment on the spot, so that research workers could themselves use it, test out different methods and examine the results as they were obtained, that we could exploit its full potentialities."[727] It was tinkering with the mainframe that new results in statistics were possible, and such tinkering meant for the most part the development of statistical software. However, programming in machine language the Elliott 401 was a heavy task, as the computer had not been designed having ease of programming as a main goal.[728]

The optimal programming strategy for the Elliott 401 imposed to locate successive instructions in the computer store – a magnetic rotating disc – as to minimize the number of rotations in order to move from an instruction to the following one. Optimum programming increased the overall speed of work of the computer, an asset for the Rothamsted statisticians as their statistical routines usually required the computer to go through several hundreds

[723] RES (1955), p. 140. For the estimated cost of the equipment see Meeting with Mr Thorne ARC on 5th January 1955, NRDC internal memo, 17th January 1955, NAHC, Manchester, NAHC/NRD C10/8.
[724] NAHC, Manchester, NAHC/NRD C11/2.
[725] RES (1957), p. 167.
[726] RES (1958), p. 175.
[727] RES (1955), p. 140.
[728] The Elliott-NRDC Computer 401, SM London T/1965-445.

instructions. Frank Yates developed an automatic programming routine to facilitate the writing of efficient programs.[729] Yates' routine automatized the re-arrangement of the software instructions according to optimum programming and facilitated the discovery of faults in the program structure. The first version was completed in January 1955 and thereafter constantly improved and adapted for writing programs of increasing complexity, such as the analysis of factorial experiments and Latin squares.[730] The routine was "found to produce programmes whose timing is as good as or better than the average human programmer and which contain far fewer errors when they are tested for the first time".[731]

4.4.3 *Making space for the Elliott 401*

In 1946 the Rothamsted statistics department moved to Rivers Lodge, an independent building located close by the central premises of the experimental station.[732] The move of the department can be understood considering the plans for its expansion, in terms of people and equipment, after WWII. However, the new accommodation became inadequate already in 1950, despite the vacancies among the staff and the promise of three more rooms in a cottage adjoining Rivers Lodge. Yates appealed to the Rothamsted director and the ARC for more space. In particular, Yates insisted on the necessity to separate scientific officers from computers as "all workers who have to do mathematical work, writing or reading, when someone else is working a calculating machine in the same room, find the strain almost intolerable".[733]

Plans were thus made for a new accommodation for Yates' department, either in the form of a new building or an extension to the present department.[734] Eventually, for lack of funding, it was decided to build a temporary hut in the garden next to Rivers Lodge. The hut was completed in March 1954, just in time for the arrival of the Elliott 401. To the new accommodation moved also the Hollerith equipment that was hosted until the late 1950s in the same room of the digital computer.[735]

---

729 F. Yates and S. Lipton (1957); further details on the automatic programming routine are in the technical notes available for programing the Elliott 401 in the Rothamsted statistics department (Automatic Programming Routine, F. Yates and S. Lipton, 27th July 1956, NAHC, Manchester, NAHC/NRD C39/3). The automatic programming routine prepared for the Elliott 401 was inspired by a similar programme written for the IBM 650, a popular mainframe during the 1950s, whose memory was a magnetic rotating drum.

730 RES (1957), p. 169; RES (1960), p. 153.

731 RES (1957), p. 235.

732 Letter from E. G. Davy to J. B. F. Cooper, 7th November 1946, RR Library and Archive, SITE 1/5/2.

733 Memorandum on the statistical department accommodation written by F. Yates, 16th May 1950, RR Library and Archive, SITE 1/5/2.

734 See the correspondence in TNA MAF 117/219 and RR Library and Archive, SITE 1/5/2, among the ARC, the Ministry of Agriculture and Fisheries, the Rothamsted director W. Ogg and F. Yates.

735 Letter from D. H. Rees to P. A. C. Thorne, 14th February 1958, RR Library and Archive, SITE 1/5/2.

As the construction of the hut and the negotiations for the acquisition of the mainframe were almost contemporary, the 'machine room' in the hut was adapted since the design phases as a possible area for the digital computer, in particular the Elliott 401.[736] Early mainframes, in fact, required for their physical dimensions, weight and power consumption an *ad hoc* location. The main body of the Elliott 401, for instance, was 3.96m long, 0.61m deep and 2.29m high, weighted over one (metric) ton, consumed five KVA of power, and demanded some precautions, as for the motor alternator that should be mounted on a plinth and kept detached from the floor.[737] Soundproofing was also necessary for the machine room because both the digital computer and the Hollerith equipment were very noisy.

The creation of a proper physical space in the hut for the digital computer was the material counterpart of the necessity to fit the Elliott 401 into the working economy of the Rothamsted statistics department. Although the mainframe had been acquired mainly for research purposes, it ended up also supporting the main routine tasks of the department, that is the analysis of experiments and surveys. The former was the first set of problems for which digitisation was attempted. Programs were written for the more common experimental designs adopted in agriculture and were then used for the examination of all the field and animal trials that fitted that pattern. The analysis of agricultural experiments represented a consistent share of the annual workload in the Rothamsted statistics department and emphasis was placed on the fact that once the program tape for the Elliott 401 had been punched and checked and another paper tape had been prepared for the data, the experiments were analysed quickly and without heavy interventions of the staff, in contrast "with the time and skill required to plug a Hollerith plug board for even quite simple operations".[738] Considering the post-war expansion of the Rothamsted statistics department as a consultancy centre for the analysis of agricultural experiments, the mainframe was thus a welcome addition to the equipment already available.

The use of the Elliott 401 for survey analysis started later than the computerization of the analysis of experiments because it was necessary to have flexibility in input and output from punched tape to Hollerith punched-cards and vice versa to make computerization of survey analysis convenient.[739] An interesting example is the data analysis of the survey of fertilizers practice. Until 1950 the survey information was collected using edge-punched cards that could be sorted by hand through a knitting needle. In 1950 there was the first attempt at the

[736] Letter from J. B. Bennett to R. A. Hughes, 28th October 1953, TNA, MAF 117/219.
[737] For the features of the Elliott 401 see S. Lavington (2011), p. 161. For the solutions to be adopted in setting the Elliott 401 in the hut of the statistics department see Letter from H. G. Carpenter to D. H. Rees, 16th February 1954, NAHC, Manchester, NAHC/NRD C10/8.
[738] RES (1955), p. 138. See also Appendix 4.3.
[739] RES (1958), p. 175; RES (1956), p. 148.

mechanization of the survey analysis using Hollerith cards, but it proved unsuccessful and it was not repeated. In 1957 the survey was instead analysed with the Elliott 401, but again with mixed results. On the one hand the mainframe "enabled the main results to be obtained much more speedily than had previously been possible", but when the Rothamsted statisticians tried to use the mainframe for more subtle investigations on the fertilizers data it "proved troublesome [...] partly because of the use of paper tape input, and partly because of programming difficulties".[740] By 1960 the Rothamsted statistics department had instead available a general survey program, using punched-card input, which was employed for the detailed analysis of surveys, including the fertilizer one. The 1962 fertilizer survey, for instance, was examined using such program and the Elliott 402 given by Elliott Brothers to the station.

Yates and his co-workers appreciated in the Elliott 401 the possibility "to speed up the routine analyses which are at present done on desk machines, but which are of sufficiently standard type to be programmed electronically", but mainly they insisted on the opportunity to tackle with the mainframe new statistical problems, such as multivariate analysis, which required heavy computations like matrix inversion.[741] It was this philosophy of experimenting statistics and computing together that gave to the Elliott 401 its own place, material and 'moral', in the Rothamsted statistics department, side by side with the desk calculators and the punched-card equipment, already available in the department and that had been acquired with the same attitude.

Fig. 4.6 Hut used by the Rothamsted statistics department (2010).
Credits: Photo taken by the author and published with permission of Rothamsted Research Ltd.

[740] F. Yates and D. A. Boyd (1965), p. 205.
[741] RES (1955), p. 140.

4.5 *Selected applications of the Elliott 401*

4.5.1 *Analysis of replicated experiments*

Since the 1920s, the analysis of agricultural experiments had represented the main share of the annual workload in the Rothamsted statistics department, but in the early 1950s, Frank Yates and his colleagues were expected to examine roughly four times more experiments than in the pre-war years.[742] Moreover, since agricultural experiments are seasonal, the results accumulated in the department at the same time leaving a consistent backlog of data analysis during the winter months.[743] Thus, not all the data could be examined promptly and "research workers [...] had to wait for many months, and even years, before the results of the full analysis of their material became available".[744]

The analysis of agricultural experiments *per se* was not one of the research tasks for which the Elliott 401 had been acquired in Yates' department.[745] Neither was it immediate that digitization could be a convenient solution, due to the variety of experimental designs and the time required for programming. Yates, instead, envisioned in the long run an economy of time with the use of the Elliott 401, at least for the experimental designs more frequently examined in his department, and since the mid-1950s programs suitable for the task began to be written at Rothamsted.

Yates proved right. In 1955, the first year in which the analysis of agricultural experiments was partly digitized, over eight hundred experiments were analysed by Yates' department, twice the work done in 1951.[746] The trend was confirmed in the following years and increased with the availability of new computer programs for the more complex designs. In 1959 over two thousand and seven hundred experiments were examined in the statistics department, all except a negligible amount, sixty-seven, with the Elliott 401. In terms of time and labour-saving, the digitization of agricultural research made the difference for Yates' department, considering that the staff had not grown at a rate comparable to the workload of the department, and indeed, as noted by Yates, it "increased very little since the computer was installed".[747]

Statisticians at Rothamsted were not the only ones in Britain concerned with the digitization of agricultural experiments, neither the first. Already in 1954 the statistician J. G. Rowell, working at the School of Agriculture in Cambridge, had described the analysis of a factorial experiment

[742] F. Yates (1960b), p. 202.
[743] F. Yates et al. (1957), p. 246.
[744] F. Yates (1960b), p. 210.
[745] Report on statistics department, ARC 18/54 (Appendix J), TNA, MAF 200/21.
[746] F. Yates (1960b), p. 202. In 1955 about half of the experiments were examined by hand and half with the electronic computer.
[747] F. Yates (1960b), p. 202.

with the EDSAC, the mainframe available in the Cambridge Mathematical Laboratory.[748] Unlike the Elliott 401, which was devoted only to statistics, for the EDSAC it was never a priority to create a library of routines that could be used to handle all the experimental designs common in agriculture, thus Rowell's work stands out as a one-off attempt rather than a comprehensive effort for the digitization of the analysis of agricultural experiments.

Instead, the programming effort paid-off when the result was a general program that could be used for several cases, as suggested before the Royal Statistical Society since 1952 by the mathematician, statistician, operational researcher and computer scientist, Keith D. Tocher.[749] For agriculture this meant that one general program should be able to handle most, if not all, the types of experiments. Tocher's proposal, however, did not take into account the several limitations of the early mainframes. At Rothamsted, in fact, Frank Yates and his colleagues decided "to write separate programmes for the different types of design, starting with the simpler and more commonly used types".[750] The choice was explained with the limited memory of the Elliott 401 and the lack of programming experience of Yates and his staff.

On the Elliott 401, the Rothamsted statisticians tackled at first the analysis of replicated experiments, that is experiments in which the same treatment was applied to more than one experimental unit.[751] This is a fairly wide category and several agricultural trials with different designs belong to it. The first program written in the department was concerned with the analysis of randomized block experiments and was used to examine several agricultural trials for the National Institute of Agricultural Botany and the NAAS.[752] By 1960 in the Rothamsted department there were available computer programs also for dealing with Latin squares, 3^3 experiments, randomized blocks with split plots, randomized blocks with a $p \times q$ factorial system, 2^n experiments, general orthogonal block analysis.[753] For all the designs mentioned the programs could handle confounding, i.e. neglect some interactions, estimation of error from higher interactions, and missing plots, that is the lack or unreliability of some experimental results.

Although the programs developed by Yates and his colleagues were specific for each experimental design, the input routine for the data was common to them all. This routine had

---

748 J. G. Rowell (1954). The experiment examined by Rowell was a factorial experiment concerned with the addition of three different supplements to bull semen in order to increase conception rate in artificial insemination.
749 K. D. Tocher (1952).
750 F. Yates (1960b), p. 203.
751 On experiments in agriculture see chapter 1.2.2. An experimental unit can be a plot of land, a group of pigs in a pen, a batch of seeds. A detailed analysis of experimental design with a close examination of the possible experimental layouts is W. G. Cochran and G. M. Cox (1957).
752 F. Yates et al. (1957), p. 238. See also RES (1955), p. 139.
753 F. Yates (1960b), p. 203. On randomized blocks and Latin squares see chapter 1.4.2. For all the other experimental designs mentioned here see W. G. Cochran and G. M. Cox (1957). J. C. Gower (1960) gives a detailed description of the program for general orthogonal blocks experiments written for the Elliott 401.

been written in order that the Elliott 401 carried out as much as possible of the clerical work necessary before the experimental data could be considered a suitable input for the mainframe. The routine read the experimental results, performed on them the necessary computations, such as metrical transformations, and stored the resulting quantities in their proper locations in the computer memory.[754]

The routine contained also several checks on the data, as to avoid that some of them could be lost while the computer read the tape or that gross recording and punching errors could be left unnoticed. The routine transcribed also the input headings in the output tape to facilitate the interpretation of the final figures.[755] The use of the same input routine not only reduced the programming labour, but standardized also the punching of the data tape easing the routine of the assistant staff.[756]

The statistical analysis performed by the Elliott 401 on the replicated experiments followed the one traditionally done on desk machines:

> the required tables of means are printed, together with their standard errors; an analysis of variance more or less in standard form is also provided. In addition the residuals are calculated and used as a general check. They can also be printed if required; this is usually done as it enables the experimenter to detect anomalous values and provide evidence of fertility gradients, etc.[757]

However, computerization permitted to refine the analysis of the data, beyond what was possible with a desk calculator. Further data collected by the research workers could now be examined along with the experimental results at a very limited cost of time and effort increasing the number of random variables examined by the program. The analysis of covariance, i.e. test whether two variables examined did not vary independently, was not routinely conducted, but the Elliott 401 could easily compute the residual sum of products between any pair of random variables and then the covariance analysis could be completed on a desk calculator. Digitization made easier also to deal with experiments in which not all the final data were available. For such experiments, an iterative algorithm allowed to minimize the influence of the missing plots on the general computation until this was negligible.[758]

With the Elliott 401, once the data tape had been prepared and checked, the analysis of a typical experiment required one and a half minute for each random variable examined.[759] The output could be presented in the form of a two-way table, if required, and labels could also be introduced to facilitate the research worker in reading the results. But labelling was "expensive

[754] F. Yates et al. (1957), p. 238.
[755] F. Yates et al. (1957), p. 238.
[756] F. Yates et al. (1957), p. 238.
[757] F. Yates (1960b), p. 204.
[758] F. Yates (1960b), p. 204.
[759] F. Yates et al. (1957), p. 246.

both in machine time and orders" and so it was restricted to basic elements such as the ± sign before standard errors and a few keywords to indicate the meaning of the figures in the printout.[760]

According to Frank Yates, besides the economy of time, two were the main advantages brought by computerization in the analysis of replicated experiments. The first was that the computer program was now in charge of minor issues such as confounding and missing plots, while the statistician was just concerned with the interpretation of the results, the second that it became unnecessary to train the human computers in the analysis of variance.

That said digitization in the analysis of agricultural experiments was not without limitations. The programs written by Yates and colleagues dealt only with some experimental designs within the category of replicated experiments. The Elliott 401, due to its technical limitations was unable to cope with some of the other popular experimental arrangements in agriculture. Handling a larger set of experiments was a goal set by Yates for the new Ferranti Orion that the Rothamsted statistics department acquired in the 1960s.[761] The new mainframe, designed to use floating-point arithmetic, was also a decisive advance over the Elliott 401, which could handle only fixed-point arithmetic and thus required programs with scaling factors in order to return consistent numbers at the end of the calculation. With the Orion, also the clerical work involved in the preparation of the input and output reached the autonomy that Yates claimed already with the Elliott 401, but which is denied at that stage by the memories of the assistant staff (Appendix 4.3).

Nevertheless, the replacement of the Elliott 401 required a new programming effort, because order codes changed consistently between the Elliott 401 and the Ferranti Orion. However, Yates hoped that the programs already available in his department could be transposed for the Orion "with relatively little work, certainly far less than that which would be required if the whole program had to be reconsidered *ab initio*".[762]

4.5.2 *Analysis of surveys*

By 1960, when the third edition of Frank Yates' *Sampling Methods for Censuses and Surveys* was published, a final section was added to the book describing how mainframes could be used in the analysis of surveys.[763] Yates explained to his readers how digital computers could ease the task of going through the consistent amount of data collected in a survey and help in the preparation of

[760] F. Yates et al. (1957), pp. 241-242.
[761] F. Yates (1960b), p. 209. For a description of the Orion see RES (1961), pp. 178-180. This mainframe remained in use at Rothamsted until 1972.
[762] F. Yates et al. (1957), p. 246.
[763] F. Yates (1960a), chapter 11.

the summary tables.[764] The book's section outlined also the structure of a general program for survey analysis, later developed in the Rothamsted statistics department.[765]

Unlike the case of replicated experiments, for which computerization had started soon after the acquisition of the Elliott 401, at Rothamsted "[m]ethods for handling survey data [...] developed more slowly, partly because of the limited storage capacity of the computer and partly because card input was not available until early in 1957".[766] The problems at stake in the digitization of the survey analysis were similar to the ones already discussed for replicated experiments: on the one hand the software should be sufficiently general to handle several types of surveys and on the other it had to optimize the performance of the Elliott 401, coming to terms with the technical limitations of the machine.

If in the analysis of experiments, Yates and his co-workers had preferred to write several programs able to cope with different designs, for surveys, after a few programs tailored to deal with specific cases, they had prepared a general program, which was slow, but effective.[767] Due to the limited memory of the Elliott 401, the program was split in two sections, the first for the acquisition of data and the preliminary calculations, the second for making and printing the summary tables.

The instructions related to the two-stage analysis of the surveys were not held in the computer memory at the same time to gain more space for the data.[768] Other solutions intended to optimize the use of the computer memory were the packing of data, that is the allocation of two different types of data in the same position, the omission of space for marginal totals, that is the results gained totalling the cells of a table in relation to one or more variable, and the analysis of data by stratum, that is considering only homogeneous subsets of the original sample.[769]

The first part of the survey program controlled the input of the data. It was expected to use as a standard punched-card input, but the program could be adjusted for paper tape input, if required.[770] Once the data were read, the survey program checked the information looking for reading errors or missing values. After these initials operations, tabulations were prepared for both qualitative and quantitative variables and derived variates, and stored in the computer

[764] Even beyond surveys, the computing methods adopted and described by textbooks of statistics during the 1950s and 1960s presupposed the use of desk calculators, while algorithms and programs for computers could be found only in specialised journals.
[765] F. Yates (1962), p. 276.
[766] H. R. Simpson (1961), p. 219.
[767] RES (1961), p. 180.
[768] F. Yates and H. R. Simpson (1960), p. 136.
[769] F. Yates and H. R. Simpson (1960), p. 138.
[770] H. R. Simpson (1961), p. 221. The use of punched-card input was linked to the goal of generality set for the survey program. Punched-cards, in fact, were a common tool for recording survey data and they allowed, when useful, to link the computer work with the punched-card equipment available in the department.

memory.[771] Some surveys required to collect information at two levels, for instance in the National Cattle Disease Survey general information was collected on herds and detailed information on cases of disease. The two types of information were recorded on distinct cards, but this was no obstacle for the digitization as the Rothamsted survey program allowed to process both types of cards in different reading cycles.[772]

Once the data had been read and the basic tabulations formed and stored in the computer memory, the second part of the program was fed in the Elliott 401 to prepare the final tables and print them, keeping into account in this final stage any necessary scaling or rounding factor. Although a reasonable expectation was to produce with the Elliott 401 an output "easily understandable by the user" and such that it "could [...] be reproduced for publication without further editing", the programming effort and the space occupied in the computer memory by the labelling of tables prevented the Rothamsted statisticians from integrating this facility in their general survey program.[773] Each table printout was just identified by the number assigned to the table by the program.[774]

In the analysis of surveys, computerization meant once more, not the achievement of novel results in statistics, but an economy of time and labour. The analysis of the survey of fertilizer practice, a yearly task of the Rothamsted statistics department, "occupied a team of about three computers for something like half a year", when ten counties were surveyed, while the computer analysis of the already mentioned cattle disease survey, which comprised overall about eleven thousand and five hundred cards, required only twenty-five hours for part one and forty minutes for part two.[775] Moreover, the digitization of surveys, as in the case of experiments, removed the need to train human computers for this task. The duty of the assistant staff was only to feed the cards in the mainframe or to punch the data tape, when paper tape input was used.

When surveys were at stake, the time and labour saving gained with computerization acquired a further meaning. For instance, the choice to digitize the analysis of the fertilizer survey in 1957 was prompted by the collaboration with the Fertilizers' Manufacturers Association. The support of the association allowed on the one hand to enlarge consistently the survey with forty-five districts examined in a single year, but on the other urged Yates to produce the analysis promptly, "if the information provided is to be of real value to the fertilizer industry".[776] The digitization of

[771] In statistics qualitative variables are variables whose values cannot be expressed with a number, such as gender or colour. On the contrary, quantitative variables are described by numerical values.
[772] H. R. Simpson (1961), p. 224-225.
[773] H. R. Simpson (1961), p. 222; F. Yates and H. R. Simpson (1961), p. 22.
[774] F. Yates and H. R. Simpson (1960), p. 139.
[775] 'A discussion on the use of electronic data processing equipment' (1957), p. 306.
[776] 'A discussion on the use of electronic data processing equipment' (1957), p. 306.

surveys, thus, can be also conceived as a fulfilment of the aims of operational research, in which data should not only be reliable, but also promptly available to policy makers.

Computerization allowed also a refinement in the examination of the survey data, improving the initial checking and giving the opportunity to tabulate not only the data directly collected, but also quantities related to them by simple mathematical functions.[777] Digitization increased also the possibility to prepare multiway tables, well beyond the limitations of punched-card equipment.[778]

The use of an autocode in the survey program prepared for the Elliott 401 deserves a comment as well.[779] In the survey program autocode facilities were provided "for specifying the calculation of derived variates, tabulations and subsequent processing".[780] They allowed to program the mainframe using an easier set of instructions than machine language: tabulations and analyses were specified in a simple numerical form, approachable also by research workers without specific knowledge of programming.[781]

The use of autocode facilities in the survey program suggested a similar approach to the analysis of experiments, for which a new more general program was written. The new software used autocode instructions to operate on the tables of experimental data in a similar way as done for the survey program.[782]

The creation of autocodes for the analysis of surveys and experiments can be interpreted as a first attempt at the development of statistical software of very wide applicability and more user-friendly than programming in machine language. The effort led in the late 1960s-early 1970s to the making of GenStat, a statistical language as well as a software package still in use today, whose development started at Rothamsted with John Nelder, the statistician who followed Frank Yates as head of the statistics department.[783] Since the 1970s GenStat offered facilities for the analysis of variance and covariance, multiple regression, multivariate analysis, tabulation and data management. It satisfied thus "most of the [Rothamsted] Statistics Department's computational

777 H. R. Simpson (1961), p. 219.
778 The increased freedom allowed by the Elliott 401 in the manipulation of the survey data recalls the point made by Edward Higgs on the mechanization of the 1911 census of England and Wales (E. Higgs (1996), mentioned also in chapter 1). When large amount of data are involved, technology matters in increasing the amount of information, which can be extracted using statistics.
779 The 'autocode' written by Frank Yates and his colleagues is nowadays usually referred to as assembly code or assembly language and is a symbolic representation of the machine code. By today's standards assembly language is quite intricate, but during the 1950s it was still an improvement if compared to machine language. The creation of high level languages like ALGOL and FORTRAN in the late 1950s and early 1960s superseded autocodes, which were computer-specific, and not general languages. An example of autocode for the Ferranti Mark I is given in D. E. Knuth and L. Trabb Pardo (1980), pp. 227-233. A more general contribution on the relevance of software issues in the history of computing is U. Hashagen et al. (2002).
780 H. R. Simpson (1961), p. 223.
781 F. Yates (1962), p. 276.
782 F. Yates et al. (1963), p. 313.
783 A scientific obituary of J. Nelder is R. Payne and S. Senn (2010). On GenStat see J. C. Gower (1986), pp. 229-232.

needs for routine analysis" and provided a language suitable for "assembling a microlibrary for statistical computations of a less routine nature".[784]

4.5.3 *The estimation of linkage*

The last scientific application of the Elliott 401 that I want to account is the estimation of linkage in human genetics.[785] Unlike the programs written for the analysis of agricultural experiments and surveys, which were employed over the years for the examination of a wide array of data, this was a one-off effort conducted in the Rothamsted statistics department to analyse the pedigree data provided by two workers at the Galton Laboratory, the physicians Sylvia D. Lawler and James H. Renwick.[786]

Since its foundation at the beginning of the twentieth century the Galton Laboratory had been planned as a storehouse of statistical material on human inheritance, and its first two directors, Karl Pearson and Ronald Fisher, had pursued this mission strengthening the statistical and computing features of the laboratory (**chapter 3**). Lionel Penrose, the physician who took over the management of the laboratory after WWII, instead was "unconcerned with the development of mathematical statistics for its own sake" and shifted the emphasis of the laboratory in a medical and biological direction, as proved by his editorship of the *Annals of Eugenics* (later *Annals of Human Genetics*), which went "non-mathematical".[787]

Lawler, a physician who had previously worked with the serologist Robert Race (**chapter 3**) at the Lister Institute, and Renwick, who was completing his PhD under Penrose, were bound to work in a laboratory, in which "[f]or the most part, people sat at tables and desks working with numbers and papers", thus, they could not count on a computing equipment able to match their ambitions, let alone a digital computer that still was a rarity in 1950s Britain. The "courtesy" – or

[784] J. C. Gower (1986), p. 231.
[785] In genetics linkage means the tendency of two genes to be inherited together. This usually happens because the genes are located on the same chromosome, and the chance of inheriting the two genes together is a rough indication of how their positions, i.e. the loci, are closely related. In the 1950s linkage studies were an opportunity for drawing maps of human chromosomes, maps that had already proved valuable in animal and plant genetics. Among the genes used for testing linkage the ones related to blood groups were a particularly efficient marker, because, unlike rare genetic anomalies, they were widespread in the population, had multi-allelic or complex loci, and standard antisera guaranteed the reliability of their identification. On the use of serology in locating genes on chromosomes see L. Hogben (1931), pp. 68-90; on the role and limitations of linkage studies in human genetics during the 1950s see J. H. Renwick (1956) and S. D. Lawler and J. H. Renwick (1959).
[786] RES (1957), p. 172; RES (1958), p. 181.
[787] For the Galton Laboratory under Penrose see D. J. Kevles (1995), pp. 213-222. The first quotation is from D. J. Kevles (1995), p. 214. The second quotation, instead, is taken from a letter written by Frank Yates to Ronald Fisher: "the 'Annals of Eugenics' [...] as you can see from Penrose's letter has gone non-mathematical" (Letter from F. Yates to R. A. Fisher, 22nd June 1950, R. A. Fisher Papers, Barr Smith Library, The University of Adelaide).

more probably the willingness to experiment – of Frank Yates, who made available the mainframe in his department, was for them a stroke of luck.[788]

That is not to say that their analysis would have been impossible without an electronic computer. Several methods, employing hand calculations, were already available for the examination of the pedigree of a single family, but the estimation of linkage in a pedigree formed by several families "when every possible genotype for each individual has to be considered in the light of population gene frequencies" was time consuming and rather impractical to be undertaken by hand.[789] Thus, Lawler and Renwick turned to the Rothamsted mainframe "[i]n the hope of extracting the maximum of information" from their data.[790]

The pedigrees considered by the Galton Laboratory researchers related to families with nail-patella syndrome, a genetic disease that can affect nails, knees, elbows, ilium, eyes and kidneys.[791] Already in 1955 Lawler and Renwick had proved that the genetic locus for this inherited condition was linked to the locus for the ABO blood groups and they wanted to refine their analysis to test whether linkage could also be proved with another blood group locus, P, and with a form of cataract, as suggested by some pedigree data.[792]

The program written for the Elliott 401 by Howard Simpson, the Rothamsted statistician and computer programmer who supported Lawler and Renwick in the analysis, tested the linkage between the ABO and the P blood groups, but only in families of two generations. To examine the pedigree data Simpson adopted the criterion for testing linkage proposed by the geneticist J. B. S. Haldane and the statistician Cedric A. B. Smith and commonly applied by the researchers at the Galton Laboratory.[793] The method of analysis consisted in computing the probability of occurrence in the pedigree in terms of the gene frequencies and of the recombination fraction (i.e. the quantity that measures the linkage between two genetic loci), and in choosing the value of the latter which maximized the probability.[794]

Unlike the programs for the analysis of replicated experiments and surveys mentioned above, a description of the computer program for the Elliott 401, written by Simpson, was not published

788 "I think that Frank Yates never turned anybody away whatever [disciplinary] area they came from", interview with J. C. Gower, September 2010. On the other hand the research workers at Rothamsted Experimental Station were not particularly interested in the mainframe (see Appendix 4.2).
789 S. D. Lawler and J. H. Renwick (1959), p. 148.
790 S. D. Lawler et al. (1958), p. 345.
791 For more details on the nail-patella syndrome and its use in linkage studies see J. H. Renwick (1956), pp. 149-152 and I. McIntosh et al. (2005).
792 S. D. Lawler and J. H. Renwick (1955). S. D. Lawler et al (1958) is the cooperative study undertaken with the Elliott 401.
793 For a detailed description of Haldane and Smith's method see A. W. F. Edwards (2005b), pp. 525-527. Edwards (2005b) is also a comprehensive account of the statistical methods developed for testing linkage before the coming of the digital computer.
794 H. R. Simpson (1958), p. 356.

in a computer or statistics journal, but alongside the genetic analysis of Lawler and colleagues in the *Annals of Human Genetics*. Thus, it was brought to the immediate attention of geneticists interested in the complex linkage analysis, although few of them in 1958 could have access to an electronic computer.

In his contribution Simpson pointed out the problems that tackling linkage on the Elliott 401 had involved. First of all, the necessity to code the phenotype and genotype information in the pedigree in a suitable numerical form. Once the coding of the data had been done, the computer program determined the probabilities for each individual in the pedigree and for each one of the seven recombination fractions examined (0.1; 0.2; 0.25; 0.3; 0.35; 0.4; 0.5).[795]

Time was a critical factor in the execution of the pedigree analysis. As observed by Simpson, in fact,

> In the ABO:P linkage investigation it was found that the time taken to evaluate a probability for a given recombination fraction varied from a few seconds to as much as 20 min.; since seven values of χ [...] were used this meant that for some pedigrees a total time of over 2 hr. was necessary. It was apparent therefore that the analysis of families of three or more generations in general will require a machine with a faster arithmetical unit than the Elliott 401.[796]

Simpson remarked also that "a programme for an electronic computer is merely a mechanical method of producing an answer following set rules" and that the statistical problems related to the estimation of the recombination fraction should be dealt with and solved by the investigator, as the digitization of the linkage analysis did not remove them.[797]

The pedigree analysis with the Elliott 401 was not conclusive: "[t]he results obtained [...] neither prove nor disprove the original suggestion that the four loci (ABO: nail-patella: P: cataract) are linked to each other, but they make it unlikely".[798] Nevertheless, Lawler and Renwick considered the application of a digital computer to genetic research a matter to pursue further and suggested to their fellow physicians that the use of electronic computers could become in time common for testing and estimating linkage.[799]

Lawler went on working "at the interface between genetic research and patient care" and digital computers never became a priority in her career, but Renwick, who had been more involved in the making of the computer program, contributed in person to transform such promise in reality. In 1958-1959, he worked with Victor A. McKusick at the John Hopkins Hospital department of human genetics, where an IBM 704, a computer much more powerful

---

[795] H. R. Simpson (1958), pp. 358-359. The flow diagram of the computer program written by Simpson is reported in the paper (p. 360).
[796] H. R. Simpson (1958), p. 361.
[797] H. R. Simpson (1958), p. 361.
[798] S. D. Lawler et al. (1958), p. 348.
[799] S. D. Lawler and J. H. Renwick (1959), p. 148.

than the Elliott 401, was already in operation and employed for linkage analysis of human pedigrees.[800]

In 1961 Renwick, by then a lecturer in the department of genetics at Glasgow University, presented with Jane Schulze, a former colleague working in McKusick's department, a computer program for the analysis of large pedigrees.[801] The programme was written for the IBM 709 and 7090 computers and allowed "to analyse pedigrees of up to 9 generations and of average complexity".[802] The results of the machine calculations were tested against linkage calculations already available in the literature, and Renwick and Schulze pointed out that "[s]o far [...] the machine has proved to be completely accurate and, on occasion, has brought to light an error in the published hand calculation".[803] Renwick constantly referred to the program written with Schulze and re-used it with adaptations in later publications promoting constantly the use of digital computers in the estimation of linkage.[804] For Renwick, thus, the linkage program developed for the Rothamsted mainframe was just the beginning of a lifetime engagement in the digitization of genetics.

At Rothamsted the collaboration with the Galton Laboratory researchers, instead, was part of a larger effort in the development of computer programs for genetics. In the same years in which the analysis of Lawler and Renwick's pedigree data was developed, the Rothamsted statistics department was involved in the investigation of linkage and inbreeding and the application of Monte Carlo methods to the study of animal populations.[805]

Michael Healy, one of the programmer of the Elliott 401, later to become a medical statistician, contributed also to the preparation of the book *The ABO Blood Groups: Comprehensive Tables and Maps of World Distribution* co-edited by the serologist Arthur E. Mourant.[806] Starting from Mourant's data on the worldwide population, Healy fitted the A_1A_2BO blood-groups frequencies by maximum likelihood (**chapter 3**) preparing a ten-page table for the book.[807] Overall, thus, the Rothamsted statistics department offered to several British geneticists a starting point to experiment with digital computers, despite the shortcomings of the Elliott 401.

---

800 For more information on McKusick and his department at John Hopkins see D. L. Rimoin (2008). On the IBM mainframes mentioned in the section a useful source of technical information is E. W. Pugh (1995). The use of digital computers for linkage studies since the late 1950s is proved by a research proposal written by McKusick and now available among his digitized papers http://profiles.nlm.nih.gov/ps/access/JQBBBW.pdf.

801 A. W. F. Edwards (2005b), p. 528.

802 J. H. Renwick and J. Schulze (1961), E145.

803 J. H. Renwick and J. Schulze (1961), E145.

804 J. H. Renwick and D. Bolling (1967).

805 RES (1957), pp. 171-172; RES (1958), p. 181.

806 A. E. Mourant et al. (1958).

807 "For table II, which contains the A_1A_2BO data, Mr. M. Healy of Rothamsted Experimental Station very kindly arranged for us to have the use of the electronic calculator at that Institute and he himself worked out the calculating programme to give the maximum likelihood solution for gene frequencies" (A. E. Mourant et al. (1958), p. 4).

4.6 *Conclusion*

Since its opening in the 1920s the Rothamsted statistics department was both a statistics and a computing centre and its first two directors, Ronald Fisher and Frank Yates, were particularly interested in efficient systems of computations. After WWII the transformation of the department in a general statistical advisory service for agriculture and biology in Britain and its involvement in sampling surveys for agriculture consistently increased its computing workload and Frank Yates coped with the new activity not only hiring more staff, but also exploring the potentialities of up-to-date computing tools.

Scientific needs devoid of patronage ties, however, could not have brought to Rothamsted a digital computer already in 1954. It was the wartime engagement of Frank Yates in operational research and the connections gained at that stage that promoted the early computerization of his department. William Slater, the ARC secretary, and Patrick Blackett, the ARC scientific advisor and NRDC consultant, who were crucial for the acquisition of the Elliott 401 at Rothamsted, had been involved with Frank Yates in operational research and favoured his project for the exploration of the potentialities of electronic computers in agriculture and biology.

So far, the computerization of the Rothamsted statistics department fits the pattern that Agar has discussed in his comparative study on scientific computerization in the 1950s. There is, however, a peculiarity in the acquisition of the Elliott 401 at Rothamsted. The main argument for the computerization of the Rothamsted statistics department was not to pursue an economy of time and labour in handling the growing amount of experimental data sent to the department for analysis, nor did it mean the dismissal of the computing equipment previously used. Rather, Frank Yates conceived computing tools in statistics as research tools and envisioned the acquisition of the Elliott 401 as an opportunity to tackle new statistical problems in agriculture and biology. His vision was certainly fulfilled, also because, due to the scarcity of mainframes provided of a statistical library, Rothamsted became a centre of attraction for a wide variety of research workers. Over time the Elliott 401 proved to be an asset also for the routine work of the department: by the beginning of the 1960s the analysis of replicated experiments and surveys were partly done with the mainframe freeing the local statisticians from tedious work and turning the assistant staff in the department from human computers to clerks in charge of punching tapes and cards.

However, there is hardly anything revolutionary in the role played by digital computers in the Rothamsted statistics department. The same willingness to experiment with computing tools that prompted the acquisition of the Elliott 401 in the 1950s had brought to the department the Hollerith equipment in the 1940s and since the 1920s and 1930s had promoted the use of desk calculators and the preparation of statistical tables.

For nearly half a century – at least until Yates' retirement and the division of his department in two distinct units – statistics and computing were two sides of the same coin at Rothamsted. Therefore, there is an historical truth in Yates' statement that the early computerization of his department stemmed directly from the Millionaire calculating machine that Ronald Fisher had acquired upon his arrival at the experimental station and from the tradition started by Fisher that "to be a good theoretical statistician one must also compute, and must therefore have the best computing aids".[808] The revolution in statistics advocated by Yates with the introduction of digital computers, if ever it took place, started thirty years before the acquisition of the Elliott 401 and was rooted in the choice to link statistical applications and computing practices in the Rothamsted statistics department.

Yates' revolutionary assessment of the role of digital computers in statistics is also ill suited to describe the time-span of the process and its institutional and disciplinary complexity. Building a library of statistical routines required years and the appointment of new staff with programming skills. Yates himself was aware of his colleagues' reluctance to approach digital computers and even at Rothamsted it could not be avoided an age divide in the use of the mainframe between statisticians of the younger and older generation.[809] Yates and his colleagues working with the computer had also to face technological and disciplinary challenges. They had to cope with the changes undergone by computer hardware and software during the 1950s and 1960s, develop their own autocode for the Elliott 401 and negotiate their position as programmers and statisticians in relation to the raising discipline of computer science, whose practitioners were merely interested in data processing rather than data analysis.[810]

---

808 F. Yates (1966), p. 233.

809 "Initially many statisticians evinced extreme reluctance to attempt to utilize computers. [...] In part this reluctance stems from the feeling that only by actually working over the figures could a statistician elucidate their meaning. [...] Two other reasons for this reluctance to turn to computers which are not so commonly mentioned are, I think, first that many skills which have been acquired are rendered obsolete and new skills have to be learned, and second that considerable disruption of existing computing organization is inevitable." (F. Yates (1962), p. 274)

810 F. Yates (1966), p. 250.

Appendix 4.1: Expansion of the Rothamsted statistics department, 1930s-1960s.

Year	Head	Scientific staff	Voluntary workers	Assistant staff
1938 Last staff list available before WWII. (10 people)	Frank Yates, Sc.D	W. G. Cochran, M.A. J. W. Weil, M.A. (Prof.) R. A. Fisher, honorary consultant	D. A. Boyd, B.A. R. N. P. Luddington, B.A.	Peggy Hubbard Peggy McLarty Gladys A. Smith Peggy Hurt (Secretary)
1945 First staff list available at the conclusion of WWII. (22 people)	Frank Yates, Sc.D	F. J. Ascombe, M.A. D. A. Boyd, Ph.D Rose O. Cashen, Ph.D A. E. Jones, B.A. O. Kempthorne, M.A. Irene Mathison, M.Phil. M. H. Quenouille, B.A. J. W. Weil, M.A.	T. Chinloy, Dip. Agric. C. E. Ford, B. Sc., Ph.D.	Hansi Barrett G. Graylen, Phyllis Hawkins Barbara Hodge Peggy Hurt Avis James Margaret Pellant Mrs V. Roberts Kathleen Simpson Sylvia Smith Alice M. Wright
1950 Staff list after the acquisition in 1948-49 of Hollerith tabulating equipment (38 people)	Frank Yates, Sc.D.	P. R. D. Avis, B.Sc. (Econ.) D. A. Boyd, Ph.D B. M. Church, B.Sc. A. T. Dunn, B.Sc. (Econ.) J. H. A. Dunwoody, B.A. G. V. Dike, M.A. M. J. R. Healy, B.A. G. E. Hodnett, B.Sc. (colonial office) G. M. Jolly, B.Sc. W. J. Lessells, B.Sc. (Econ.) Rowena Lord, B.Sc. Averil Munns, B. Sc. H. D. Patterson, M. Sc. Emily P. Poulton, B. A. D. H. Rees, B. Sc.	Muriel E. Davies, B.Sc. K. Slater, B.Sc. J. K. R. Wood, B.Sc	Patricia M. Barrett Dorothy Browning Doris Court Mrs. W. Dawson Marjorie Fulford Gwendoline Halsey Constance Hunt Mrs. F. Jordan Mrs. G. Lucas Barbara Needham Susan Oliphant Audrey Pincombe Mrs. V. Roberts Freda A. Smith Sylvia Smith Doreen Tucker W. J. Walters Joan Wilkinson Ruth T. Hunt (Secretary)
1954 Staff list in the year of the acquisition of the Elliott 401 (43 people)	Frank Yates, Sc.D.	D. A. Boyd, Ph.D B. M. Church, B.Sc. Sheila Cohen, B.Sc. Muriel E. Davis, B.Sc. J. H. A. Dunwoody, B.A. M. J. R. Healy, B.A. G. E. Hodnett, B.Sc. D. J. Knight (technician for electronic computer) F. B. Leech, B.Agr., M.R. C.V.S. W. J. Lessells, B.Sc. (Econ.) S. Lipton, B.Sc. Averil M. Munns, B.Sc. H. D. Patterson, M.Sc. D. H. Rees, B.Sc. Emily P. Simpson, B.A. Edith Spetch, B.Sc. M. H. Westmacott, B.A. A. W. Whitwell, B.Sc.	P. T. Nelson, B.Sc. (Agric.) R. W. Sharp, B.Sc. P. Sprent, B.Sc.	Doreen Bispham Edith E. Challenger Doris A. Court Winifred L. Dawson Gwendoline Halsey Constance Hunt Florence V. W. Jordan Marie Keasley Gwendoline Lucas Stella Mardle Doreen Mitchell Muriel C. Must Audrey Pincombe Verona A. Roberts Dorothy Slack Freda Smith Berenice Sorrell Katharine Thomas Ann Watkinson Margaret P. Wittingham Ruth Hunt (Secretary)

1958 See Fig. 2 (45 people)	Frank Yates, Sc.D.	Joan Anderson, B.A. D. A. Boyd, Ph.D O. B. Chedzoy, B.Sc. B. M. Church, B. Sc. Muriel E. Davis, B.Sc. J. H. A. Dunwoody, B.A. J. C. Gower, M.A. M. J. R. Healy, B.A. Mary G. Hills, B.A. D. J. Knight (computer technician) F. B. Leech, B.Agr., M.R.C.V.S. W. J. Lessells, B.Sc. (Econ.) S. Lipton, B.Sc. Marjory G. Morris, B.Sc. H. D. Patterson, M.Sc. D. H. Rees, B.Sc. Emily P. Simpson, B.A. H. R. Simpson, B.A. Edith Spetch, B.Sc. Beryl Turner, B.Sc. A. J. Vernon, B.Sc. (Agric.)	S. F. Buck, B.Sc.	Pauline G. Ambrose Betty E. Artiss Doreen Bispham Marion Bridge Edith E. Challenger Joan E. Coiley Monica R. Coleman Doris A. Court Rosemary E. Cumming Winifred L. Dawson Evelyn M. Hawkins Constance Hunt Florence V. W. Jordan Sheila A. Lawrence Hazel P. Oggelsby Mary D. Rainbow Verona A. Roberts Barbara A. Seymour Betty Sparkes Freda Walker Margaret P. Wittingham Ruth Hunt (Secretary)
1965 Staff list in the retirement year of the Elliott 401 (52 people)	Frank Yates, Sc.D.	K. E. Bicknell D. A. Boyd, Ph.D. B. M. Church, B.Sc. J. H. A. Dunwoody, B.A. C. W. Fearne J. C. Gower, M.A. Mary G. Hills, B.A. Winifred A. Johnson, B.Sc. F. B. Leech, B.Agr., M.R.C.V.S. W. J. Lessells, B.Sc. (Econ.) J. A. Lewis, B.A. Bridget I. Lowe, B.Sc. Alison Macfarlane, B.A. A. H. Martin, B. Sc. Marjory G. Morris, B.Sc. H. D. Patterson, M.Sc. D. A. Preece, M.A. D. H. Rees, B.Sc. G. J. S. Ross, B.A. Joyce R. Rossiter, B.Sc. H. R. Simpson, B.A. A. J. Vernon, B.Sc. (Agric.)	N. G. Alvey	Betty E. Artiss Sandra M. Baldwin Jennifer D. Biggs Anne Butterworth Sheila M. Callewaert Dorothy C. Clark Doris A. Court Margaret Cox Elsie Davies Janet I. Elsmere Margaret Green Christine C. M. Helliar Constance Hunt Penelope K. Leech Lucille V. Legg Susan C. Maddy Ena M. Nicholls Sheila O'Reilly Dorothy G. Palk Mary J. Paynter Verona A. Roberts Olive Sparkes Anne Vessey Lucille Y. Ward Pamela Whitfield Margaret P. Wittingham Jane H. Wright Ruth Hunt (Secretary)

Appendix 4.2: Interview with Prof John C. Gower.

Skype interview collected by the author (GP) during September 2010.

GP: *Which kind of education and professional background did you have when you began to work at Rothamsted Experimental Station in 1955?*

JG: In 1953 I completed the Mathematical Tripos at Cambridge University and then I moved to Manchester University, where I gained a diploma in mathematical statistics with specialization in industrial statistics and electronic computing methods. During my time in Manchester I used for the first time a digital computer, a Ferranti Mark I, and worked on stochastic models of epidemics with the statistician Maurice Bartlett. While I was in Manchester I received several job proposals, one for instance came from Elliott Brothers, the company that manufactured the Elliott 401, but eventually, when I finished my degree in 1955, I went to Rothamsted. The salary there was modest, but certainly reasonable, and the work seemed very promising. It was interesting work I was looking for.

GP: *Which were your duties in the department?*

JG: Since 1955, when I started my work at Rothamsted, to the late 1960s my main task was writing computer programs for statistical analysis. When Frank Yates retired and the department was taken over by John Nelder, however, I drifted away from programming to concentrate on the development of statistical methods. I enjoyed programming, but it took up all one's time and I wanted to follow-up also theoretical developments. But in my first years in the department I was certainly one of the people who constantly worked with the department mainframe. I remember also Howard Simpson, Michael Healy, Steve Lipton as keen programmers of the Elliott 401.

GP: *Which kind of programs did you write?*

JG: I wrote programs for the analysis of replicated experiments and collaborated with several research workers who came to the Rothamsted statistics department with data to analyse. For instance, I collaborated with Eric C. R. Reeve, from the Institute of Animal Genetics in Edinburgh, on the effects of linkage and selection in inbreeding. For P. H. Leslie of the Bureau of Animal Population in Oxford I did a Monte Carlo analysis of a population with two competing species. I gave statistical advice and computing assistance in the estimation of gene frequencies to a researcher of the Animal Health Trust, G. C. Ashton.[811] Apart from collaborations with geneticists, I helped also people interested in classification problems, the first of whom was a researcher at the Low Temperature Research Station, Margaret Pleasance, who

[811] RES (1957), p. 171; RES (1958), p. 181; RES (1959), p. 159.

was interested in the classification of bacteria; followed by Grace Waterhouse of the Commonwealth Mycological Institute, who worked on the classification of fungi. Ian Galbraith of the Natural History Museum asked me to classify the Pacific Islands in relation to their bird populations; he was among the first of many clients from the Natural History Museum. Albert Maxwell of the Maudsley Hospital, London, was interested in the classification of psychiatric patients in relation to their symptoms and I helped him as well.[812] On taxonomic problems I also collaborated with researchers at Kew Gardens and a linguist from the University of Uppsala in Sweden, but I cannot remember their names. I did many many things in taxonomy since these early stages of my career.

GP: *Did you collaborate also with researchers working at Rothamsted Experimental Station?*

JG: Yes, a few of them. Although, apart from James Rayner, a crystallographer working in the pedology department, I cannot recall any other staff member at Rothamsted actually concerned with using the computer at that stage. Rayner knew about Fourier analysis, new about computing, and was interested in writing programs to handle the computations that more frequently arose in crystallography, as structure factors and summation of Fourier series. By the late 1950s, when Rayner and I set to work, crystallographic calculation had been done on electronic computers for about a decade, but still we succeeded in publishing a paper describing how these calculations could be handled on a small mainframe like the Elliott 401.[813]

GP: *Do you think that researchers at Rothamsted, due to the availability of a digital computer in the station, were more interested in the new technology or not?*

JG: I feel that the research workers at Rothamsted, on the whole, did not know much about the potentialities of the digital computer in the statistics department. An example explains what I mean. In the entomology department at Rothamsted at one point arrived a visitor from Notre Dame University, in Indiana, United States, who knew my work and how I had used the department mainframe for statistical analysis and it was this researcher who informed his Rothamsted colleagues of what could be done for them on a digital computer. Paradoxically, the information, instead of travelling straight for a couple of a hundred meters, had to go all the way across the Atlantic and back to reach them.

---

812 RES (1961), p. 181.

813 J. C. Gower and J. C. Rayner (1958). Rayner had gained his PhD in crystallography at the University of Oxford working with Dorothy Hogdkin, a pioneer in the crystallography of organic molecules. Hodgkin had resorted to sophisticated computing equipment (at first punched-card equipment, then analogue and digital computers) for the Fourier calculations required in the determination of the more complex structures (J. Agar (2006), p. 884-888). Rayner's interest for computing did not peter out after his experience with the Elliott 401. In the 1970s he collaborated with the soil scientist D. Jenkinson, another member of the Rothamsted staff, in the creation of a computer model able to simulate the carbon cycle in soil. (P. Merchant interviews D. Jenkinson, 29th March 2010, Part 6, British Library).

GP: *I am particularly interested in the classification programs you wrote for the Elliott 401. Can you tell me which was the starting point for them?*

JG: The first work I did in this area came through Margaret Pleasance, the researcher I already mentioned. Pleasance's institution, the Low Temperature Research Station in Cambridge, an institute of the ARC, worked on food preservation and refrigeration. For this reason Pleasance was particularly interested in bacteria and she had read the paper published in 1957 by the microbiologist Peter Sneath on the *Journal of General Microbiology*.[814] Sneath described a method to produce a numerical classification of different strains of bacteria on a digital computer. In a nutshell Sneath's idea was to calculate coefficients of similarity between pairs of bacterial strains and then to group the strains in such a way that, set a level of similarity, each strain in a group had a similarity greater than the one under consideration with at least another member of the same group. The procedure went on until all the strains had been grouped together in decreasing order of similarity. Margaret Pleasance wanted us to do something along the lines described by Sneath and so I got involved writing a computer program for the Elliott 401. After Pleasance, other researchers came with similar problems to solve and so what I did was to generalize the program I had originally written for her.

GP: *Which kind of data did the taxonomists bring to you?*

JG: The data that came in were species and a list of their properties in relation to which the species should be classified. The program I wrote for the Elliott 401 was able to make a hierarchical classification of these data through the use of a general similarity coefficient, which was more comprehensive than the one originally suggested by Sneath. Although the data were of very different types, statisticians are used to recognise similar patterns and general concepts able to measure similarity/dissimilarity between classes could be applied to very different sets of data.

GP: *Was this work published?*

JG: I never wrote a joint paper with the researchers I helped in taxonomy in these early stages of my career. Later, these people wrote their own papers about what it had been done and they were embarrassed because they did not know how to refer to my statistical and computational contribution. Actually I did write a paper in *Biometrics* about the general similarity coefficient I had proposed and then they could refer to this publication, but it was much later, at the beginning of the 1970s.[815]

---

[814] P. H. A. Sneath (1957). The computer used in the preparation of the paper was made available to Sneath by Elliott Brothers, the same company which manufactured the Elliott 401. A first-hand account of Sneath's interest for the use of computer in taxonomy is P. H. A. Sneath (2010). For an historical contribution on the role of digital computers in systematics see J. B. Hagen (2001, 2003).

[815] J. C. Gower (1971).

GP: *How did you engage with research workers? For instance, you had to know something on the species to be classified working as a consultant of taxonomists.*

JG: Even though I have spent large part of my scientific career as a consultant of life scientists and agronomists, I have never consulted textbooks on any specific discipline while engaged in these collaborations. Botany is a passion of mine and so I knew something about taxonomy in advance and this certainly helped, but my real introduction to the problem was the paper written by Peter Sneath and, of course, my scientific resource were the taxonomists themselves, who had the deep knowledge of their subject. They knew what they wanted and I said what I could do for them as a statistician and programmer and we found an agreement starting from our different perspectives. Sometimes not even the research workers had very clear ideas about the best use of their data and so what I could offer was at least a starting point.[816]

[816] Further reflections on the role of the statistician as consultant are in J. C. Gower and R. W. Payne (1987). This contribution focuses on the classification of yeast, a field in which Gower has been engaged for several years.

Appendix 4.3: Interview with Mrs (Lucille) Vera Wiltsher.

Interview collected by the author (GP) at Rothamsted Research on 1st September 2011.

GP: *You began to work in the Rothamsted statistics department in 1961. Which kind of education and professional background did you have when you arrived there?*

VW: I went to college, after I went to school, and I got a GCE in accountancy, so I was familiar with numbers. When they advertised the job they said a liking and aptitude for numerical work was the main thing, that's why I came here.

GP: *Did you work somewhere else before coming to Rothamsted?*

VW: Only temporary jobs. When I arrived at Rothamsted I was just about seventeen.

GP: *Which kind of duties did you have?*

VW: We were employed as data processors, which means we assisted in running the computer. I worked with the Elliott 401 at the beginning. It was not a computer like today, we used tape and we had to learn all about it. You had to punch the tape, and then you put it through the verification program. Another person would find out any mistakes you had made. Also you had to do a lot of hand-work because the output we were getting was not very sophisticated, so you had to do a lot of calculations, standard errors etc. To do all these computations we used the calculating machines that are still in the archive [of the biomathematics and bioinformatics department, now department of computational and systems biology]. Mainly we used Monroe calculating machines. The first time I was here for six and a half years – I left in February 1967 – you had to do a lot of these hand calculations, areas, conversion factors and things like that, and it was very interesting. You had to work with a colleague because everything was double checked. I really liked it. I was busy all the time.

GP: *Did you work also with the Hollerith equipment?*

VW: No, that was a different section to us. When the first two computers arrived one was for running tape [the Elliott 401] and one was for punched-cards [the Elliott 402]. But when I came back in later years, in 1979, we were supposed to be doing the surveys. This work had always been done on cards and so, all of a sudden, I had to learn to work on cards instead of the punched-tape I had used in the 1960s. It was so different from anything I had ever touched before!

GP: *In 1961 the department was housed in Rivers Lodge and the computer in the near-by hut. Where did you work?*

VW: I was in Rivers Lodge to start off with, but our equipment was in the hut, as you said. The computer was there and also the machines for punching the computer tape were there. Rivers

Lodge was basically a house made into an office and it was not possible to have all the machines there because they would have made so much noise!

GP: *With how many people did you interact for your job?*

VW: With my manager and with the colleagues in charge of the data processing. I think there were about six of us. We had computing machines on our desks in the office, but if we needed to punch the tape, we had to go to the hut. In the assistant staff there were young people like me, but also older ones. I remember one who to me was old because she was my mother's age and she came after I had started.

GP: *How much was your salary?*

VW: It was about twelve pounds per week. It does not sound much now, but it was all right then.

GP: *How was your work organized?*

VW: My manager would receive data either from the [Rothamsted] farm or from people who were doing experiments at Rothamsted or in other institutions such as the John Innes Institute and then he would write a set of instructions that went round the data and he gave these to us and we had to insert the data to prepare a roll of tape. Then you had to verify the tape just to make sure that it was all right. And then, after the manager had looked at the data tape and thought it was all right, he would put it in a box with the program tape and it was sent to the computer [Elliott 401]. I used to do overtime on the Elliott 401 so I was actually running it [out of office hours]. You put the program tape in and then you run the tape and you got an output tape, which you had to print out. The printer was just a sort of typewriter.

GP: *What happened when something went wrong?*

VW: Well, it might happen that my manager had made an error in programming and so the output tape would say "Failed", that would be the only way you would know. But we did not often make mistakes.

GP: *During you first stay the statistics department moved also to another building...*

VW: We were in Rivers Lodge for about a year before we moved into the new building of the department. At that stage arrived also the new computer, the Orion. I did not actually run the new computer, but I helped for overtime work – go and get magnetic tapes and so on. I think that with the Orion the number of people working on the data processing increased from six to eight.

GP: *Which kind of interaction did you have with the people in charge of the scientific work of the department?*

VW: We actually did not have anything to do with the scientific side, we just did the punching of the tape and verified it, and the tape was sent off to be run. If there was any mistake, our boss would just say "you need to change that" and you did as requested.

GP: *Did you have any technical knowledge about the computer?*

VW: No, not really. But I loved to work on the 401 because you were producing everything. With the Orion instead you had not to do hand calculations because the computer did all by itself and also the output was much more sophisticated. If you think that I was born in 1944 and the 401 arrived in the 1950s... When I came here for the first time and I saw these two computers [the 401 and 402], they were just like metal cabinets, valves and wires and stuff like that. I just thought it was wonderful.

Archive materials

Barr Smith Library, The University of Adelaide

1) R. A. Fisher Digital Archive

http://digital.library.adelaide.edu.au/dspace/handle/2440/3860

2) R. A. Fisher Papers

The National Archives of the UK

1) The Security services personal files: Frank Yates alias Francis Gray, Ref. KV 2/3082.

2) Ministry of Agriculture and Fisheries and Ministry of Agriculture, Fisheries and Food: Rothamsted Experimental Station, organisation of research statistics, Ref. MAF 117/88.

- Joint Committee of the AIC (England and Wales), the ARC and the Scottish Agricultural Advisory Council, J.C./P.65.

- Review of Hollerith equipment at Rothamsted Experimental Station, F. Yates, 27th December 1950.

3) Ministry of Agriculture and Fisheries and Ministry of Agriculture, Fisheries and Food: Rothamsted Experimental Station, provision of accommodation for statistical department, Ref. MAF 117/219.

- Letter from J. B. Bennett to R. A. Hughes, 28th October 1953.

4) ARC minutes and papers, 123rd Meeting, Ref. MAF 200/19.

- Statistics department, Rothamsted: proposal to acquire electronic calculating machine, office note, ARC 15/54.

5) ARC minutes and papers, 125th Meeting, Ref. MAF 200/21.

- Rothamsted Experimental Station, Report of the Visiting Groups, 1953, ARC 18/54.
- Rothamsted Experimental Station, Report of the Visiting Groups, 1953, Report on statistics department, ARC 18/54 (Appendix J).

Museum of English Rural Life, University of Reading

Letter from F. Yates to E. J. Russell, 19th March 1948, E. J. Russell Correspondence, FR HERT 11/1/1.

Rothamsted Research, Library and Archive

1) Statistics department accommodation, Ref. SITE 1.5.2.

- Letter from E. G. Davy to J. B. F. Cooper, 7th November 1946.

- Statistical department accommodation, summary of memorandum compiled by F. Yates for W. Ogg, 16th May 1950.

- Memorandum on the statistical department accommodation written by F. Yates, 16th May 1950.

- Letter from D. H. Rees to P. A. C. Thorne, 14th February 1958.

2) Rothamsted Staff Council Minutes, Ref. STA 2.1.

-Volume 2, May 1928 – October 1934.

3) Correspondence between R. A. Fisher and F. Yates, 1930s-1960s, Ref. STATS 7.5.

The Royal Society Centre for the History of Science

Letter from F. Yates to P. M. S. Blackett, 15th June 1965, Blackett Papers, Ref. PB/1/33/3/8.

Science Museum London, T/1965-445

- The Elliott- NRDC Computer 401
- Letter from W. S. Elliott to J. C. P. Miller, 10th March 1954
- Speech of Dr. H. R. Calvert, Closing Down Ceremony of the Elliott/NRDC 401, 30th July 1965

Special Collections Department, Iowa State University Library

Report of the ASA task force on statistical computing 1970-1971, American Statistical Association Collection, Box 56 Folder 7.

The Victor A. McKusick Papers, National Library of Medicine (digitized)

Project proposal for "Heritable Disorders of Connective Tissue", ca. 1958-1959 [accessed 15th November 2012, http://profiles.nlm.nih.gov/ps/access/JQBBBW.pdf]

UK National Archive for the History of Computing, Manchester

National Research Development Corporation Papers, Ref. NAHC/NRD.

- Computers, Elliott 401, user enquiries: correspondence 1952-56, NAHC/NRD C15/2.
- Letter from W. K. Slater to Lord Halsbury, 2nd November 1953, NAHC/NRD C10/8.
- Note of the meeting between the NRDC and the Rothamsted statistics department, 12th February 1954, NAHC/NRD C10/8.
- Letter from H. G. Carpenter to D. H. Rees, 16th February 1954, NAHC/NRD C10/8.
- 16th Sub-Committee Report, 23rd February 1954, NAHC/NRD C10/8.
- Letter from D. Hennessey to F. Yates, 24th February 1954, NAHC/NRD C10/8.
- Letter from H. G. Carpenter to D. J. Knight, 21st April 1954, NAHC/NRD C39/1.
- 55th Board, 28th April 1954, NAHC/NRD C10/8.
- Letter from R. A. E. Walker to F. Yates, 4th May 1954, NAHC/NRD C10/8.
- Letter from F. S. Ellis to Messers Creed & Co, 14th June 1954, NAHC/NRD C10/8.
- Letter from H. G. Carpenter to D. H. Marlow, 28th October 1954, NAHC/NRD C10/8.
- Letter from F. S. Ellis to the Secretary, Ministry of Supply, 22nd November 1954, NAHC/NRD C10/8.
- Meeting with Mr Thorne ARC on 5th January 1955, NRDC internal memo, 17th January 1955, NAHC/NRD C10/8.
- Letter from F. Yates to H. H. Brazier, 29th April 1955, NAHC/NRD C15/2.
- Summary of the performance of the 401 at Rothamsted during May 1955, NAHC/NRD C10/8.
- Agreement for hire of the 401 electronic digital computer, 22nd November 1955, NAHC/NRD C10/8.
- The electronic computer at Rothamsted, progress report, 1954-1956, NAHC/NRD C10/8.
- Automatic programming routine, F. Yates and S. Lipton, 27th July 1956, NAHC/NRD C39/3.
- Letter from D. H. Rees to H. J. Crawley, 10th January 1957, NAHC/NRD C10/8.

- Letter from D. Hennessey to P. M. S. Blackett, 19th February 1957, NAHC/NRD C10/8.
- Letter from D. H. Rees to F. S. Ellis, 7th June 1957, NAHC/NRD C10/8.
- Letter from H. R. Calvert to H. J. Crawley, 2nd March 1965, NAHC/NRD C11/3.
- Letter from D. H. Rees to D. Hennessey, 21st July 1965, NAHC/NRD C11/3.

<u>University of East Anglia, S. Zuckerman Papers</u>

- Letter from B. Schafer to S. Zuckerman, 21st September 1940, SZ/TQ/2/2/6
- Letter from S. Zuckerman to R. S. Capon, 24th June 1945, SZ/BBSU/1/22

CONCLUSION

Since the teatime event that brought together on the lawn of Rothamsted Experimental Station in the early 1930s statisticians and research workers, statistical methods, computing tools and information technologies have met a consistent success in agriculture and biology changing in depth the outlook of these disciplines. In my research I have accounted such change in the half century between the 1920s and the 1960s using as a case study analysis of variance and experimental design, and examining the careers of Ronald Fisher and of his lifetime collaborator Frank Yates.

My case studies have involved technologies as diverse as index cards and digital computers, and I have drawn together in a single narrative data archives, number crunching, mathematical reasoning and research practices in the field and in the laboratory. Moving from Rothamsted Experimental Station to the Galton Laboratory and following the careers of Ronald Fisher and Frank Yates, I have assessed on the one hand the development of Fisher's statistical methods for the planning and analysis of experiments and the role of the statisticians as consultants of agronomists and life scientists, and on the other I have examined the role played by computing tools and information technologies in building this collaboration.

I am going now to discuss the general trends and ideas that the four case studies developed in the thesis have suggested in relation to three main points: 1) the spread of statistical methods in agriculture and biology as a pattern for the mathematization of these disciplines; 2) the role of the statisticians as a new expert group; 3) the part played by computing tools and information technologies in statistics applied to experimental research.

I. *Statistical methods and mathematization*

The mathematical physicist Eugene Wigner begins his contribution on 'The unreasonable effectiveness of mathematics in the natural sciences' with the story of two former classmates who are talking about their respective jobs. One of the classmates is a statistician by profession and explains to the other how population trends can be represented through a Gaussian distribution. The statistician's interlocutor is rather sceptical about the feasibility of pairing the natural phenomena and the bell curve and suspects that his friend is mocking him. His suspicion becomes a certainty for him when he realizes that even pi is involved in the Gaussian formula.

The startling ending of the story, in fact, is that: "surely the population has nothing to do with the circumference of the circle"![817]

Wigner's contribution identifies the natural sciences with just one science, physics, and the paper subscribes to the tradition started by Galileo Galilei in the seventeenth century: the natural world is identified with a book written in mathematical language and only the knowledge of such language allows to understand it.[818] Wigner's conclusion is that

> [t]he miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve. We should be grateful for it and hope that it will remain valid in future research and that it will extend, for better or for worse, to our pleasure even though perhaps also to our bafflement, to wide branches of learning.[819]

If Wigner's immediate concern is for an hard science, physics, the questions it raises with its initial tale are relevant also for the disciplines, agriculture and biology, I have examined in my thesis, although I would never use the categories of 'miracle' or 'gift' to account for their mathematization, neither I am convinced, as Wigner, that we cannot understand the phenomenon.

A simple change of perspective from mathematization as an intrinsic property of nature – a point that despite Wigner's examples remains to me unsatisfactory, as human beings have accounted the natural world in non-mathematical terms for thousands of years before Galileo and yet they have been able to predict phenomena – to mathematization as a convenient choice to describe the natural world is enough to solve the issue.

Recognising that mathematization has been a voluntary choice rather than a necessity is not a way to belittle science and its results, but rather an opportunity to make sense of the current outlook of the experimental sciences, acknowledging that it has been shaped by the cultures and values of their investigators, rather than by an improbable god fascinated by mathematics.

In soft sciences, such as agriculture and biology, the intrinsic variability of organisms – plants, animals or human beings – immediately points to the conventional value of mathematical tools and since the early decades of the twentieth century statistical methods became allied of the research workers in domesticating this variability. On the one hand they enabled to examine problems in which several phenomena interacted. That is the case of field experiments in agriculture, in which the seasonal changes of the weather, the inhomogeneity of the soil and the

[817] E. P. Wigner (1960), p. 1.

[818] "Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders in a dark labyrinth". (G. Galilei [English transl.] (1957), pp. 237-238)

[819] E. P. Wigner (1960), p. 14.

intrinsic variations of crops and fertilizers are at stake simultaneously. Statistics avoided the necessity to resort to ideal systems – popular instead in hard sciences like physics – that could be only a poor substitute for the complexity of nature. On the other hand, analysis of variance and experimental design, the statistical methods on which I have focused in my research, were flexible enough to adapt to laboratory investigations in which only one factor at a time was examined.

Moreover, the rules of statistics – whether for randomising an experiment or making tests of significance on the results – were seen as a guarantee of intellectual honesty, besides being a way to standardise procedures in and across disciplines. Inferential statistics was deemed able to transform the results gained in field trials or at the lab bench in something of value beyond the single experiment, moving from the sample to the population. At the bottom of this idea there is what has been called "mechanical objectivity", that is a decision making process achieved following rules and thus writing off in principle any form of personal bias.[820]

Moreover, in both laboratory and field research, as I have explicitly argued in the case of agricultural science at Rothamsted Experimental Station, Fisher's statistical methods offered the opportunity to achieve increased precision in experimental research and to quantify the error of field trials. For an institution like Rothamsted – whose scientific mission encompassed also the promotion of sound agricultural practices – the improved accuracy of the experimental results was an opportunity to give more detailed suggestions to farmers.

The mathematization of agriculture and biology through statistics, however, was a matter of both gains and losses. Constraints had to be imposed on the variability of nature in order to guarantee the applicability of statistical methods. I discuss one such example in the case of serology. For the adoption of Mendelian algebra in dealing with blood groups, and thus for the mathematical treatment of the problem, the acceptance of Ehrlic's receptor theory and of the unit-character concept with the consequent one-to-one correspondence between antibody and antigen were a necessary pre-requisite. As argued by Pauline Mazumdar, "[n]o overlapping specificities, multiple antibodies, or more or less good fit would have supported such calculations".[821]

A crucial question that remains open is whether statistical methods were accepted in agriculture and biology for convenience or for firm belief in their values. Research workers, in fact, became aware of tests of significance and applied them to their data, but often in a cookbook fashion, without really understanding the limitations of these tools and considering as

[820] For the concept of mechanical objectivity see L. Daston and P. Galison (2010) and T. M. Porter (1996).
[821] P. M. H. Mazumdar (1995), p. 303; p. 357. Quotation p. 357.

an unquestionable limit rather than a subjective choice the five per cent threshold for statistical significance that Fisher had suggested in his textbooks and embodied in his statistical tables.

Did research workers appreciate the chi-square and Student's test for their intrinsic value or because these statistical tools could be applied in a repetitive way and even their pedantic use gave back a numerical value of apparent objectivity? Robert Race's resigned acceptance of the principle of maximum likelihood quoted in chapter three seems to suggest the latter explanation rather than the former.[822] In the mathematization of agriculture and biology adoption and understanding did not always go hand in hand, especially when the subtleties of the understanding required a new set of skills, in this case familiarity with numbers and formulae, which did not belong to the standard background of the research worker.

In hindsight it was largely Ronald Fisher who favoured the uncritical acceptance of his mathematical tools. We should not forget that *Statistical Methods for Research Workers* was written without mathematical proofs, suggesting that research workers could content themselves with the results without following the reasoning behind. No mathematician, instead, would have ever accepted Fisher's methods at face value. It might be that Ronald Fisher envisioned a future of daily collaboration between statisticians and research workers and believed that research workers would have had in any case an 'expert' to consult for their own problems, when the situation required the adoption of non-standard methods or the careful application of the standard ones. But the situation developed in a very different way. Fisher's methods diffused much more rapidly in experimental research than the availability of statisticians in research institutions. For decades and certainly until the 1950s, neither the traditional training system for mathematicians, neither the funding institutions that promoted experimental research in Britain were able to provide enough statisticians and to support their appointment in research institutions.

Even when statistics expanded after WWII and more trained statisticians became available, their presence next to research workers could not be taken for granted. In the planning of complex – and therefore costly – experiments as many in agriculture and medicine, statisticians were usually called as consultants, while for experiments done on a much smaller scale, the research worker was often left alone and had to care for both the planning of the trial and the analysis of the results, with the problems described above. Fisher's suggestion for the serologist Robert Race to require a mathematical assistant to help him out with mathematics was wishful thinking rather than a real assessment of Race's possibilities in the late 1940s, and throughout

---

822 Letter from R. R. Race to R. A. Fisher, 1st July 1947, WL, SA/BGU/F4/3/1. Quoted in section 3.7.

Fisher's life the serologist went on struggling with numbers and formulae, and asking advice to Fisher when the matter became too complicated.[823]

II. ***Statisticians as a new expert group***

In the twentieth century probability and statistics have fostered the authority of experts in a wide array of disciplines.[824] The case of operational research mentioned in chapter four is just one example of this general trend. On the other hand, due to their knowledge of formal techniques for extracting information from large collections of data, statisticians themselves have become an expert group. As such, I want now to discuss the mathematical practitioners encountered in my narrative. The idea is not novel. Jon Agar, for instance, has already examined as an expert movement the British statisticians engaged in governmental matters during the nineteenth century and the first decades of the twentieth century.[825]

Agar argues that these statisticians struggled to gain power within the British civil service: their technical expertise was not enough to gain a central statistical office before WWII, neither they were able to secure the acceptance of sophisticated tools, such as random sampling and index numbers, as their political patrons valued methodological rigor less than the immediate usability of the statistics produced.[826]

If Agar describes a "relative failure" of the statisticians within the British civil service before WWII, the case studies I have examined suggest instead that statisticians succeeded in establishing themselves as an expert group in British research on agriculture and biology already in the 1920s and 1930s. Their advisory role touched upon two crucial areas of experimental life, that is planning and analysis of experiments and even though they had no knowledge over the practicalities of experimentation and they were not directly engaged in it, statisticians were nonetheless governing this activity 'at a distance', because requirements, such as randomisation, redefined practices and instruments of experimentation, as it has been evident in my discussion on the Field Plots Committee at Rothamsted Experimental Station.

To claim that this was a success story does not imply that it was a large scale phenomenon. Ronald Fisher's work was influential and several people from research institutions and industries came to learn his statistical methods, but the number of British statisticians involved in

---

823 Letter from R. A. Fisher to R. R. Race, 1st February 1949, WL, SA/BGU/E7/1.
824 G. Gigerenzer et al. (1989), p. 235.
825 J. Agar (2003), chapter 3.
826 J. Agar (2003), pp. 115-116.

experimental research in the first half of the twentieth century was certainly limited. The Rothamsted statistics department counted at most ten people throughout the 1920s and 1930s, and only during WWII and in its aftermath consistently expanded.

In the post-war era statisticians engaged in research benefited from both the increased funding available for science and the connections established in the conflict, especially in operational research, as discussed for the computerization of Frank Yates' department. WWII was a turning point for statistical expertise not only in the case studies I have examined in my thesis. Harry Marks in his investigation of the development of medical experimentation in the U. S. argues as well for the "triumph of statistics after 1950" and for the new role gained by statisticians in medicine in the second half of the century.[827]

As pointed out by Jon Agar, expert movements have to convince outsiders of the legitimacy of their claims.[828] But persuasion was not always possible and even in my narrative that sketches a substantially positive collaboration between statisticians and research workers not everyone could be convinced of the value of analysis of variance and experimental design. In chapter one I have described, for instance, the case of the botanist Edgar Anderson, who came to Rothamsted to learn Fisher's statistical methods, but decided in the end not to apply them to his own research as too unnecessarily complicated. More generally, the historian Joel Hagen has pointed out the difficulties encountered by systematists in accepting Fisher's methods in the study of evolutionary phenomena.

> The towering figure of R.A. Fisher who made seminal contributions both to evolutionary theory and formal statistics provided a particularly attractive model for the new systematists. Yet, although Simpson and other new systematists deeply admired Fisher, they found his approaches problematic on at least two counts. First, even for those with considerable mathematical training Fisher's work seemed esoteric and the repetitive calculations involved in his statistical tests tedious. Second, there was a nagging suspicion that many of the statistical methods developed to analyze experimental data were inappropriate for the problems faced by systematists.[829]

Nevertheless, the power gained by numbers translated for the statisticians engaged in experimental research in new career opportunities within institutions concerned with agriculture and biology, and statisticians eventually gained permanent positions even out of the centres, such as Rothamsted Experimental Station and the Galton Laboratory, where the development of statistical methods had taken place. Already in 1927 the East Malling Research Station, an institute sponsored by the Agricultural Research Council, hired its first statistician T. N. Hoblyn, a former student of Fisher, for designing horticultural experiments and after WWII more

[827] H. Marks (2000), p. 130.
[828] J. Agar (2003), p. 3.
[829] J. Hagen (2003), p. 355.

statistical units were created and sponsored by the Agricultural Research Council in Britain.[830] When no local statistician was available, the research institutions in post-war Britain and the empire could submit any problem linked, even loosely, to statistics in agriculture and biology to Frank Yates' department, which became a general statistical advisory service, as described in chapter four.

Even beyond official appointments, statisticians went on offering informal advice, as in the case of the lifetime collaboration between Ronald Fisher and the serologists. Although unofficial, these collaborations were not less effective in promoting the application of statistical tools and thus in reinforcing the role of the statisticians as an expert group within agriculture and biology.

III. *Making room for computing tools and information technologies*

When Ronald Fisher negotiated his appointment to the chair of genetics at Cambridge University in the midst of WWII, he listed among the priorities the requisition of a few calculating machines from the Galton Laboratory. The department of genetics in Cambridge so far had "never even possessed a slide rule", remarked Fisher, but for him and his two assistants some computing equipment should be provided, otherwise they would have "absolutely sunk for lack of calculating machines".[831]

Fisher's predecessor in Cambridge, Reginald Punnett, had been an eager Mendelian like Fisher himself, but was sceptical about mathematics and in his department computing had never been linked to experimental research. Instead, for Fisher, who was both a statistician and a geneticist, calculating machines were necessary as breeding stocks and laboratory facilities for research. The contrasting attitudes of Fisher and Punnett are quite telling of the change brought by statistical methods in scientific research. Once mathematization had been accepted in agriculture and biology, also number crunching could not be left out, because the latter was integral to the former.

The same can be said of the information technologies that aided in collecting and recording data. Researchers had been more familiar with them as standardized forms and index cards had been employed for long time, and in the case of human genetics researchers had even resorted to the most up-to-date office equipment, as mentioned in relation to the Eugenics Record Office

[830] D. J. Finney and F. Yates (1981), p. 229.
[831] Letter from R. A. Fisher to W. W. C. Topley, 23rd February 1943, R. A. Fisher Papers (digitized), BSL, The University of Adelaide [http://digital.library.adelaide.edu.au/dspace/bitstream/2440/68066/1/1943-02-23.pdf].

started by Charles Davenport at Cold Spring Harbor. Nonetheless, statistics required more cogent rules in the amount of data to be stored, and in their management and conservation.

With the adoption of punched-card equipment and digital computers, moreover, computing tools and information technologies became interdependent. Punched-cards, for instance, were the preferred method for recording the results of surveys, because Hollerith equipment was in the pre-computer era the tool with which large amounts of data could be handled. On the other hand, the availability of data recorded in a certain format influenced the computing equipment chosen for their analysis, as exemplified by one of the external users of the Elliott 401, the statistician David Cox. On behalf of the National Institute of Industrial Psychology, Cox contacted the National Research Development Corporation to purchase computing time on the Elliott 401 because the experimental data collected by the institute had been already stored on punched tape and the Rothamsted mainframe used precisely this input for the data.[832]

Even beyond the institutions and statisticians accounted in my thesis computing tools and information technologies were constantly coupled to statistics in the practice of agricultural and biological research. Computing, in particular, was a relevant share of the workload of the statisticians in applied research and an expensive one, as it required not merely the acquisition of equipment, but also the appointment of a dedicated staff of human computers. George Snedecor's statistics and computing laboratory at Iowa State College has been mentioned at several points and represent a relevant example of this tradition.

Therefore, statisticians were also computers and managers of computing laboratories and in some cases computation became their main task. According to Frank Yates for this reason some of his colleagues were reluctant to turn to digital computers, as "they would feel lost without computations of the kind with which they are familiar to fill their days".[833] As I have mentioned in chapter four, in fact, computerization meant – unlike the adoption of desk calculators and punched-card equipment – a reshaping of the skills of the statistician and the translation of statistical theory from the language of mathematical analysis to the one of algorithms.

The matching of scientific knowledge and new tools described in my thesis is not a peculiarity of applied statistics. Expert groups have often supported the introduction of new technologies, associating the success of their technical knowledge to the use of these instruments.[834] In particular, the relation of both Ronald Fisher and Frank Yates with computing tools was not merely utilitarian. Fisher and Yates considered computing equipment a research tool in statistics

832 Computers, Elliott 401, user enquiries: correspondence 1952-56, NAHC, Manchester, NAHC/NRD C15/2.
833 F. Yates (1962), p. 274.
834 J. Agar (2003), p. 3.

and were directly engaged in making statistical tables and, in the case of Frank Yates, also computer programs for statistical analysis. Technologies, thus, deserve a place in the history of statistics, not only for the material contributions they gave to the working life of the statisticians, but also for their impact on the development of statistical theory. From the age of slide rules and mathematical tables to the one of the digital computer, computing tools have constantly challenged statisticians to work out more efficient numerical methods, have offered them new possibilities with increased computing potentialities, have frustrated their ambitions when theoretical aspirations and calculating power did not match. And the adoption of new computing equipment has entailed reforms of statistical textbooks and labour routines.

If computing and information technologies were the tools of the trade for statisticians, the research workers, who usually received only a basic mathematical training, were less at ease in dealing with these technologies and had available more basic equipment. In the 1920s and 1930s they mainly relied on slide rules and mathematical tables, while calculating machines were still rare among biologists and agronomists. Moreover, unlike statisticians, research workers were not usually interested in discussing and comparing the efficiency of different computing instruments. At Rothamsted Experimental Station even the arrival of the mainframe found a cold acceptance among research workers and except James Rayner – by training not an agronomist or a biologist, but a crystallographer with strong interests in mathematics – they did not approach programming during the 1950s and 1960s.

Nonetheless, research workers were compelled to resort to number crunching and data management in the application of statistical methods when there was no statistician to shoulder the task, and room for computing tools and information technologies had to be made also on the lab bench of the research worker and not only on the statistician's desk.

IV. *The bottom line*

There are many stories intertwined in my narrative: the scientific careers of Ronald Fisher and Frank Yates, the constitution of the statistics department at Rothamsted Experimental Station and its computerization after WWII, the reshaping of research practices in agricultural experiments, the making of computing instruments for the promotion of analysis of variance and experimental design, the use of statistical methods in serology and the exploration of human biological diversity in Britain through the use of numbers and blood groups. However, only two are the main topics around which the stories develop, on the one hand the adoption of statistical methods in experimental research and their impact on the research practices and instruments adopted in agriculture and biology, and on the other the role of computing tools and information technologies in this process.

The lesson I learnt from the case studies examined is that the mathematization of agriculture and biology through statistics was a two way process in which both statisticians and research workers took part. Fisher's statistical methods were developed not only at the statistician's desk, but also in the experimental fields and at the lab bench, in cooperation with ecologists, plant physiologists and laboratory workers and their application required both the technical knowledge of the experimentalists and the mathematical competence of the statisticians. From this point of view quite revealing is the case of serology, in which the technical complexity of the laboratory work required the constant collaboration of the statistician and the research worker because the mathematical understanding of the data could not be disjoint from a knowledge of how they had been obtained.

In the two way process that brought statistics into agriculture and biology computing tools and information technologies played a relevant role. Number crunching and data management were a natural counterpart of the mathematization of these disciplines and even research workers had to confront themselves with the new tasks, when no organized statistical service was available. Statistics required a new outlook in collecting information – as argued for instance in the case of the transfusion records during the ABO survey – and made more and more valuable the historical series of data that allowed to examine natural phenomena over time. Calculating machines, mathematical tables, slide rules, Hollerith equipment, mainframes on the other hand were research tools for the statistician and influential in the promotion of statistical methods. Thus, an uninterrupted partnership with computing instruments marks the history of applied statistics in the period here examined. In this alliance continuity between old and new computing equipment was never lost making Fisher's Millionaire calculator in principle, if not in practice, a parent technology of Yates' Elliott 401.

I claim thus the relevance of an history of statistics written bottom-up, beginning with the instruments and not with the theoretical achievements, discussing the developments of statistical science along with the computing tools and information technologies that were instrumental for such developments, and presenting the mathematization of agriculture and biology through statistics as a two voice invention in which took part both statisticians and research workers.

BIBLIOGRAPHY

AGAR, Jon (2003) *The Government Machine: A Revolutionary History of the Computer*. Cambridge, Massachusetts, and London, England: the MIT Press.

AGAR, Jon (2006) What difference did computers make? *Social Studies of Science*, Vol. 36, No. 6: 869-907.

AGAR, Jon and BALMER, Brian (1998) British scientists and the Cold War: the defence research policy committee and information networks, 1947-1963. *Historical Studies in the Physical and Biological Sciences*, Vol. 28, No. 2: 209-252.

ALDRICH, John (2007) 'Variance, analysis of variance and ANOVA'. In: Earliest known uses of some of the words of mathematics [accessed 5th July 2012, http://jeff560.tripod.com/v.html].

ALLEN, Michael T. and HECHT, Gabrielle (2001) *Technologies of Power: Essays in Honour of Thomas Parke Hughes and Agatha Chipley Hughes*. Cambridge, Massachusetts and London, England: The MIT Press.

ANONYMOUS (1931) Mr W. D. Christmas. *Nature*, Vol. 127, No. 3195: 136.

ANONYMOUS (1940) *Science in War*. Penguin Special.

ANONYMOUS (1951) Anthropology and Serology. *British Medical Journal*, Vol. 1, No. 4712: 942-943.

ASHTON, E. H., HEALY, Michael J. R. and LIPTON, Steve (1957) The descriptive use of discriminant functions in physical anthropology. *Proceedings of the Royal Society of London Series B*, Vol. 146, No. 925: 552-572.

BANCROFT, T. A. (1982) Roots of the Iowa State University statistical center: 1914-1950. *Iowa State Journal of Research*, Vol. 57, No. 1: 3-10.

BARKAN, Elazar (1992) *The Retreat of Scientific Racism. Changing Concepts of Race in Britain and the United States Between the World Wars*. Cambridge: Cambridge University Press.

BAXANDALL, D. [revised by PUGH, Jane] (1975) *Calculating Machines and Instruments*. London: Science Museum.

BENNETT, John H. (ed.) (1971) *Collected Papers of R. A. Fisher*, Volumes 1-5. South Australia: The University of Adelaide.

BENNETT, John H. (ed.) (1990) *Statistical Inference and Analysis: Selected Correspondence of R. A. Fisher*. Oxford: Clarendon Press.

BERNSTEIN, Felix (1966a) [English translation of 'Ergebnisse einer biostatistichen zusammenfassenden Betrachtung ber die Erbstrukturen der Menschen' (1924) *Klinische Wochenscrift*, Vol. 3, No. 33: 1495-1497] Results of a biostatistical, summarizing study of hereditary human blood structures. In: *Selected Contributions to the Literature of Blood Groups and Immunology. Volume 1: The ABO System* (1966) Blood Transfusion Division, US Army Medical Research Laboratory, Fort Knox, Kentucky; pp. 83-89.

BERNSTEIN, Felix (1966b) [English translation of 'Zusammenfassende Betrachtungen über die erblichen Blutstrukturen des Menschen', *Zeitscrift für Induktive Abstammungs- und Vererbungslehre* (1925) Vol. 37, No. 1: 237-270] A summary of considerations concerning the genetic blood structures in man. In: *Selected Contributions to the Literature of Blood Groups and Immunology. Volume 1: The ABO System* (1966) Blood Transfusion Division, US Army Medical Research Laboratory, Fort Knox, Kentucky; pp. 91-138.

BERNSTEIN, Felix (1930) Fortgesetze Untersuchungen aus der Theorie der Blutgruppen. *Zeitscrift für Induktive Abstammungs- und Vererbungslehre*, Vol. 56, No. 1: 233-273.

BILLING, Edward (1939) Racial origins from blood groupings. *The British Medical Journal*, Vol. 2, No. 4108: 712.

BIRD, G. W. G. (1992) Percy Lane Oliver OBE (1878-1944): founder of the first voluntary blood donor panel. *Transfusion Medicine*, Vol. 2, No. 2: 159-160.

BLISS, Chester I. (1943) Review: Statistical Tables. *Science*, Vol. 98, No. 2546: 346-347.

BOWKER, Geoffrey C. and STAR, Susan Leigh (2000) *Sorting Things Out: Classification and its Consequences*. Cambridge, Massachusetts, and London, England: The MIT Press.

BOX, George E. P. (1976) Science and statistics. *Journal of the American Statistical Association*, Vol. 71, No. 356: 791-799.

BOYD, William C. (1939) Blood groups. *Tabulae Biologicae*, Vol. 17: 134-240.

BRASSLEY, Paul (1995) Agricultural research in Britain, 1850-1914: failure, success and development. *Annals of Science*, Vol. 52, No. 5: 465-480.

BRASSLEY, Paul (2004) Hall, Sir (Alfred) Daniel (1864–1942). *Oxford Dictionary of National Biography*. Oxford: Oxford University Press; online edition January 2008, [accessed 21st May 2012, http://www.oxforddnb.com/view/article/33647].

BRIGGS, G. E., CANNY, M. J. P., BARKER, J. (1961) Ernest John Maskell, 1895-1958. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 7: 161-171.

BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (1874) *Report of the Forty-Third Meeting Held at Bradford in September 1873*. London: John Murray.

BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (1900) *Report of the Sixty-Ninth Meeting Held at Dover in September 1899*. London: John Murray.

BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (1948) Mathematical Tables: report of Committee. *The Advancement of Science*, No. 5, Vol. 17: 67-72.

BRITISH RED CROSS SOCIETY (1933) *Blood Transfusion Service Quarterly Circular*, No. 1.

BROWNING, T. O. (1981) James Davidson (1885-1945). *Australian Dictionary of Biography* [accessed 17th July 2012, http://adb.anu.edu.au/biography/davidson-james-5900/text10047].

CAJORI, Florian (1994) *A History of the Logarithmic Slide Rule and Allied Instruments* [first edition 1909]. Mendham, New Jersey: Astragal Press.

CAMPBELL-KELLY, Martin (1989) *ICL: A Business and Technical History*. Oxford: Clarendon Press.

CAMPBELL-KELLY, Martin (1996) Information technology and organizational change in the British census, 1801-1911. *Information Systems Research*, Vol. 7, No. 1: 22-36.

CAMPBELL-KELLY, Martin, CROARKEN, Mary, FLOOD, Raymond and ROBSON, Eleanor (2003) *The History of Mathematical Tables: From Sumer to Spreadsheets*. Oxford: Oxford University Press.

CASSATA, Francesco (2006) *Il fascismo razionale: Corrado Gini tra scienza e politica*. Roma: Carocci.

DE CHADAREVIAN, Soraya (2002) *Designs for Life: Molecular Biology after World War II*. Cambridge: Cambridge University Press.

CHADWICK, James (1961) Frederick John Marrian Stratton, 1881-1960. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 7: 280-293.

CHARNLEY, Berris (2011) *Agricultural Science, Plant Breeding and the Emergence of a Mendelian System in Britain, 1880-1930*. Doctoral Dissertation, University of Leeds.

CLARK, Ronald (1984) *The Life and Work of J. B. S. Haldane*. Oxford: Oxford University Press.

CLARKE, Ernest [rev. JOHNSTON, A. E.] (2004) Gilbert, Sir Joseph Henry (1817–1901). *Oxford Dictionary of National Biography*. Oxford: Oxford University Press; online edition, May 2006 [accessed 30th June 2012, http://www.oxforddnb.com/view/article/33399].

COCHRAN, William G. (1976) Early development of techniques in comparative experimentation. In: Owen, D. B. (ed.) *On the History of Statistics and Probability*, New York: Marcel Dekker Inc.

COCHRAN, William G. and COX, Gertrude M. (1957) *Experimental Designs* (second edition). New York and London: Wiley.

COOKE, George W. (ed.) (1981) *Agricultural Research 1931-1981*. Published by the Agricultural Research Council.

CRAWLEY, H. John (1993) NRDC's role in the early British computer industry. *Resurrection*, No. 8 [accessed 27th August 2012, http://www.cs.man.ac.uk/CCS/res/res08.htm#h].

CROARKEN, Mary (1990) *Early Scientific Computing in Britain*. Oxford: Clarendon Press.

CROARKEN, Mary (2000) L. J. Comrie: a forgotten figure in the history of numerical computation. *Mathematics Today*, Vol. 36, No. 4: 114-118.

CROARKEN, Mary (2003) Table making by committee: British table makers, 1871-1965. In: Campbell-Kelly, M. et al. (eds.) *The History of Mathematical Tables: From Sumer to Spreadsheets*. Oxford: Oxford University Press; pp. 235-263.

CROARKEN, Mary and CAMPBELL-KELLY, Martin (2000) Beautiful numbers: the rise and decline of the British Association Mathematical Tables Committee, 1871-1965. *IEEE Annals of the History of Computing*, Vol. 22, No. 4: 44-61.

CROWTHER, Edward M. and YATES, Frank (1941) Fertilizer policy in war-time. *Empire Journal of Experimental Agriculture*, Vol. 9, No. 34: 77-97.

DANIELS, Geoff and BROMILOW, Imelda (2010) *Essential Guide to Blood Groups*. Chichester: Wiley-Blackwell.

DASTON, Lorraine (2008) Analogies and migration of scientific ideas: the strange career of the normal curve. *Nova Acta Leopoldina*, Vol. 97, No. 358: 169-185.

DASTON, Lorraine and GALISON, Peter (2010) *Objectivity*. New York: Zone Books.

DAVIDSON, James (1922) Biological studies of *Aphis rumicis*: reproduction on varieties of *Vicia faba*. *Annals of Applied Biology*, Vol. 9, No. 2: 135-142.

DEAN, H. R. (1946) George Lees Taylor, born 26th June 1897- died 9th March 1945. *Journal of Pathology and Bacteriology*, Vol. 58, No. 3: 593-597.

DOBSON, Aileen M. and IKIN, Elizabeth W. (1946) The ABO blood groups in the United Kingdom; frequencies based on a very large sample. *Journal of Pathology*, Vol. 58, No. 2: 221-227.

DODSWORTH, Helen (1996) Blood Transfusion Services in the UK. *Journal of the Royal College of Physicians of London*, Vol. 30, No. 5, 457-464.

DYKE, George (1995) Frank Yates. *Journal of the Royal Statistical Society Series A*, Vol. 158, No. 2: 333-338.

EDEN, Thomas (1935) The development of field experiments in agricultural research. *Tropical Agriculturist*, Vol. 84: 63-69; 131-149; 188-195.

EDEN, Thomas and MASKELL, Ernest J. (1928) The influence of soil heterogeneity on the growth and yield of successive crops. *Journal of Agricultural Science*, Vol. 18, No. 2: 163-185.

EDGERTON, David (2011) *Britain's War Machine: Weapons, Resources and Experts in the Second World War*. London: Allen Lane.

EDWARDS, Anthony W. F. (1976) Fiducial probability. *Journal of the Royal Statistical Society D*, Vol. 25, No. 1: 15-35.

EDWARDS, Anthony W. F. (2003) R. A. Fisher – twice professor of genetics: London and Cambridge, or 'a fairly well-known geneticist'. *The Statistician*, Vol. 52, No. 3: 311-318.

EDWARDS, Anthony W. F. (2005a) R. A. Fisher, *Statistical Methods for Research Workers*, first edition (1925). In: Grattan-Guinness, Ivor (ed.) *Landmark Writings in Western Mathematics*. Amsterdam: Elsevier; pp. 856-870.

EDWARDS, Anthony W. F. (2005b) Linkage methods in human genetics before the computer. *Human Genetics*, Vol. 118, No. 3: 515-530.

EDWARDS, Anthony W. F. (2007) R. A. Fisher's 1943 unravelling of the Rhesus blood-group system. *Genetics*, Vol. 175, No. 2: 471-476.

EDWARDS, J. H. and BUCKLE, Veronica J. (2004) Lawler, Sylvia Dorothy (1922-1996). *Oxford Dictionary of National Biography*, Oxford University Press [accessed 15th November 2012, http://www.oxforddnb.com/view/article/60757].

F. G. H. (1931) Thomas Barlow Wood, 1869-1929. *Proceedings of the Royal Society of London Series B*, Vol. 108, No. 759: *i-iii*.

FIENBERG, Stephen E. and HINCKLEY, David V. (1980) *R. A. Fisher: An Appreciation*. Berlin, Heidelberg: Springer-Verlag.

FINE, Paul (1994) Obituary: Professor James H. Renwick. *The Independent*, 13th October 1994.

FINNEY, David J. (1995) Frank Yates, 12th May 1902 – 17th June 1994. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 41: 555-573.

FINNEY, David J. and YATES, Frank (1981) Statistics and computing in agricultural research. In: Cooke, G. W. (ed) *Agricultural Research 1931-1981*. Published by the Agricultural Research Council; pp. 219-236.

FISHER, Ronald A. (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, Vol. 52: 399-433. [CP 9]

FISHER, Ronald A. (1921) Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk. *Journal of Agricultural Science*, Vol. 11, No. 2: 107-135. [CP 15]

FISHER, Ronald A. (1922a) On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, Vol. 222, No. 594-604: 309-368. [CP 18]

FISHER, Ronald A. (1922b) Appendix: Statistical considerations involved in Tables I and II of the above paper. *Annals of Applied Biology*, Vol. 9, No. 2: 142-145. [CP 23]

FISHER, Ronald A. (1925a) *Statistical Methods for Research Workers* (1st edition). Edinburgh: Oliver and Boyd. [accessed 6th June 2012, http://psy.ed.asu.edu/~classics/Fisher/Methods/index.htm]

FISHER, Ronald A. (1925b) Applications of Student's distribution. *Metron*, Vol. 5, No. 3: 90-104. [CP 43]

FISHER, Ronald A. (1925c) Expansion of Student's integral in powers of *n*-1. *Metron*, Vol. 5, No. 3: 109-120. [CP 44]

FISHER, Ronald A. (1930) *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press. [accessed 7th February 2012,

http://ia700301.us.archive.org/29/items/geneticaltheoryo031631mbp/geneticaltheoryo031631mbp.pdf].

FISHER, Ronald A. (1933) Review of *Tables for Statisticians and Biometricians*, edited by K. Pearson. *Nature*, Vol. 131, No. 3321: 893-894. [Not listed in the CP]

FISHER, Ronald A (1934) The contributions of Rothamsted to the development of the science of statistics. In: Rothamsted Experimental Station, *Report 1933*, St. Albans: Gibbs and Bamforth; pp. 43-50. [CP 103]

FISHER, Ronald A. (1935) The logic of inductive inference. *Journal of the Royal Statistical Society*, Vol. 98, No. 1: 39-54. [CP 124]

FISHER, Ronald A. (1937) The relation between variability and abundance shown by measurements of the eggs of British nesting birds. *Proceedings of the Royal Society B*, Vol. 122, No. 826: 1-26. [CP 153]

FISHER, Ronald A. (1939a) "Student", *Annals of Eugenics*, Vol. 9, No. 1: 1-9. [CP 165]

FISHER, Ronald A. (1939b) The Galton Laboratory. *Science*, Vol. 90, No. 2341: 436 [originally printed in *The Times* (London), 29th September 1939]. [accessed 5th February 2012,

http://digital.library.adelaide.edu.au/dspace/bitstream/2440/15130/1/172.pdf] [CP 172]

FISHER, Ronald A. (1946) *Statistical Methods for Research Workers* (10th edition). Edinburgh: Oliver and Boyd.

FISHER, Ronald A. (1947) *The Design of Experiments* (4th edition). Edinburgh: Oliver and Boyd.

FISHER, Ronald A. (1950) Gene frequencies in a cline determined by selection and diffusion. *Biometrics* Vol. 6, No. 4: 353-361. [CP 238]

FISHER, Ronald A. and FRASER ROBERTS, John A. (1943) A sex difference in blood-group frequencies. *Nature*, Vol. 151, No. 3840: 640-641. [CP 197]

FISHER, Ronald A. and MACKENZIE, Winifred A. (1922) The correlation of weekly rainfall. *Journal of the Royal Meteorological Society*, Vol. 48, No. 203: 234-245. [CP 21]

FISHER, Ronald A. and MACKENZIE, Winifred A. (1923) Studies in crop variation II. The manurial response of different potato varieties. *Journal of Agricultural Science*, Vol. 13, No. 3: 311-320. [CP 32]

FISHER, Ronald A. and TAYLOR, George L. (1939) Blood groups in Great Britain. *The British Medical Journal*, Vol. 2, No. 4111: 826. [CP 169]

FISHER, Ronald A. and TAYLOR, George L. (1940) Scandinavian influence in Scottish ethnology. *Nature*, Vol. 145: 590. [CP 179]

FISHER, Ronald A., THORNTON, Henry G. and MACKENZIE, Winifred A. (1922) The accuracy of the plating method of estimating the density of bacterial populations. *Annals of Applied Biology*, Vol. 9, No. 3-4: 325-359. [CP 22]

FISHER, Ronald A. and VAUGHAN, Janet (1939) Surnames and blood groups. *Nature*, Vol. 144, No. 3660: 1047-8. [CP 171]

FISHER, Ronald A. and YATES, Frank (1963) *Statistical Tables for Biological, Agricultural and Medical Research* (6th edition). Edinburgh: Oliver and Boyd [accessed 20th January 2013, http://digital.library.adelaide.edu.au/dspace/bitstream/2440/10701/1/stat_tab.pdf].

FISHER BOX, Joan (1978) *R. A. Fisher: the Life of a Scientist.* New York: John Wiley and Sons.

FISHER BOX, Joan (1987) Guinness, Gosset, Fisher and small samples. *Statistical Science*, Vol. 2, No. 1: 45-52.

FLETCHER, A., Miller, J. C. P., ROSENHEAD, L. and COMRIE, L. J. (1962) *An Index of Mathematical Tables* (2nd edition). Published for the Scientific Computing Service Ltd. Oxford: Blackwell Scientific Publications Ltd.

FORTUN, Michael and SCHWEBER, Silvan S. (1993) Scientists and the legacy of World War II: the case of operations research (OR). *Social Studies of Science*, Vol. 23, No. 4: 595-642.

FRASER ROBERTS, John A. (1942a) Surnames and blood groups, with a note on a probable remarkable difference between North and South Wales. *Nature*, Vol. 149, No. 3770: 138.

FRASER ROBERTS, John A. (1942b) Blood group frequencies in North Wales. *Annals of Eugenics*, Vol. 11, No. 1: 260-271.

FRASER ROBERTS, John A. (1947) The frequencies of the ABO blood-groups in South-Western England. *Annals of Eugenics*, Vol. 14, No. 1: 109-116.

GALILEI, Galileo (1957) *Discoveries and Opinions of Galileo* [translated with an Introduction and Notes by Stillman Drake]. New York: Doubleday Anchor Books.

GANNETT, Lisa and GRIESEMER, James R. (2004) The ABO blood groups: mapping the history and geography of genes in *Homo Sapiens*. In: Rheinberger, Hans-Jörg and Gaudillière, Jean-Paul (eds) *Classical Genetics Research and Its Legacy: The Mapping Cultures of Twentieth Century Genetics*. London and New York: Routledge, pp. 57-87.

GARCÍA-SANCHO, Miguel (2012) *Biology, Computing, and the History of Molecular Sequencing: From Proteins to DNA, 1945-2000*. Basingstoke: Palgrave Macmillan.

GARNER, H. V. (1962) Field Plots Committee. In: Rothamsted Experimental Station, *Report 1961*, Harpenden: Rothamsted Experimental Station; pp. 180-183.

GARRATTY, George (1996) Association of blood groups and disease: do blood groups antigens and antibodies have a biological role? *History and Philosophy of the Life Sciences*, Vol. 18, No. 3: 321-344.

GEMELLI, Giuliana, PICARD, Jean-François and SCHENEIDER, William H. (eds) (1999) *Managing Medical Research in Europe: The Role of the Rockefeller Foundation (1920s-1950s)*. Bologna: CLUEB.

GIGERENZER, Gerd, SWIJTINK, Zeno, PORTER, Theodore, DASTON, Lorraine, BEATTY, John, KRÜGER, Lorenz (1989) *The Empire of Chance*. Cambridge, Massachusetts: Cambridge University Press.

GINI, Corrado (1957) Felix Bernstein, 1878-1956 [in Italian]. *Review of the International Statistical Institute*, Vol. 25, No. 1: 185-186.

GOLD COAST SURVEY (1929) *Tables for Use in the Department* (1st edition). Accra: Government Printing Office.

GOSSET, William S. (1962) *Letters from W. S. Gosset to R. A. Fisher, 1915-1936*. Dublin: Printed for private circulation by Arthur Guinness and Sons.

GOWER, John C. (1960) The analysis of experiments on the Rothamsted computer. *Manchester Statistical Society Group Meetings*, Session 1959-60: 40-49.

GOWER, John C. (1971) A general coefficient of similarity and some of its properties. *Biometrics*, Vol. 27, No. 4: 857-871.

GOWER, John C. (1986) The development of statistical computing at Rothamsted. In: Rothamsted Experimental Station, *Report 1985 Part II*; pp. 221-235.

GOWER, John C. and PAYNE, Roger W. (1987) On identifying yeasts and related problems. In: Hand, D. J. and Everitt B. S. (eds) *The Statistical Consultant in Action*, Cambridge: Cambridge University Press; pp. 108-120.

GOWER, John C. and RAYNER, James C. (1958) Crystallographic programmes for a computer. *British Journal of Applied Physics*, Vol. 9, No. 11: 446-447.

GREENBERG, Bernard G. (1983) Joseph Oscar Irwin, 1898-1982. *Biometrics*, Vol. 39, No. 2: 527-528.

GRIDGEMAN, N. T. (1959) The lady tasting tea and allied topics. *Journal of the American Statistical Association*, Vol. 54, No. 288: 776-783.

GRIER, David A. (2000) Agricultural computing and the context for John Atanasoff. *IEEE Annals of the History of Computing*, Vol. 22, No. 1: 48-61.

GRIER, David A. (2007) *When Computers Were Human*. Princeton: Princeton University Press.

HACKEL, Emanuel (1984) R. R. Race Obituary. *News Briefs (of the) American Association of Blood Banks*: 1.

HACKING, Ian (1976) *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.

HAGEN, Joel B. (2001) The introduction of computers into systematics research in the United States during the 1960s. *Studies in the History and Philosophy of Biological and Biomedical Sciences*, Vol. 32, No. 2: 291-314.

HAGEN, Joel B. (2003) The statistical frame of mind in systematic biology from 'Quantitative Zoology to Biometry'. *Journal of the History of Biology*, Vol. 36, No. 2: 353-384.

HALD, Anders (1998) *A History of Mathematical Statistics from 1750 to 1930*. New York: John Wiley and Sons.

HALD, Anders (2007) *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935*. New York: Springer.

HALL, Alfred D. (1905) *The Book of the Rothamsted Experiments*. London: J. Murray [accessed 26th May 2011, http://www.archive.org/stream/bookofrothamsted00hallrich#page/30/mode/2up].

HALL, Nancy S. (undated) *Did Fisher's voluntary workers at Rothamsted make a difference in the spread of statistical techniques in agriculture?* Unpublished talk.

HALL, Nancy S. (2002) *R. A. Fisher and Randomized Experimental Design*. Doctoral Dissertation, University of Maryland, College Park.

HALL, Nancy S. (2007) R. A. Fisher and his advocacy of randomization. *Journal of the History of Biology*, Vol. 40, No. 2: 295-325.

HANLEY, Francis (1998) *The Honour is Due: a Personal Memoir of The Blood Transfusion Service now Known as The Greater London Red Cross Blood Transfusion Service, 1921-1986*. Surbiton: JRP.

HART, Edwin W. (1943) An analysis of the blood group composition of a population in Northern Ireland. *Annals of Eugenics*, Vol. 12, No. 1: 89-101.

HARTLEY, Herman O. (1976) The impact of computers on statistics. In: Owen, D. B. (ed) *On the History of Statistics and Probability*, New York: Marcel Dekker; pp. 421-442.

HASHAGEN, Ulf, KEIL-SLAWICK, Reinhard and NORBERG, Arthur (eds) (2002) *History of Computing: Software Issues*. Berlin: Springer-Verlag.

HEALY, Michael J. R. (1995) Frank Yates, 1902-1994: The work of a statistician. *International Statistical Review*, Vol. 63, No. 3: 271-288.

HEIDE, Lars (2009) *Punched-Card Systems and the Early Information Explosion, 1880-1945*. Baltimore: the John Hopkins University Press.

HENDRY, John (1990) *Innovating for Failure: Government Policy and the Early British Computer Industry*. Cambridge, Massachusetts, and London, England: The MIT Press.

HIGGS, Edward (1996) The statistical Big Bang of 1911: ideology, technological innovation and the production of medical statistics. *Social History of Medicine*, Vol. 9, No. 3: 409-426.

HIGGS, Edward (2004) *The Information State in England*. Basingstoke: Palgrave Macmillan.

HILL, I. D. (1964) Review: Statistical Tables. *The British Medical Journal*, Vol. 1, No. 5376: 172.

HOGBEN, Lancelot (1931) *Genetics Principles in Medicine and Social Science*. London: Williams and Norgate.

HORSBURGH, E. M. (ed) (1914) *Handbook of the Exhibition of Napier Relics and of Books, Instruments and Devices for Facilitating Calculation*. Edinburgh: Royal Society of Edinburgh.

HOTELLING, Harold (1938) Review: Statistical Tables. *Science*, Vol. 88, No. 2295: 596-597.

HUGHES, Agatha C. and HUGHES, Thomas P. (eds) (2000) *Systems, Experts, and Computers: The Systems Approach in Management and Engineering: WWII and after*. Cambridge, Massachusetts, and London, England: The MIT Press.

IKIN, Elizabeth W., PRIOR, Aileen M., RACE, Robert R. and TAYLOR, George L. (1939) The distributions in the A_1A_2BO blood groups in England. *Annals of Eugenics*, Vol. 9, No. 4: 409-411.

INSTITUT INTERNATIONAL D'AGRICULTURE (1908) *Assemblée générale Actes*. Rome: Imprimerie de l'Institut International d'Agriculture.

INSTITUT INTERNATIONAL D'AGRICULTURE (1927) *L'Institut International d'Agriculture: Organisation, Activité, Résultats*. Rome: Imprimerie de l'Institut International d'Agriculture.

INSTITUT INTERNATIONAL D'AGRICULTURE (1933) *Les Institutions d'Expérimentation Agricole dans les Pays Tempérés*. Rome: Imprimerie de la Chambre des Députés.

JOHNSTON, James F. W. (1849) *Experimental Agriculture, Being the Results of Past and Suggestions for Future Experiments in Scientific and Practical Agriculture*. Edinburgh: W. Blackwood and Sons.

KAIJSER, Arne and TIBERG, Joar (2000) From operations research to future studies: the establishment, diffusion and transformation, of the systems approach in Sweden. In: Hughes, Agatha C. and Hughes, Thomas P. (eds) *Systems, Experts, and Computers: The Systems Approach in Management and Engineering: WWII and after*, Cambridge, Massachusetts, and London, England: The MIT Press; pp. 385- 412.

KAY, H. D. (1971) William Kershaw Slater, 1893-1970. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 17: 663-680.

KAY, Lily E. (1993) *The Molecular Vision of Life: Caltech, the Rockefeller Foundation, and the Rise of the New Biology*. New York and Oxford: Oxford University Press.

KEITH, S. T. (1981) Inventions, patents and commercial development from governmentally financed research in Great Britain: the origins of the National Research Development Corporation. *Minerva*, Vol. 19, No. 1: 92-122.

KEMPTHORNE, Oskar (1946) The use of a punched-card system for the analysis of survey data, with special reference to the analysis of the national farm survey. *Journal of the Royal Statistical Society*, Vol. 109, No. 3: 284-295.

KEVLES, Daniel J. (1995) *In the Name of Eugenics: Genetics and the Uses of Human Heredity*. Cambridge, Massachusetts, and London, England: Cambridge University Press.

KNUTH, Donald E. and TRABB PARDO, Luis (1980) The early development of programming languages. In: Metropolis, N., Howlett, J. and Rota, Gian Carlo (eds) *A History of Computing in the Twentieth Century*, New York: Academic Press; pp. 197-273.

KOHLER, Robert E. (1991) *Partners in Science: Foundations and Natural Scientists, 1900-1945*. Chicago and London: The University of Chicago Press.

KOHLER, Robert E. (2002) *Landscapes and Labscapes*. Chicago and London: The University of Chicago Press.

LANCASTER, Henry O. (1994) *Quantitative Methods in Biological and Medical Sciences*. New York: Springer.

LAVINGTON, Simon (2011) *Moving Targets: Elliott-Automation and the Dawn of the Computer Age in Britain, 1947-1967*. London: Springer-Verlag.

LAWLER, Sylvia and RENWICK, James H. (1955) Genetical linkage between the ABO and nail-patella loci, *Annals of Human Genetics*, Vol. 19, No. 4: 312-331.

LAWLER, Sylvia and RENWICK, James H. (1959) Blood groups and genetic linkage. *British Medical Bulletin*, Vol. 15, No. 2: 145-149.

LAWLER, Sylvia, RENWICK, James H., HAUGE, M., MOSBECH, J. and WILDERVANCK, L. S. (1958) Linkage tests involving the P blood group locus and further data on the *ABO* : nail-patella linkage. *Annals of Human Genetics*, Vol. 22, No. 4: 342-355.

LEHMANN, Eric L. (1993) The Fisher, Neyman-Pearson theories of testing hypothesis: one theory or two? *Journal of the American Statistical Association*, Vol. 88, No. 424: 1242-1249.

LEHMANN, Eric L. (2011) *Fisher, Neyman and the Creation of Classical Statistics*. New York: Springer.

VON LIEBIG, Justus (1840) [edited by L. Playfair] *Organic Chemistry in its Applications to Agriculture and Physiology*. London: Taylor and Walton [accessed 14th July 2012, http://archive.org/stream/organicchemistry00liebrich#page/n5/mode/2up].

LIPPHARDT, Veronika (2010) The Jewish community of Rome: an isolated population? Sampling procedures and bio-historical narratives in genetic analysis in the 1950s. *BioSocieties*, Vol. 5, No. 3: 306-329.

LOVELL, Bernard (1975) Patrick Maynard Stuart Blackett, Baron Blackett, of Chelsea, 18th December 1897 – 13 July 1974. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 21: 1-115.

LUDBROOK, John (2005) R. A. Fisher's life and death in Australia, 1959-1962. *The American Statistician*, Vol. 59, No. 2: 164-165.

MACKENZIE, Donald A. (1981) *Statistics in Britain, 1865-1930*. Edinburgh: Edinburgh University Press.

MAGNELLO, M. Eileen (1993) *Karl Pearson: Evolutionary Biology and the Emergence of a Modern Theory of Statistics (1884-1936)*. Doctoral Dissertation, University of Oxford.

MAGNELLO, M. Eileen (1999a), The non-correlation of biometrics and eugenics: rival forms of laboratory work in Karl Pearson's career at University College London (Part I). *History of Science*, Vol. 37, No. 115: 79-106.

MAGNELLO, M. Eileen (1999b), The non-correlation of biometrics and eugenics: rival forms of laboratory work in Karl Pearson's career at University College London (Part II). *History of Science*, 37, No. 116: 123-150.

MAINLAND, Donald (1938) *The Treatment of Clinical and Laboratory Data: An Introduction to Statistical Ideas and Methods for Medical and Dental Workers*. Edinburgh: Oliver and Boyd.

MARKS, Harry M. (2000) *The Progress of Experiment. Science and Therapeutic Reform in the United States, 1900-1990*. Cambridge: Cambridge University Press.

MARKS, Jonathan (1996) The legacy of serological studies in American physical anthropology. *History and Philosophy of the Life Sciences*, Vol. 18, No. 3: 345-362.

MARTIN, Ernst (1992) [English translation by P. Aldrich Kidwell and M. R. Williams] *The Calculating Machines*. Cambridge, Massachusetts, and London, England: The MIT Press [accessed 5th July 2012, http://www.rechenmaschinen-illustrated.com/Martins_book/Ernst%20Martin%20-%20Rechen%20Machinen%20OCR%204.pdf].

MASKELL, Ernest J. (1929) Experimental error: a survey of recent advances in statistical methods. *Tropical Agriculture Trinidad*, Vol. 5: 306-309; Vol. 6: 5-11, 45-48, 97-99.

MAZUMDAR, Pauline M. H. (1995) *Species and Specificity: An Interpretation of the History of Immunology*. Cambridge: Cambridge University Press.

MAZUMDAR, Pauline M. H. (1996) Two models for human genetics: blood grouping and psychiatry in Germany between the world wars. *Bulletin of the History of Medicine*, Vol. 70, No. 4: 609-657.

MAZUMDAR, Pauline M. H. (2011) *Eugenics, Human Genetics and Human Failings: The Eugenics Society, its Sources and its Critics in Britain*. London and New York: Routledge.

MCINTOSH, I., DUNSTON, J. A., LIU, L., HOOVER-FONG, J. E. and SWEENEY, E. (2005) Nail patella syndrome revisited: 50 years after linkage. *Annals of Human Genetics*, Vol. 69, No. 4: 349-363.

MCMULLEN, Lance (1939) "Student" as a Man. *Biometrika*, Vol. 30, No. 3-4: 205-210.

MERCER, W. B. and HALL, Alfred D. (1911) The experimental error of field trials (with an Appendix by Student). *The Journal of Agricultural Science*, Vol. 4, No. 2: 107-132.

MILLER, J. C. P. (1949) The BAASMTC now RSMTC. In: *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 25: 333-340.

MINISTRY OF INFORMATION (1945) *Life Blood: The Official Account of the Transfusion Services*. London: His Majesty's Stationary Office.

MISA, Thomas J. (2007) Understanding 'How computing has changed the world'. *IEEE Annals of the History of Computing*, Vol. 29, No. 4: 52-63.

MOORE, James (2006) R. A. Fisher: a faith fit for eugenics. *Studies in the History and Philosophy of Biological and Biomedical Sciences*, Vol. 38, No. 1: 110-135.

MORGAN WATKIN, I. (1964) English and Welsh racial elements in Western Shropshire and in the adjacent Welsh borderland: ABO blood group evidence. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, Vol. 94, No. 1: 52-65.

MOURANT, Arthur E. (1954) *The Distribution of the Human Blood Groups*. Oxford: Blackwell Scientific Publications.

MOURANT, Arthur E., KOPEC, Ada C. and DOMANIEWSKA-SOBCZAK, Kazimiera (1958) *The ABO Blood Groups: Comprehensive Tables and Maps of World Distribution*. Oxford: Blackwell Scientific Publications.

MÜLLER-WILLE, Staffan (2005) Early Mendelism and the subversion of taxonomy: epistemological obstacles as institutions. *Studies in History and Philosophy of Biological and Biomedical Sciences*, Vol. 36, No. 3: 465-487.

NATHAN, Henry (1970) 'Bernstein, Felix'. In: Gillispie, Charles C. (ed) *Dictionary of Scientific Biography*. New York: Charles Scribner's Sons; pp. 58-59.

NEYMAN, Jerzy (1934) On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, Vol. 97, No. 4: 558-625.

NOVEMBER, Joseph (2012) *Biomedical Computing: Digitizing Life in the United States*. Baltimore: The John Hopkins University Press.

NUTMAN, P. S. (1977) Henry Gerard Thornton, 22 January 1892-6 February 1977. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 23: 557-574.

NYE, Mary J. (2004) Blackett, Patrick Maynard Stuart, Baron Blackett (1897–1974). *Oxford Dictionary of National Biography*. Oxford: Oxford University Press; online edition January 2011, [accessed 8th January 2013, http://www.oxforddnb.com/view/article/30822].

OFFICE MACHINES RESEARCH INC. (1940) *American Office Machines Research Service*, Vol. 3. New York: Office Machines Research.

OLBY, Robert (1991) Social imperialism and state support for agricultural research in Edwardian Britain. *Annals of Science*, Vol. 48, No. 6: 509-526.

PALLADINO, Paolo (1990) The political economy of applied research: plant breeding in Great Britain, 1910-1940. *Minerva*, Vol. 28, No. 4: 446-468.

PAYNE, Roger and SENN, Stephen (2010) John Nelder obituary. *The Guardian*, 23rd September 2010 [accessed 28th November 2012, http://www.guardian.co.uk/technology/2010/sep/23/john-nelder-obituary/print].

PEARSON, Egon S. (1926) Review of *Statistical Methods for Research Workers* (R. A. Fisher). *Science Progress*, Vol. 20: 733-734.

PEARSON, Egon S. (1929) Statistics in biological research. Review of *Statistical Methods for Research Workers* (2nd edition) by. Dr R. A. Fisher. *Nature*, Vol. 123, No. 3110: 866-867.

PEARSON, Egon S. (1936) Karl Pearson. An appreciation of some aspects of his life and work. Part I: 1857-1906. *Biometrika*, Vol. 28, No. 3-4: 193-257.

PEARSON, Egon S. (1938) Karl Pearson. An appreciation of some aspects of his life and work. Part II: 1857-1906. *Biometrika*, Vol. 29, No. 3-4: 161-248.

PEARSON, Egon S. (1939) "Student" as statistician. *Biometrika*, Vol. 30, No. 3-4: 210-250.

PEARSON, Egon S. (1957) John Wishart, 1898-1956. *Biometrika*, Vol. 44, No. 1-2: 1-8.

PEARSON, Egon S. (1990) *Student: A Statistical Biography of William Sealy Gosset*. Based on the writings by E. S. Pearson; edited and augmented by R. L. Plackett with the assistance of G. A. Barnard. Oxford: Clarendon Press.

PEARSON, Egon S. and HARTLEY, Herman O. (1966) *Biometrika Tables for Statisticians*, Vol. 1 (3rd edition). Cambridge: Cambridge University Press.

PEARSON, Egon S. and WISHART, John (eds) (1942) *"Student's" Collected Papers*. London: *Biometrika* Office.

PEARSON, Karl (1914) *Tables for Statisticians and Biometricians*. Cambridge: Cambridge University Press.

PICARD, Jean-François and SCHENEIDER, William H. (1999) The Rockefeller Foundation and the Development of Biomedical Research in Europe. In: Gemelli, Giuliana, Picard, Jean-François, Scheneider, William H. (eds) *Managing Medical Research in Europe: The Role of the Rockefeller Foundation (1920s-1950s)*. Bologna: CLUEB; pp. 13-50.

PIRIE, N. W. (1966) John Burdon Sanderson Haldane, 1892-1964. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 12: 218-249.

PORTER, Theodore M. (1986) *The Rise of Statistical Thinking, 1820-1900*. Princeton: Princeton University Press.

PORTER, Theodore M. (1996) *Trust in Numbers*. Princeton: Princeton University Press.

PORTER, Theodore M. (2006) *Karl Pearson: The Scientific Life in a Statistical Age*. Princeton: Princeton University Press.

PROVINE, William B. (2001) *The Origins of Theoretical Population Genetics*. Chicago and London: The University of Chicago Press.

PUGH, Emerson W. (1995) *Building IBM: Shaping an Industry and its Technology*. Cambridge, Massachusetts, and London, England: The MIT Press.

QUIRKE, V. M. (2004) Haldane, John Burdon Sanderson (1892–1964). *Oxford Dictionary of National Biography*. Oxford: Oxford University Press; online edition January 2011, [accessed 8th January 2013, http://www.oxforddnb.com/view/article/30822].

RACE, Robert R. and SANGER, Ruth (1954) *Blood Groups in Man*. Oxford: Blackwell Scientific Publications.

RÅDE, Lennart (1990) Statistics and the Ccomputer. *International Conference for Teaching Statistics 3* [accessed 2nd December 2012, http://www.stat.auckland.ac.nz/~iase/publications/18/BOOK1/C8-2.pdf].

RASMUSSEN, Nicolas (1996) Making a machine instrumental: RCA and the wartime origins of biological electron microscopy in America, 1940-1945. *Studies in the History and Philosophy of Science Part A*, Vol. 27, No. 3: 311-349.

RAU, Erik P. (2001) Technological systems, expertise and policy making: the British origins of operational research. In: Allen, Michael T. and Hecht, Gabrielle (eds), *Technologies of Power: Essays in honor of Thomas Parke Hughes and Agatha Chipley Hughes*, Cambridge, Massachusetts and London, England: The MIT Press.

RAU, Erik P. (2005) Combat science: the emergence of operational research in World War II. *Endeavour*, Vol. 29, No. 4: 156-161.

REID, Constance (1998) *Neyman*. New York: Springer.

RENWICK, James H. (1956) Linkage in human genetics. *The Eugenics Review*, Vol. 48, No. 3: 149-153.

RENWICK, James H. and BOLLING, D. (1967) A program-complex from encoding, analyzing, and storing human linkage data. *American Journal of Human Genetics*, Vol. 19, No. 3: 360-367.

RENWICK, James H. and SCHULZE, Jane (1961) A computer programme for the processing of linkage data from large pedigrees (Abstr.). *Excerpta Medica, International Congress Series*, Vol. 32: E145.

RIMOIN, David L. (2008) Obituary: Victor A. McKusick, 1921-2008. *Nature*, Vol. 40, No. 9: 1037.

ROBSON, John (1929) Labour-saving in clerical work. *Public Administration*, Vol. 7, No. 3: 278-284.

ROCKEFELLER FOUNDATION (1936) *The Rockefeller Foundation Annual Report 1935*. New York: Rockefeller Foundation.

ROSS, Gavin (undated) *Ronald and his line*. Unpublished poem.

ROSS, Gavin (undated) Using the Elliott 401 (personal memories).

ROTHAMSTED EXPERIMENTAL STATION (undated) *Programme of Work at the Rothamsted Experimental Station for the Years 1919-1920*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1913) *Annual Report for 1912*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1921) *Report 1918-1920 With the Supplement to the 'Guide to Experimental Plots'*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1923) *Report 1921-22 With the Supplement to the 'Guide to Experimental Plots'*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1925) *Report 1923-24 With the Supplement to the 'Guide to Experimental Plots'*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1927) *Report 1925-26 With the Supplement to the 'Guide to Experimental Plots'*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1929a) *Report 1927-28 With the Supplement to the 'Guide to Experimental Plots'*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1929b) *Records of the Rothamsted Staff Harpenden*, No. 1.

ROTHAMSTED EXPERIMENTAL STATION (1930) *Report for 1929*. Harpenden: Printed by D. J. Jeffery.

ROTHAMSTED EXPERIMENTAL STATION (1931a) *Report for 1930*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1931b) *Guide to the Laboratories and Selected Field Experiments*.

ROTHAMSTED EXPERIMENTAL STATION (1931c) *Records of the Rothamsted Staff Harpenden*, No. 3.

ROTHAMSTED EXPERIMENTAL STATION (1932) *Report for 1931*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1933) *Report for 1932*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1934) *Report for 1933*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1935a) *Report for 1934*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1935b) *Records of the Rothamsted Staff Harpenden*, No. 5.

ROTHAMSTED EXPERIMENTAL STATION (1936) *Report for 1935*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1937) *Report for 1936*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1938) *Report for 1937*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1939) *Report for 1938*. St. Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1943) *Rothamsted Experimental Station Centenary Celebrations, 1843-1943*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1946) *Report for the War Years 1939-1945*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1948a) *Report for 1946*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1948b) *Report for 1947*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1949) *Report for 1948*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1950) *Report for 1949*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1951) *Report for 1950*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1952) *Report for 1951*. St Albans: Gibbs and Bamforth.

ROTHAMSTED EXPERIMENTAL STATION (1953) *Report of the Rothamsted Experimental Station for 1952*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1954) *Report of the Rothamsted Experimental Station for 1953*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1955) *Report of the Rothamsted Experimental Station for 1954*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1956) *Report of the Rothamsted Experimental Station for 1955*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1957) *Report of the Rothamsted Experimental Station for 1956*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1958) *Report of the Rothamsted Experimental Station for 1957 With 50-Year Index (1908-57)*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1959) *Report of the Rothamsted Experimental Station for 1958*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STation (1960) *Report of the Rothamsted Experimental Station for 1959*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1961) *Report of the Rothamsted Experimental Station for 1960*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1962) *Report of the Rothamsted Experimental Station for 1961*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1963) *Report for 1962*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1964) *Report for 1963*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1965) *Report for 1964*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1966) *Report for 1965*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1967) *Report for 1966*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1968) *Report for 1967*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED EXPERIMENTAL STATION (1969) *Report for 1968 Part 1*. Harpenden: Rothamsted Experimental Station.

ROTHAMSTED RESEARCH (2006) *Guide to the Classical and Other Long-term Experiments, Datasets and Sample Archive*. [accessed 29th May 2012,

http://www.rothamsted.ac.uk/Content.php?Section=Resources&Page=ExperimentsGuide].

ROWELL, J. G. (1954) The analysis of a factorial experiment (with confounding) on an electronic computer. *Journal of the Royal Statistical Society Series B*, Vol. 16, No. 2: 242-246.

ROYAL STATISTICAL SOCIETY (1934) *Annals of the Royal Statistical Society, 1834-1934*. London: The Royal Statistical Society.

RUSSELL, E. John (1935) Rothamsted in 1933-1934. In: *Records of the Rothamsted Staff Harpenden*, No. 5, pp. 7-9.

RUSSELL, E. John (1942) Alfred Daniel Hall, 1864-1942. *Obituary Notices of Fellows of the Royal Society*, Vol. 4, No. 11: 228-250.

RUSSELL, E. John (1956) *The Land Called Me*. London: George Allen and Unwin Ltd.

RUSSELL, E. John (1966) *A History of Agricultural Science in Great Britain, 1620-1954*. London: George Allen and Unwin Ltd.

RUSSELL, E. John and VOELCKER, J. A. (1936) *Fifty Years of Field Experiments at the Woburn Experimental Station. With a Statistical Report by W. G. Cochran*. London: Longmans, Green and Co.

SALSBURG, David (2001) *The Lady Tasting Tea. How Statistics Revolutionized Science in the Twentieth Century*. New York: W. H. Freeman.

SALVAGE, Leonard J. (1976) On re-reading R. A. Fisher. *The Annals of Statistics*, Vol. 4, No. 3: 441-500.

SCHAPPACHER, Norbert (2005) Felix Bernstein. *International Statistical Review*, Vol. 73, No. 1: 3-7.

SCHNEIDER, William H. (1983) Chance and social setting in the application of the discovery of blood groups. *Bulletin of the History of Medicine*, Vol. 57, No. 4: 545-562.

SCHNEIDER, William H. (1995) Blood group research in Great Britain, France, and the United States between the world wars. *Yearbook of Physical Anthropology*, Vol. 38, No. S21: 87-114.

SCHNEIDER, William H. (1996a) Introduction. *History and Philosophy of the Life Sciences*, Vol. 18, No. 3: 275-276.

SCHNEIDER, William H. (1996b) The history of research on blood group genetics: initial discovery and diffusion. *History and Philosophy of the Life Sciences*, Vol. 18, No. 3: 277-303.

SEIDENFELD, Teddy (1979) *Philosophical Problems of Statistical Inference: Learning from R. A. Fisher*. Dordrecht, Boston: Reidel.

SIMPSON, Howard R. (1958) The estimation of linkage on an electronic computer. *Annals of Human Genetics*, Vol. 22, No. 4: 356-361.

SIMPSON, Howard R. (1961) The analysis of survey data on an electronic computer. *Journal of the Royal Statistical Society Series A*, Vol. 124, No. 2: 219-226.

SNEATH, Peter H. A. (1957) The application of computers to taxonomy. *Journal of General Microbiology*, Vol. 17, No. 1: 201-226.

SNEATH, Peter H. A. (2010) Reflections on microbial systematics. *Bulletin of the Bergey's International Society for Microbial Systematics*, Vol. 1, No. 1: 77-83 [accessed 24th November 2012, http://www.bergeys.org/bulletin/BulletinofBISMIS%201-1.pdf].

SNEDECOR, George W. (1928) Uses of punched card equipment in mathematics. *The American Mathematical Monthly*, Vol. 35, No. 4: 161-169.

STAR, Susan Leigh and GRIESEMER, James R. (1989) Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907-1939. *Social Studies of Science*, Vol. 19, No. 3: 387-420.

STEVENS, Wilfred L. (1938) Estimation of blood-group gene frequencies. *Annals of Eugenics*, Vol. 8, No. 4: 362-375.

STIGLER, Stephen M. (1986) *The History of Statistics: The Measurement of Uncertainty Before 1900*. Cambridge, Massachusetts and London, England: Harvard University Press.

STIGLER, Stephen M. (2002) *Statistics on the Table*. Cambridge, Massachusetts and London, England: Harvard University Press.

STUDENT (alias W. S. Gosset) (1908) The probable error of a mean. *Biometrika*, Vol. 6, No. 1: 1-25.

STUDENT (alias W. S. Gosset) (1917) Tables for estimating the probability that the mean of a unique sample of observations lies between $-\infty$ and any given distance of the mean of the population from which the sample is drawn. *Biometrika*, Vol. 11, No. 4: 414-417.

STUDENT (alias W. S. Gosset) (1925) New tables for testing the significance of observations. *Metron*, Vol. 5, No. 3: 25-32.

STUDENT (alias W. S. Gosset) (1926) Mathematics and agronomy. *Journal of the American Society of Agronomy*, Vol. 18: 703-720.

TAYLOR, George L. and PRIOR, Aileen M. (1938a) Blood groups in England I. Examination of family and unrelated material. *Annals of Eugenics*, Vol. 8, No. 4: 343-355.

TAYLOR, George L. and PRIOR, Aileen M. (1938b) Blood groups in England II. Distribution in the population. *Annals of Eugenics*, Vol. 8, No. 4: 356-361.

TAYLOR, George L. and PRIOR, Aileen M. (1939a) Blood groups in England III. Discussion of the family material. *Annals of Eugenics*, Vol. 9, No. 1: 18-44.

TAYLOR, George L. and PRIOR, Aileen M. (1939b) The distribution of the M and N factors in random samples of different races. *Annals of Eugenics*, Vol. 9, No. 2: 97-108.

TAYLOR, George L., RACE, Robert R., PRIOR Aileen M. and IKIN, Elizabeth W. (1942) A reliable technique for the diagnosis of the ABO blood groups. *The Journal of Pathology and Bacteriology*, Vol. 54, No. 1: 81-87.

THOMPSON, F. M. L (2004) Lawes, Sir John Bennet, first baronet (1814-1900). *Oxford Dictionary of National Biography*. Oxford: Oxford University Press [accessed 30th June 2012, http://www.oxforddnb.com/view/article/16159].

THORNTON, H. Gerard (1966) Edward John Russell, 1872-1965. *Biographical Memoirs of Fellows of the Royal Society*, Vol. 12: 456-477.

THURTLE, Phillip (2007) *The Emergency of Genetic Rationality*. Seattle and London: University of Washington Press.

TOCHER, Keith D. (1952) The design and analysis of block experiments. *Journal of the Royal Statistical Society Series B*, Vol. 14, No. 1: 45-100.

VAUGHAN, Janet M. and PANTON, (Sir) Philip (1952) The civilian blood transfusion service. In: C. L. Dunn (ed.) *The Emergency Medical Services*. Vol. I: England and Wales. London: Her Majesty's Stationary Office; pp. 334-355.

WARWICK, Andrew (1995) The laboratory of theory or what's exact about the exact sciences? In: Wise, N. (ed.) *The Values of Precision*, Princeton: Princeton University Press; pp. 311-351.

WESTON, B. (1962) *Fifty Years at Rothamsted, 1910-1960*. Privately printed.

WIGNER, Eugene P. (1960) The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, Vol. 13, No. 1: 1-14.

WILKES, Maurice V., WHEELER, David J. and GILL, Stanley (1951) *The Preparation of Programs for an Electronic Digital Computer*. Cambridge, Massachusetts: Addison-Wesley Press.

WINNER, Langdon (1980) Do artifacts have politics? *Daedalus*, Vol. 109, No. 1: 121-136.

WINNER, Langdon (1983) *Autonomous Technology*: *Technics-out-of-Control as a Theme in Political Thought.* Cambridge, Massachusetts, and London, England: The MIT Press.

WINNER, Langdon (1989) *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago and London: The University of Chicago Press.

WINNER, Langdon (1993) Upon opening the black box and finding it empty: social constructivism and the philosophy of technology. *Science, Technology and Human Values*, Vol. 18, No. 3: 362-378.

WISHART, John (1939) Review: Tables for Statisticians. *Nature*, Vol. 144, No. 3647: 533.

WOOD, Thomas B. (1910) The feeding value of mangels. *Journal of Agricultural Science*, Vol. 3, No. 3: 225-232.

WOOD, Thomas B. and STRATTON, Frederick J. M. (1910) The interpretation of experimental results. *The Journal of Agricultural Science*, Vol. 3, No. 4: 417-440.

WOOD, Thomas B. and YULE, George Udny (1914) Statistics of British feeding trials and the starch equivalent theory. *The Journal of Agricultural Science*, Vol. 6, No. 2: 233-251.

YATES, Frank (1943) Methods and purposes of agricultural surveys. *Paper read before the Royal Society of Arts* on 24th March 1943: 1-10.

YATES, Frank (1950) Agriculture, sampling and operational research. *Bulletin of the International Statistical Institute*, Vol. 32: 220-227.

YATES, Frank (1951) The influence of *Statistical Methods for Research Workers* on the development of the science of statistics. *Journal of the American Statistical Association*, Vol. 46, No. 253: 19-34.

YATES, Frank (1952) George Udny Yule, 1871-1951. *Obituary Notices of Fellows of the Royal Society*, Vol. 8, No. 21: 308-323.

YATES, Frank (1953) *Sampling Methods for Censuses and Surveys* (2nd edition). London: Charles Griffin and Company.

YATES, Frank (1959) Wilfred Leslie Stevens. *Journal of the Royal Statistical Society Series A*, Vol. 122, No. 3: 403-404.

YATES, Frank (1960a) *Sampling Methods for Censuses and Surveys* (3rd edition). London: Charles Griffin and Company.

YATES, Frank (1960b) The use of electronic computers in the analysis of replicated experiments, and groups of experiments of the same design. *Bulletin de l'Institut Agronomique et des Stations de Recherches de Gembloux*, Hors série, Volume 1: 201-210.

YATES, Frank (1962) Computers in research – promise and performance. Presidential addressed delivered to the British Computer Society in London on 26th September 1961. *Computer Journal*, Vol. 4, No. 4: 273-279.

YATES, Frank (1966) Computers, the second revolution in statistics. *Biometrics*, Vol. 22, No. 2: 233-251.

YATES, Frank (1970) *Experimental Design: Selected Papers of Frank Yates.* London: Charles Griffin and Co. Ltd.

YATES, Frank (1973) The analysis of surveys on computers – Features of the Rothamsted survey program. *Journal of the Royal Statistical Society Series C*, Vol. 22, No. 2: 161-171.

YATES, Frank (1981) Rothamsted in the 1930s. *Bulletin in Applied Statistics*, Vol. 8, No. 2: 134-156.

YATES, Frank (1990) Foreword. In: Fisher, Ronald A. *Statistical Methods, Experimental Design and Scientific Inference*. Oxford: Oxford University Press.

YATES, Frank and BOYD, David A. (1965) Two decades of surveys of fertilizer practice. *Outlook on Agriculture*, Vol. 4, No. 3: 203-210.

YATES, Frank and HEALY, Michael J. R. (1951) Statistical methods in anthropology. *Nature*, Vol. 168, No. 4287: 1116-1117.

YATES, Frank, HEALY, Michael J. R. and LIPTON, Steve (1957) Routine analysis of replicated experiments on an electronic computer. *Journal of the Royal Statistical Society Series B*, Vol. 19, No. 2: 234-254.

YATES, Frank and LIPTON, Steve (1957) An automatic programming routine for the Elliott 401. *Journal of the Association for Computing Machinery*, Vol. 4, No. 2: 151-156.

YATES, Frank and MATHER, Kenneth (1963) Ronald Aylmer Fisher (1890-1962). *Biographical Memoirs of Fellows of the Royal Society*, Vol. 9: 91-129.

YATES, Frank and REES, Douglas H. (1958) The use of an electronic computer in research statistics: four years' experience. *Computer Journal*, Vol. 1, No. 2: 49-58.

YATES, Frank and SIMPSON, Horace R. (1960) A general program for the analysis of surveys. *Computer Journal*, Vol. 3, No. 3: 136-140.

YATES, Frank and SIMPSON, Horace R. (1961) The analysis of surveys: processing and printing the basic tables. *Computer Journal*, Vol. 4, No. 1: 20-24.

YATES, Frank, GOWER, John C. and SIMPSON, Horace R. (1963) A specialized autocode for the analysis of replicated experiments. *Computer Journal*, Vol. 5, No. 4: 313-319.

YATES, JoAnne (1993) *Control through Communication: The Rise of System in American Management*. Baltimore and London: The John Hopkins University Press.

YATES, JoAnne (2008) *Structuring the Information Age: Life Insurance and Technology in the Twentieth Century*. Baltimore: The John Hopkins University Press.

YOUNG, Arthur (1771) *A Course of Experimental Agriculture*. Dublin: J. Exshaw.

YULE, George Udny (1922) *An Introduction to the Theory of Statistics* (6th edition). London: Charles Griffin and Co.

YULE, George Udny and FILON, Louis N. G. (1936) Karl Pearson, 1857-1936. *Obituary Notices of Fellows of the Royal Society*, Vol. 2, No. 5: 72-110.

ZILIAK, Stephen T. (2008) Is the *p*-form of the *t*-table fraudulent? Yes, it misrepresents scientific thought (and credit) [accessed 26th October 2012

http://stephentziliak.com/doc/Ziliak%20article%20on%20p%20form%20of%20t%20table%20Web%20version%20Oct.%202008.pdf].

ZILIAK, Stephen T. and MCCLOSKEY, Deirdre N. (2008) *The Cult of Statistical Significance*. Ann Arbor: the University of Michigan Press.

ZUCKERMAN, Solly (1988) *From Apes to Warlords, 1904-46*. London: Collins.

Interviews

Sir David Cox interviews Frank Yates, 24th April 1986, Royal Statistical Society.

Paul Merchant interviews David Jenkinson, 29th March 2010, Part 6, British Library (accessed, 23rd November 2012, http://sounds.bl.uk/Oral-history/Science/021M-C1379X0006XX-0006V0).

Giuditta Parolini interviews Prof. John C. Gower, September 2010, skype interview.

Giuditta Parolini interviews Mrs (Lucille) Vera Wiltsher, 1st September 2011, Rothamsted Research.