



UNIVERSITY OF LEEDS

This is a repository copy of *The resisted rise of randomisation in experimental design: British agricultural science, c.1910-1930*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/87554/>

Version: Accepted Version

Article:

Berry, DJ (2015) The resisted rise of randomisation in experimental design: British agricultural science, c.1910-1930. *History and Philosophy of the Life Sciences*, 37 (3). 242 - 260. ISSN 0391-9714

<https://doi.org/10.1007/s40656-015-0076-8>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The resisted rise of randomisation in experimental design: British agricultural science, c.1910-1930

Dominic Berry
University of Leeds
Philosophy
Michael Sadler Building
LS2 9JT
d.berry@leeds.ac.uk

Abstract

The most conspicuous form of agricultural experiment is the field trial, and within the history of such trials, the arrival of the randomised control trial (RCT) is considered revolutionary. Originating with R.A. Fisher within British agricultural science in the 1920s and 30s, the RCT has since become one of the most prodigiously used experimental techniques throughout the natural and social sciences. Philosophers of science have already scrutinised the epistemological uniqueness of RCTs, undermining their status as the ‘gold standard’ in experimental design. The present paper introduces a historical case study from the origins of the RCT, uncovering the initially cool reception given to this method by agricultural scientists at the University of Cambridge and the (Cambridge based) National Institute of Agricultural Botany (NIAB). Rather than giving further attention to the RCT, the paper focuses instead on a competitor method – the half-drill strip – which both predated the RCT and remained in wide use for at least a decade beyond the latter’s arrival. In telling this history, John Pickstone’s *Ways of Knowing* is adopted, as the most flexible, and productive way to write the history of science, particularly when sciences and scientists have to work across a number of different kinds of place. It is shown that those who resisted the RCT did so in order to preserve epistemic and social goals that randomisation would have otherwise run a tractor through.

Keywords

Field science; Agriculture; Randomised Control Trials; Epistemic goals; Ways of knowing

Acknowledgments

Firstly my thanks to Giuditta Parolini for conceiving of this special issue and inviting my contribution. Giuditta gave extensive feedback on numerous drafts, for which I am very grateful. Two anonymous reviewers and Staffan Müller-Wille greatly improved the article with excellent criticisms and suggestions. Greg Radick, Lijing Jiang, and Carl Warom gave generously with their time scrutinising final drafts. Matt Paskins and Mike Finn discussed the paper with me on a number of occasions, suggesting important changes. Lastly I thank those at the 10th iHPS meeting in Durham (and the conference organisers) who responded to this work with exciting questions (particularly Sabina Leonelli, William Peden, Alper Bilgili and Nadia Moro), and Liz Watkins and Kiara White for always wanting to know more about agricultural science. This work was funded by the Arts and Humanities Research Council.

1. Introduction

When it comes to the history and philosophy of agricultural experiments, Britain in the 1920s and 30s is remembered above all for the pioneering research of R.A. Fisher at Rothamsted Experimental Station (RES).¹ While there is a small scholarly disagreement over who pioneered randomisation in the setup of experimental design (Hacking 1988), it is generally agreed that Fisher did the most to establish and then popularise the randomised control trial (RCT) (Armitage 2003; Gower 1988; Hall 2002; Parolini 2014). Today the virtues of the RCT are often lauded as the best means of ensuring *truly* reliable experimental results in everything from medicine to public policy. Philosophers of science have already scrutinised the uniqueness of RCTs, undermining their status as the ‘gold standard’ in experimental design (Cartwright 2007; Cartwright and Munro 2010).

While Fisher and the RCT are well known, the same cannot be said for his contemporary and collaborator, William Sealy Gosset, a fellow statistician, and analyst for Guinness breweries who is more famous today for his publications under the pen name ‘Student’ (McMullen 1939; Pearson 1939; Plackett and Barnard 1990). A disagreement eventually emerged between Gosset and Fisher that rumbled on for many years and became public in their lifetimes, one which revolved around the value of randomised as opposed to ‘systematic’ experimental design. A number of historians have had cause to comment on the debate, though none have pursued it particularly deeply, or attended to the immediate agricultural world for which it mattered (Box 1978; Gigerenzer et al. 1989; Hacking 1988; Hall 2007). Two recent authors who *have* given substantial time to the debate, have done so with a revisionist statistical programme in mind. They “lament what could have been in the statistical sciences if only Fisher had cared to understand the full import of Gosset’s insights” and describe the latter as “Charming, rustic, and mysterious...a very Woody Guthrie of mathematical statistics”, vastly different to the ‘waspy’ Fisher (McCloskey and Ziliak 2008, 212-214). The present paper will not address the more technical elements of the Fisher-Gosset debate, linking them to the wider history of twentieth century science and society, which shall be the job of a future paper. Instead here I approach the debate from the ground up, looking at a significant case – agricultural field trials organised by the University of Cambridge and the Cambridge based National Institute of Agricultural Botany (NIAB) – where the implications of Fisher’s and Gosset’s positions were materially realised. At the centre of the argument are two rival trialling methods, the half-drill strip, and the RCT. Adopting the latter over the former might appear to be nothing more than a question of good sense, of increasing the certainty of our experimental results,

¹ This research is derived in part from Chapter 2 of Berry (2014a).

and so on. On the contrary, the two methods embodied starkly different epistemological goals and values, that in some aspects were incommensurable.

Two philosophical conclusions are reached. Firstly, attention to the epistemic goals of experimental programmes (Ankeny and Leonelli 2011), is crucial for understanding choice of experimental method, explaining why experimenters sometimes sharply disagree – and in cases where different evidentiary standards are in play *cannot* agree (Hicks 2015) – as to the best method to adopt. Secondly, thanks to the material form that it imposes on experimental resources, randomisation is prejudicial to certain epistemic goals and values that otherwise might be accommodated (Elliott and McKaughan 2014), and which therefore must be accounted for whenever randomisation is employed. In addition, and at the boundary between philosophy and sociology of science, the case also provides further evidence of the social and epistemological alignment that takes place fieldside between scientists and farmers (Diser 2012; Henke 2008; Latour 1988; Maat 2011).

The paper consists of two halves. The first analyses in what sense, and to what extent, these field trials constituted experiments. The answer requires close attention to their epistemic and social goals. Having explained the latter in detail, the first half ends by returning to the question of what kind of experiment these trials were, with an explanation in terms of John Pickstone's 'ways of knowing' (Pickstone 2000).² The second half of the paper then applies ways of knowing (WoK) to the Fisher-Gosset debate. The half-drill strip trialling method is considered first, and the extent to which it achieved the epistemological goals of the first half are explained. This exercise is then repeated for the RCT. It is argued that the desire to preserve certain epistemic and social goals, and three key epistemic values (of *reliability*, *imitability*, and *significant novelty*), are sufficient explanation for the divergence of views on the usefulness of the RCT in field trialling. Different epistemic goals can cause sharp disagreements about appropriate methodology, while RCTs narrow the variety of epistemic goals and values that might otherwise be embraced.

2. What kind of experiment are field trials?

This question is complementary to asking what kinds of agricultural experiment there are (Maat 2011), but requires attention to the philosophy of experiment. Answering this question is difficult because our most thoroughgoing accounts of experiment have focussed exclusively on laboratory settings, though recently there have been signs of change. Mary Morgan has made an important intervention on behalf of the qualitative difference between

² While WoK forms only one part of Pickstone's overall framework, which partners the latter with 'ways of working', ultimately constituting 'working knowledge' (Pickstone 2007; Pickstone 2011), for simplicity's sake, only WoK will concern us here as it is the part of his framework most directly related to experiment.

'Nature's experiments' and 'natural experiments', two forms of experiment that remain under-explored but which need to be integrated into our broader histories and philosophies of experiment (Morgan 2013). However there is not much in the latter that can immediately help us to understand field trials as experiments, other than recognising them as forms of experiment in the mode of 'intervention and control designed by scientist', much as laboratory experiments are. Another author, one who has also identified the dearth of material on the field sciences in the philosophy of experiment, has dedicated a manuscript to the subject (Schwarz 2014). However Schwarz follows the lead of the majority of historians and sociologists of the field sciences, by assuming from the outset that there will be a considerable difference between a 'laboratory ideal' and a 'field ideal' when it comes to experimentation (Kohler 2002). The present paper assumes the contrary; that the same ideals and epistemic values will be found operating more or less to some extent in all places, though they might differ in the ways that they exhibit themselves and the order in which they are prioritised.

This assumption is adopted firstly because places should not be conflated with practices, and secondly because there is no good reason to think that all epistemic values cannot be found operating to more or less some extent in all places. Place, epistemic value, and practice compose a nicely complementary research cluster. Place has been most extensively developed by geographers of science, whose efforts have contributed to a disciplinary 'spatial turn' in the history and philosophy of science (Finnegan 2008; Livingstone 2010; Naylor 2005), which has prompted a re-evaluation of the laboratory as a singularly important kind of place (Gooday 2008). Epistemic values have become increasingly important in the history and philosophy of science (along with epistemic goals), as scientific research programmes that deal with social, environmental, and industrial issues have become more central (Elliott 2011; Steel 2010). Practices meanwhile have been at the heart of some of our most innovative histories and philosophies of experiment (Chang 2012; Rheinberger 1997), though in such cases the emphasis on laboratory practice makes them exceedingly difficult to extend to the field. Such an effort will not be made here, though the exercise should be highly productive for someone more philosophically sophisticated than I.

Why should places not be conflated with practices? Places can certainly constrain the practices and research programmes that might be pursued within them (Müller-Wille 2005), while allowing particular practices to flourish. However, even when a new kind of place allows for a wholly new kind of practice to emerge, that practice cannot be conflated with that place. As Bruno Strasser has recently emphasised for his own study of *collecting* practices "To explore the possibility that collecting practices have been important, not only for the field and museum sciences but also for the laboratory sciences, it is first necessary to

question the common conflation by historical actors and scholars alike of places and practices” (Strasser 2012, 310). The same is true for any practice.

What does it mean to say that all epistemic values can be found operating to more or less some extent in all places? It means that standards for conforming to epistemic values, whether those values be of precision, replicability, consilience, and so on, will be set by researchers working in their own contexts. It is inappropriate to hold all researchers to some universal set of such values, grouping together one set as the ‘lab ideal’ and another set as the ‘field ideal’. Instead we should assume that researchers everywhere are interested in reliability, precision, replicability, and so on, but that the ways in which they hold themselves to these epistemic values can be exceedingly different. Sometimes those differences will clearly be related to the kind of place in which they are being exercised, but if we approach the issue from the wrong way round (seeing a place and immediately assuming what practices and epistemic values we will find there), we will miss all of the interesting ways in which scientists work across and through places.

Clearly these arguments need further clarification. To begin unpacking them it is first necessary to be much more specific about the field trials studied here, understanding their epistemic and social goals. Doing so shall also make it easier to understand what was at stake upon the arrival of the RCT.

2.1 The epistemic and social goals of varietal trialling

Field trials have been the cornerstone of efforts to improve farming for centuries (Jonsson 2013), becoming of interest to historians and philosophers of science somewhat more recently (Hall 2002; Schaffer 2003). What are the goals of such an activity? The first thing to admit is that while some large goals are often appealed to (for instance NIAB’s original official motto was “Grow More”), these broad goals are clearly best understood as composed of smaller sub-goals (Riggs 2003). In what follows a number of different goals (some of which were either epistemic or social, and others of which were *at one and the same time* social *and* epistemic) shall be teased out.

We can begin in 1908, with the founding of the Norfolk Agricultural Station (NAS), a privately owned field station established by some of the wealthier farmers in the county (Hutchinson and Owers 1980). Cambridge University’s T.B. Wood (Russell 1930) – who had been made Drapers Professor of Agriculture in 1907 – was a close collaborator with these Norfolk organisers and became a member of the Station’s Executive Committee. He

arranged for the research programme of NAS to become the responsibility of the University of Cambridge.³ The second annual report for the Station, 1909-10, opened by explaining that

During the last few years the value of breed in cereals has been brought very prominently before farmers, and the Committee have devoted some considerable attention to the subject. They consider that in no way can the Station better assist the farmers of the County than by trying all the new varieties of cereals which are now being so rapidly introduced, and growing on those which appear to be most suitable to the district for distribution to subscribers and others.⁴

This purported new era in plant breeding was being pioneered by geneticists such as Wood, which itself adds a further dimension to the specifics of varietal trialling in the first half of the twentieth century. Geneticists cultivated a greater sense of urgency around breeding with their productivity claims (Radick and Charnley 2013), touting their abilities to practically ‘engineer’ new plants, thereby increasing the need for the assessment of old and new varieties (Charnley 2013; Palladino 1994). Note the espoused goals of aiding farmers, by A) finding the varieties most suitable for a geographic area and B) doing so in a way that makes maximal use of all the new varieties ‘now being so *rapidly* introduced’. Goal A requires further clarification. The trials here were carried out with the intention of *differentiating between varieties*, taking in characteristics such as susceptibility to disease, response to climatic changes, and ultimately yield. They were designed to answer what Jonathan Harwood has called the ‘varietal question’ (Harwood 2012), which was a concern for many states modernising their agricultural industries at the turn of the twentieth century (Bonneuil 2006; Iori 2013; Maat 2001; Wieland 2006). Only by finding the varieties best suited to particular areas, and under normal farming conditions (so it was argued), could agricultural industries modernise and become sufficiently productive.

It was with these private resources, those of the Cambridge University farm, and his experiences working closely alongside the farming interest, that Wood began to introduce more sophisticated statistical techniques into agricultural field trials. In 1910 Wood published an article co-authored with the Cambridge astronomer F.J.M. Stratton, which addressed the probable error of results produced in a field trial.

³ A large number of the reports of the Norfolk Station have been digitised and made available on the website of the Morley Agricultural Foundation website. For Wood’s involvement in particular see the report for 1922-23. www.tmaf.co.uk/archived-reports Accessed 1 March 2015.

⁴ Morley Agricultural Foundation Archive, Norfolk Agricultural Station Report:1909-10.

It might seem at first that no two branches of study could be more widely separated than Agriculture and Astronomy. A moment's consideration, however, will show that they have one point in common: both are at the mercy of the weather. The astronomer's measurements come short of absolute accuracy because of a great number of varying atmospheric conditions, each of which is equally likely to make any one result high or low. He has to obviate this unavoidable lack of accuracy by making many independent observations, and taking their average. This is, or should be, the method followed by the agriculturalist. (Wood and Stratton 1910)

Here then is a further goal, C) to manage widely variable environmental conditions, which – it should be noted – is not a problem that is confined solely to sciences in the field, but can be just as important in the lab (Ankeny et al. 2014). In the same article Wood also reflected on the other values of field experiments aside from the collection of new data; “By laying down such local plots and meeting farmers on them to inspect and discuss the results, the staffs of the various institutions have been brought into touch with the agricultural public, and a mutual understanding has resulted.” Wood thereby introduces a fourth goal, D) to achieve a *mutual understanding* with farmers.

Wood was amongst the first to adopt the half-drill strip trialling method, which he began to use in the fields of the Norfolk Station around 1912-13⁵. Section 3 will deal with the origins of the half-drill strip. Before then, trialling has one further goal that can be uncovered by attending to NIAB, established around a decade after NAS, to pursue precisely the kind of varietal trialling work organised their, but on a national scale.

Though NIAB was an independent organisation – established through a mixture of public and private money, and subsequently maintained through government grants – it was deliberately built at the heart of Cambridge University's agricultural science buildings and farms (Berry 2014a; Charnley 2013; Opitz 2011; Palladino 2002; Silvey and Wellington 1997). The genetical research of Rowland Biffen was of particular importance for NIAB's founding, and of course that of the Plant Breeding Institute (PBI) some years earlier, in 1912. There was considerable interaction between NIAB and the University, and also much discussion and agreement on the topic of field trialling. Soon after the Institute opened its doors the Director, W.H. Parker, published a pamphlet stating its aims (Parker 1922). At the top of the list: “To discover by comparison with established Standard forms, the value of new forms of farm plants, and to determine by comprehensive tests under varied conditions of soil and climate the areas in which their cultivation would be beneficial to British Agriculture.”

⁵ Morley Agricultural Foundation Archive, Norfolk Agricultural Station Report:1915-16.

This statement evidences one final goal, E) to compare new varieties directly with older, established, varieties.

2.2 What kind of experiment are field trials?

To answer this question I adopt John Pickstone's ways of knowing (WoK), which is more readily extended to field trialling than are our extant philosophies of experiment. The decision to pursue WoK is partially inspired by the successes had by the likes of Strasser, and more recently Abigail Woods in adapting this model to their own cases. Woods has considered a case all the more similar to the NIAB story told here, as she has focussed on the importance of field expertise and 'learning by doing' for veterinary administrators. Woods writes "Awarding due historical significance to "learning by doing" re-problematizes the question of why experts subsequently turned to the laboratory" (Woods 2013, 474). Following this lead, the present paper awards due historical significance to the epistemic goals and values of field trialling, which re-problematizes the question of why the RCT came to occupy such a preeminent position over experimental design.

WoK also helps us meet the challenge of not assuming that vastly different analytical categories will be needed when moving from laboratory to field places. Pickstone's ideals are dependent only upon minimally specified places and social arrangements, but this is not all that constitutes them, and infiltration or rearrangement of these places and social arrangements need not undermine a way of knowing. By making the ideals of WoK our starting point, we are permitted to follow our historical actors along whatever path they take, noticing when they appear to be indoctrinated into one way of knowing or another, or when different ways of knowing are being used in a complementary or antagonistic way. If these trajectories take in vastly different laboratory places (from CERN to the wet-labs of synthetic biology), vastly different kinds of field places (from the deep ocean to outer space), and commercial or public places, amongst other different kinds of place, then so be it. WoK foregrounds the knower and does not allow natural or social places to straightforwardly determine thought or method, they merely constrain it.

Pickstone identifies three primary WoK; natural history, analysis, and experiment. It might be expected for me to now explain the sense in which the field trials discussed here conform to the experimental WoK. Alas, they do not. Pickstone's experimentalism has a far narrower meaning than the notion of experiment typically used elsewhere. The experimentalist WoK "stresses the relation between experiment and systematised invention; it concentrates on the creation and control of novelty." (Pickstone 2000, 136) On these terms, not only is much of what historians typically refer to as experiment actually something else, but the field trials in this paper – scientific work with no ambition to create and control novelty – lose their experimental status. The same goes for RCTs, which here do not

attempt to create and control novelty. The word experiment has been allowed to describe all manner of scientific activities and practices that require access to the world (or models of the world), sometimes with the caveat that an intervention of some kind has to be performed (by a scientist or by nature), though this intervention can itself be exceedingly minimal and in some cases can amount to a mere change of perspective (Okasha 2011). There will be lots of different ways for scholars to solve these problems for themselves while preserving the things we otherwise find interesting under the umbrella of 'experiment'. (Nor, of course, is experiment the sumtotal of everything interesting about science, far from it, though this should make restriction of our use of the word all the more important).

The problem is solved for our purposes by allowing Pickstone's divesting of much scientific practice, including the trials here, of their commonly accepted status as experimental. RCTs and trialling are *analysis*, which has interesting and important epistemological characteristics of its own. The analytical WoK supposes that the properties of the material world are best explained by deconstructing them into fundamental elements and subsequently making classifications based upon them (Pickstone 2000, 84). The analytical WoK captures perfectly the field trials of interest in this paper. Indeed, in a passage on botany Pickstone even comes close to describing such varietal field trials himself, though his attention is actually on botanical gardens and does not concern varietal differentiation.

Any botanical garden had to find ways of ensuring the growth of plants...but in the eighteenth century this was a kind of craft knowledge; proper botanists were then primarily concerned with surface classifications and natural philosophy. The preferences of plants for certain kinds of soils, climatic regimes, etc., became the material of science only when analytical botanists came to construe plants as organisms in characteristic milieux. (Pickstone 1994, 127)

The analytical WoK shall provide our framing from here on out. If, dear reader, you are only interested in 'true' experiments, or 'real' experiments, then I have deceived you, though at the very least you may now have a much clearer idea of what 'true' experiment might mean, and you also might want to reconsider your own motivations studying it as an independent form of knowledge production.

3. The half-drill strip as analysis

The half-drill strip trialling method, used in NIAB's large flagship trials from the 1920s and throughout the 1930s (after which smaller and increasingly randomised trials become the norm), had been in general development and use for around 15 years before the institute

was founded. William Sealy Gosset began developing the statistical basis to what would eventually become the half-drill strip while working in the laboratories of the Guinness brewery (Dennison and MacDonagh 1998). He first published on the method in 1911, in an appendix to a paper on field trials written by Daniel Hall (Olby 1991) and W.B. Mercer, both of Rothamsted.

...if we are comparing two varieties it is clearly of advantage to arrange the plots in such a way that the yields of both varieties *shall be affected as far as possible by the same causes to as nearly as possible an equal extent*. To do this is it necessary, from what has been said above, to compare together plots which lie side by side and also to make the plots as small as may be practicable and convenient. (Appendix to Mercer and Hall 1911, emphasis added)

His description contains the statistical rationale for the half-drill strip, though the method is still not so named. In this publication Gosset left out a number of practical considerations that would need to be addressed before his method could be applied in a trial. These practicalities were being solved thanks to his long-standing collaboration with another man, one deeply embedded in farming, food processing, and plant breeding.

Decisive support for the half-drill strip at NIAB came from its Council member Edwin Sloper Beaven, a figure well known to historians of British agricultural science (Brassley 2000; Palladino 1994).⁶ His expertise was readily drawn upon, as we see here through his membership on NIAB's Council and his support for the half-drill strip, the trialling method he had himself developed over a number of years with Gosset. Beaven's farming resources in Wiltshire provided Gosset with a practical field context to wrestle with, and together they produced a workable trialling method (Beaven 1909). He had met Gosset through his work with Guinness, who valued Beaven as both a breeder of new barley varieties and maltster. When in the early 1920s NIAB embarked on its first comprehensive varietal trialling programme Beaven was finally prompted to publish the complete method (Beaven 1922a; Beaven 1922b).

In a half-drill strip trial, a seed drill is pulled along the length of the field (by horse or tractor). The seed-box (which sits in the centre and above the drill) is partitioned down the middle, and one half filled with the seed of a variety for trialling, and the other half with seed from the control variety. When the drill reaches the end of the field, it is turned around and brought back parallel to the first strips (just as in normal farming practice). The control and trial variety therefore alternate sides as the drill is pulled up and down the field. This

⁶ In the final year of the Great War Beaven's daughter married T.B. Wood, mentioned in section 2 as amongst the first to adopt Beaven's half-drill strip. UK agricultural science was an intimate business.

simplicity facilitated the easy inclusion of numerous trial varieties, as one could simply change the variety poured into that side of the seed-box. After a minimum of three years repeated trialling across a number of different stations around the country (some attached to agricultural colleges, others lent to NIAB by philanthropic farming folk), the average performance of varieties were reported as a percentage of the control varieties' performance. The clarificatory example given in an account of NIAB's methods explains that "variety "A" may be recorded as $115\% \pm 2.5\%$ of control". This means that variety "A" has yielded 15 per cent. more than control, and all that need be said about the $\pm 2.5\%$ is that if the difference found is greater than three times this Probable Error – in this case it is six times ($2.5 \times 6 = 15$) – such a difference is regarded by the Institute as worthy of serious attention" (Parker 1931). Upon such results recommendations were made.

Seeing plant varieties in these terms, as possessing an average yield, is – I argue – to see them in the terms of *analysis* (Pickstone 2000, 102). These trials were taken to reveal something essential about the hereditary capacities of these varieties. After all, the broadest goal of these trials was to eventually release information to farmers regarding their differential performances in order to influence future seed sales. Recognising field trials as conforming to the analytical WoK is important for two reasons. Firstly, doing so embeds NIAB's activities in a far wider and longer history of science. We can for instance recognise that NIAB's botanical activities were directly inspired by analytic chemistry, which is itself the most essential touchstone for the ascendancy of analysis in Pickstone's account. Gosset's analytical work in the brewery laboratories of Guinness are good evidence for this, while the debt to chemistry was even made explicit at the time of NIAB's formation. Rothamsted – one of the most important centres for analytical chemistry in the UK – was considered an inspiration to the Institute's founders. In the memorandum written to build support for NIAB, it was stated that the new Institute would "bring scientific botany to bear on the improvement of the plant in the same way that Rothamsted has revolutionised the treatment of the soil by chemistry".⁷ (This debt will become ironic in the following section.) The genetic context of NIAB's work also helps further evidence the brave new analytical world in botany, as the research programme of key geneticists such as Wilhelm Johannsen (who coined that elemental term 'gene'), reduced collections of plants down to individual pure lines in precisely the way demanded by the analytical WoK (Müller-Wille 2007). Certainly the initial creation and control of new lines conforms to the experimental WoK, but once novelties stopped being sought and aberrant plant forms instead begin to become a problem requiring 'roguing' (removing from the field and destroying), then experiment has ended, and plant varieties are put into the analytical hands of field triallers, who can work out what each is

⁷ National Archives, MAF 33/22 'Memorandum on the Establishment of a National Institute of Agricultural Botany' January 1918.

really made of. This is all the more true as some believed that the process of creating 'purer' varieties actually dehistoricized these plants (Berry 2014b), erasing their natural history just as we would expect in the analytical WoK. The second reason for seeing field trialling as analysis will become apparent when we turn to the RCT.

Before moving on, it remains to highlight the ways in which the half-drill strip achieved the epistemic and social goals listed in the first half of the paper. Firstly, the half-drill strip obviously achieves goal A, to find the varieties most suitable for a geographic area, assessed according to characters such as disease resistance, frost tolerance, and yield, because it was a method that could be picked up and applied with only a little more attention and supervision than normal farming practice. The half-drill strip also achieves goal B, to make maximal use of all the new varieties 'now being so *rapidly* introduced'. However, speed is a relative term, and as we shall see later, when the number of new varieties increased even further, NIAB was forced to rethink its trialling methodology. As for goal C, the half-drill strip was explicitly designed to manage widely variable environmental conditions. In one of Gosset's later articles on varietal differentiation, published in *Biometrika*, he writes:

The peculiar difficulties of the problem [variety testing] lie in the fact that the soil in which the experiments are to be carried out is nowhere really uniform; however little it may vary to the eye, it is found to vary not only from acre to acre but from yard to yard, and even from inch to inch. This variation is anything but random, so that the ordinary formulae for combining errors of observation which are based on randomness are even less applicable than usual. (Student 1923, 272).

Goal D, to achieve a *mutual understanding* with farmers, was a particularly important goal that the half-drill strip facilitated, because the strips of varieties that it produced on the field scale were readily recognisable to farmers. Gosset also appreciated the importance of this goal, as we see in the following comparison of small and large scale trials. "Taking first the large scale, it has the advantage that the farmer, who always has a healthy contempt for gardening, may pay some attention to the results; he is to this extent right, that large scale conditions cannot be accurately reproduced in a wire cage". Later in the same article he explains "By means of Beaven's "half-drill strip" method...This combines the advantage of growing corn on the large scale with an accuracy almost as great as that of small scale work; and is within the powers of anyone who can combine the necessary knowledge and patience with the control of skilled agricultural labour" (Student 1923). Lastly goal E is clearly directly driven towards by the half-drill strip, as the control varieties selected were always old established forms or existing varieties commonly grown in the local region, used to provide a

familiar comparison point to cutting-edge varieties. By demonstrating directly to farmers the capacities of the new in comparison to the old, it was hoped that trials could operate as a ratchet on national agricultural productivity.

Despite its lengthy development, and association with some of the most important figures in agricultural science at this time, the half-drill strip found itself undermined almost as soon as it was launched in national trials.

4. The RCT as hyper-analysis

By hyper-analysis I mean that the RCT privileges analysis to the point that it excludes epistemic goals and values that do not contribute directly to its own analytical method. The half-drill strip on the other hand, while still being an analytical method, can accommodate a wider variety of epistemic goals and values. ‘Hyper-analysis’ only makes sense in this comparative context and is not meant as an augmentation of WoK. Pickstone explains that in “all practical activities there will be a tension between the claims of formal analysis and those of ‘experience’ which are hard to formulate; we see such tensions in industry and agriculture as well as in medicine” (Pickstone 2000, p. 114). This description captures perfectly what was at stake with the emergence of the RCT. This is the second (and weaker) reason to see field trialling as an instance of the analytical WoK, because otherwise the close fit of Pickstone’s general description to our case would have to be dismissed as coincidental.

Fisher developed his appreciation for randomisation by attending to experimental work conducted in both the fields *and* laboratories of Rothamsted (Parolini 2013, 82). First explaining his requirement of randomisation in his 1925 *Statistical Methods for Research Workers*, Fisher compared randomised to systematic arrangements (like the half-drill strip), though he gave no specific examples (Fisher 1925). In a subsequent article he took his argument further, scrutinising trials that attempted to differentiate between varieties (precisely as NIAB had been doing for a number of years by this time). “Only a minority of field experiments are of the simple type, typified by variety trials, in which all possible comparisons are of equal importance. In most experiments involving manuring or cultural treatment, the comparisons involving single factors, *e.g.*, with or without phosphate, are of far higher interest and practical importance than the much more numerous possible comparisons involving several factors” (Fisher 1926). Speaking as a representative of Rothamsted, famed for its decades of research into soil and manurial treatments, one might sense here a certain amount of institutional posturing.⁸

⁸ Parolini disagrees and considers it just a matter of statistical good sense. I prefer the dramatic interpretation. This footnote now signals my potential error. My thanks to Parolini for pointing it out to me!

Randomisation produced field arrangements that would be held with considerable suspicion by those with knowledge of how variable field conditions can be. Fisher certainly recognised the variability of field sites, but sort to eliminate variability as a problem by ensuring any results derived from a trial were above a certain threshold of statistical significance (a true test of which could only be ensured by randomisation). This meant abstracting away from the particular field in question, only incorporating such knowledge of its specific variability *after* the trial had been carried out and auxiliary hypothesis brought in to explain anomalies. This kind of abstraction was also key to his ongoing efforts toward ending the disagreement between biometricians and Mendelians at this time (Morrison 2002). Addressing this issue through an example, Fisher preempted those who would be concerned. “Note what a “bad” distribution chance often supplies; the chloride plots are all bunched together in the middle of the first block, while they form a solid band across the top block on the right; in the bottom block on the right, too, all the early plots are on one side, and all the late plots on the other” (Fisher 1926). A ‘bad’ distribution was the price to be paid for high levels of certainty. In making this argument Fisher placed his statistics in a preeminent position over experimental design and robbed trials of important social and epistemological roles.

First to publicly challenge randomisation was Frank Engledow, agricultural geneticist at the University and a member of NIAB’s Crop Improvement Committee. In a two-part article co-authored with his Cambridge colleague George Udny Yule (geneticist and statistician), Engledow set out both the proper statistical basis to trialling and the practical difficulties faced by experimenters attempting to ensure their reliability. They then turn to the question of randomisation. While explaining in a perfectly satisfactory way how their example could be made to conform to the demands of randomisation, they then write:

We do not propose to follow these theoretical questions further for the reason that their importance appears to us nullified by practical considerations. The simplicity attaching to a repeating form of scatter [as used in their example trial]...carries two solid advantages. It facilitates sowing, observation of the growing plots, and harvest. Anything which facilitates work in careful yield trials at such periods is of great value. Sowing randomised plots would call for constant reshuffling of the seed packets if the plots were, as is usual, sown one after another from beginning to end of the whole series. It might be avoided by sowing all the A plots wherever they might be, and so on, but for this the sower would have to tramp over his tilth again and again, and so poach it. A second advantage of simplicity is insurance against mistakes. In any considerable piece of plot work there is great risk of mistakes in the “damn fool” order. An A label on a B sheaf may be more upsetting to results than the adoption, in

the face of theoretical objection, of a systematically repeating pattern of plots.
(Engledow and Yule 1926)

They instead recommended an alternative. “Dr E.S. Beaven of Warminster has devised the “Half-Drill-Strip” method, which very ingeniously overcomes the kind of difficulty we have considered...We commend it to all concerned with field-scale trials.” Researchers based in laboratories and fields seek to protect their experiments from wasteful impracticalities that would cast doubt over the *reliability* of their results. How one decides this issue (the extent to which tests of statistical significance should be held as more valuable than any other epistemic consideration) will decide which of the two methods achieves goals A, D and E best. The only goal that the RCT can immediately serve better than the half-drill strip is B, as RCT plots take up less space, allowing for more varieties to be dealt with more quickly. (Though obviously in the eyes of those who do not privilege randomisation, this speed comes at the cost of reliability). Nor could either method claim superiority over the other when it came to goal C, the management of widely variable conditions, because incommensurate epistemic criteria are being appealed to, so that it is not even possible to agree on the most appropriate means of solving the debate (Hicks 2015). Fortunately (and inevitably) argument spilled over into other concerns.

Secondly, and relatedly, the demand for randomisation narrowed the kinds of locations in which field experiments could be expected to take place, as they required high levels of supervision and trialling expertise. That those who collaborated with NIAB at agricultural colleges and regional centres across the country could readily adopt the half-drill strip was of considerable importance for those who relied on external partners to provide trialling resources. “I have discussed with Sir Daniel Hall the question of the methods of trial and at his suggestion have also talked over the matter with Dr. Fisher of Rothamsted” wrote Parker in 1930.

As an outcome of these discussions I remain of the opinion...[that the present methods of yield testing]...are the best of any of the methods in which ordered, as distinct from random, distribution is practised. The adoption of random distribution on a field scale is ruled out by the absence of suitable implements, by the greater technical skill required in handling and by the great increase in statistical work that they entail.⁹

⁹ NIAB, C-1, Council Papers, No. 103, ‘National Institute of Agricultural Botany, Methods of Trial’, dated June 1930.

The national information gathered through a trial method that could be conducted virtually anywhere was much more important than any putative certainty lent to the results through randomisation. The half-drill strip is more *imitable* for a wider array of would-be triallers, a characteristic that can be understood as a form of replicability.¹⁰ That the half-drill strip was more imitable than the RCT meant that it could deliver goal A better than the RCT (though remember again of course, ‘better’ requires us to fall on one side or the other of *reliability*), because it could be applied more readily across the entire country. In laboratory and field research alike, the epistemological goals of researchers (in this instance, learning about varietal characters *under a wide variety of conditions* in collaboration with farmers who *owned the resources* that could enable such an investigation) directly influenced the methods adopted (Ankeny and Leonelli 2011).

Lastly, growing field trials in a manner that was essentially ‘farm-like’ was important not only for maintaining strong relations between scientists and farmers, but so as to enable the expert farmer and plant scientist to bring their skills to bear on varietal analysis. Just as Fisher argued randomisation ensured truly important results could be identified (on his terms this meant those that passed a true test of statistical significance, ensured by randomisation), so too did the opponents of randomisation believe they were ensuring truly important results could be identified. In a draft of the 11th Report of the Council, eventually circulated in January 1931, Parker writes

Indeed many a farmer cannot grasp the difference that choice of variety may make to his own results until he has seen a set of these trials. It is all the more important that the trials should be under careful cultivation and observation throughout their course; no amount of statistical analysis can compensate for errors in the field or discover those practical points of difference measurable by eye alone...¹¹

It is not clear in this quote whether the ‘practical points of difference’ refer to the plants themselves or the conditions of the field, but perhaps there is no reason to think of the two as separable? (Recall Gosset’s earlier comments about the limited value of small cage plots.) Only by growing plants on the field scale could one hope to find *significant novelty*, or an understanding of a varieties’ true nature. Such ‘trained judgment’ was perfectly defensible within and without the laboratory at this time (Daston and Galison 2007). Another point to make here, is that one should think of field trials as more or less ‘readable’, in the same way

¹⁰ In an earlier draft I had called this ‘replicability’ outright. In discussing this paper with Matt Paskins however, he suggested that *imitability* would be more accurate, and I have adopted his suggestion (for which he has my sincere thanks).

¹¹ NIAB, C-1, Council Papers, No. 106, ‘Draft: National Institute of Agricultural Botany, Eleventh Report of the Council’, dated Jan 1931.

other forms of data representation are. Scientists can choose between many ways to display their data (graphs, maps, tables, and so on), some of which are going to be more exclusive in the kinds of readership that can access them, and others more inclusive. These choices are an excellent way for a scientist to indicate the kind of epistemological status they are claiming for themselves in relation to the public or other experts (Tufte 2001).¹² Inclusive and exclusive forms of representation can also be an effective way for scientists to make intellectual property (broad) claims over a domain of knowledge and practice (MacLeod and Radick 2013). To put the argument clearly; the half-drill strip resulted in field trials that were more accessible to a non-statistical audience, who could read such fields much as they would those on their own farm, while the RCT made field trials readable almost solely to the statistician. Once we recognise that these skills matter for varietal differentiation and analysis, then once again the half-drill strip out-competes the RCT, this time in achieving goals D and E.

5. Conclusion

We have seen how a form of *analysis* that was not prejudicial toward the expertise of experienced farmers and agricultural botanists (the half-drill strip) was initially preferred, over and above the purported certainty of a *hyper-analytical* method (the RCT) which *was* highly prejudicial to those who wished to assess varieties on a field scale by their expert eye. Field triallers in Cambridge did not immediately seize upon the RCT in their varietal differentiation trials, because it did not achieve the epistemic and social goals of field trialling to the same extent as the half-drill strip, and because the RCT was actually prejudicial to certain key epistemic goals and values. Thanks to the use of WoK, this case study can be offered up as an example of the long history of the analytical WoK being refined, in this instance in order expand the statistician's hold over agricultural science and industry. The half drill-strip remained widely in use throughout the 1930s, though by the early part of the decade it was beginning to be referred to as merely a good compromise. "The Institute's methods of trial have been thoroughly reviewed and discussed with other statistical authorities. It is not always possible, for physical reasons, to reconcile practical farming with statistical requirements, but there is no doubt that Dr. Beaven's half-drill strip system as used by the Institute, is a compromise, ensures the highest practicable standard of accuracy".¹³ How and why RCTs came to win the day is a story for another time, but some lines of future investigation are worth speculating on.

¹² My thanks to Charlotte Sleight who recommended this important work to me when I panicked on twitter.

¹³ NA, C-8.60, 11th Annual Report (1929-1930).

We know that when something like analysis or any other WoK climbs a little higher, it is because the kinds of place and the social arrangements in which scientific work is conducted have changed, or because the WoK in question has been used to engineer such a new arrangement. “We can argue that political and social changes allowed the elevation of such knowledge, either through the empowerment of protagonists and/or because it became easier and more attractive for already powerful groups to appropriate the knowledge/practices” (Pickstone 1994, 127). As I have argued elsewhere (Berry 2014c) plant breeding geneticists built power for themselves around genes and what these genes could deliver for society and industry in just this sort of maneuver. Statisticians give the appearance of going through a similar process, and more interestingly, perhaps also entered into a bootstrapping operation with genetics (but this is just speculation). My suggestion is that geneticists and statisticians – whom we know became increasingly integrated throughout the century thanks in part to the kind of work in population genetics pioneered by Fisher – found that they could command greater power by increasingly embracing analysis, to the greater exclusion of other, more experience based, forms of knowledge. An alternative or complementary hypothesis, is that the RCT may have come to prominence because it can ease the strain on over-stretched resources, such as seed supply or land. The smaller land requirements of the RCT were mentioned earlier, with regard to the RCT’s ability to achieve goal B more readily than the half-drill strip. For evidence one could point to NIAB’s 12th Annual Report (1930-1931), which explained that ‘chequer-board’ trials (more similar in their layout to a randomised block than the half-drill strip) were going to be more heavily used, in order to speed up the number of varieties that might be tested for yield in any one year.¹⁴ Obviously this comes at the cost of large field scale trialling and the epistemic goals and values that they could embody, which in turn suggests the relations between farmers and scientists fieldside were also undergoing a transformation.

Lastly, refusing the assumption that we will need to think of the field in vastly different terms to the laboratory has been highly productive. Not only do we thereby avoid conflating places and practices, but we can all the better appreciate how WoK work across, upwards, and downwards through places and people. The categories of ‘lab ideal’ and ‘field ideal’ do more harm than good, even when – as is often the case – they are problematised by historians and sociologists who want to bring attention to the places and practices ‘in between’, which get classed as messy, or hybrid, or even dirty (Kingsland 2009; Kohler 2012; Schwarz 2014). If we drop the initial idea of a controlled lab place and a wild field place, we can see that it is then inappropriate to be surprised when analysers and experimenters are working through a whole host of different places and epistemological

¹⁴ NIAB, C-8.60, 12th Annual Report, 1930-1931.

strategies, and we do not risk classing certain epistemic values as 'absent' from a place, when really we should be alert to the variety of different forms that adherence to such a value might take.

Bibliography

Ankeny, R.A. & Leonelli, S. (2011). What's so special about model organisms? *Studies in History and Philosophy of Science*, 42, 313-323.

Ankeny, R.A., Leonelli, S, Nelson, N.C., & Ramsden, E. (2014). Making organisms model human behaviour: Situated models in North-American alcohol research, since 1950. *Science in Context*, 27, 485-509.

Armitage, P. (2003). Fisher, Bradford Hill, and randomization. *International Epidemiological Association*, 32, 925-928.

Beaven, E.S. (1909). Pedigree seed corn. *Journal of the Royal Agricultural Society of England*, 70, 119-139.

Beaven, E.S. (1922a). Trials of new varieties of cereals – Part 1. *Journal of the Ministry of Agriculture*, 29, 337-347.

Beaven, E.S. (1922b). Trials of new varieties of cereals – Part 2. *Journal of the Ministry of Agriculture*, 29, 433-444.

Berry, D. (2014a). *Genetics, Statistics, and Regulation at the National Institute of Agricultural Botany, 1919-1969*. University of Leeds: Unpublished PhD thesis.

Berry, D. (2014b). The plant breeding industry after pure line theory: Lessons from the National Institute of Agricultural Botany. *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 46, 25-37.

Berry, D. (2014c). Bruno to Brunn; or the *Pasteurization* of Mendelian genetics. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 48, 280-286.

Berry, D. (2015). Agricultural Modernity as a Product of the Great War: The Founding of the Official Seed Testing Station for England and Wales, 1917–1921. *War & Society*, 34, 121-139.

Brandreth, B. (1935). County Spring oat trials, 1931-33. *Journal of the National Institute of Agricultural Botany*, 4, 5-25.

Brassley, P. (2000). Crop Varieties. In Collins, E.J.T. (ed.), *The Agrarian History of England and Wales, Volume VII 1850-1914*. Cambridge: Cambridge University Press.

Bonneuil, C. (2006). Mendelism, plant breeding and experimental cultures: Agriculture and the development of genetics in France. *Journal of the History of Biology*, 39, 281-308.

- Box, J.F. (1978). *R.A. Fisher The Life of a Scientist*. New York: John Wiley & Sons.
- Cartwright, N. (2007). Are RCT's the gold standard?. *BioSocieties*, 2, 11-20.
- Cartwright, N. & Munro, E. (2010). The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16, 260-266.
- Chang, H. (2012). *Is Water H₂O?: Evidence, Realism and Pluralism*. Dordrecht: Springer.
- Charnley, B. (2011). *Agricultural Science, Plant Breeding and the Emergence of a Mendelian System in Britain, 1880-1930*. University of Leeds: Unpublished PhD thesis.
- Charnley, B. (2013). Experiments in empire-building: Mendelian genetics as a national, imperial, and global agricultural enterprise. *Studies in History and Philosophy of Science*, 44, 292-300.
- Charnley, B. and Radick, G. (2013). Intellectual property, plant breeding and the making of Mendelian genetics. *Studies in History and Philosophy of Science*, 44, 222-233.
- Daston, L. & Galison, P. (2007). *Objectivity*. Cambridge, MA: MIT Press.
- Dennison, S.R. & MacDonagh, O. (1998). *Guinness 1886-1939: From Incorporation to the Second World War*. Cork: Cork University Press.
- Diser, L. (2012). Laboratory science in Belgian agriculture at the end of the nineteenth century. *Agricultural History*, 86, 31-54.
- Elliott, K.C. (2011). Direct and Indirect Roles for Values in Science. *Philosophy of Science*, 78, 303-324.
- Elliott, K.C. & McKaughan (2014). Nonepistemic Values and the Multiple Goals of Science. *Philosophy of Science*, 81, 1-21.
- Engledow, F. & Yule, G.U. (1926). The principles and practice of yield trials. *Empire Cotton Growing Review*, 3, 235-267.
- Finnegan, D.A. (2008). The spatial turn: Geographical approaches in the history of science. *Journal of the History of Biology*, 41, 369-388.
- Fisher, R.A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- Fisher, R.A. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture*, 33, 503-515.

- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J. & Kruger, L. (1989). *The Empire of Chance: How probability changed science and everyday life*. Cambridge: Cambridge University Press.
- Gooday, G. (2008). Placing or replacing the laboratory in the history of science? *Isis*, 99, 783-795.
- Gooding, D., Pinch, T., & Schaffer, S. (1989). *The uses of experiment*. Cambridge: Cambridge University Press.
- Gower, J.C. (1988). Statistics and Agriculture. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 151, 179-200.
- Hacking, I. (1983). *Representing and Intervening: introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Hacking, I. (1988). Telepathy: Origins of randomization in experimental design. *Isis*, 79, 427-451.
- Harwood, J. (2012). *Europe's Green Revolution and Others Since: The rise and fall of peasant-friendly plant breeding*. London and New York: Routledge.
- Hall, N.S. (2002). *R.A. Fisher and Randomized Experimental Design*. University of Maryland: Unpublished PhD thesis.
- Hall, N.S. (2007). R.A. Fisher and his advocacy of randomization. *Journal of the History of Biology*, 40, 295-325.
- Henke, C.R. (2008). *Cultivating Science, Harvesting Power: Science and Industrial Agriculture in California*. Cambridge, MA: The MIT Press.
- Hicks, D.J. (2015). Epistemological depth in a GM crops controversy. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 50, 1-12.
- Hutchinson, J. & Owers, A.C. (1980). *Change and Innovation in Norfolk Farming*. Norfolk: Packard Publishing.
- Iori, L. (2013). *Agricultural Genetics and Plant Breeding in Early Twentieth-Century Italy*. University of Bologna: Unpublished PhD thesis.
- Jonsson, F.A. (2013). *Enlightenment's Frontier: The Scottish Highlands and the Origins of Environmentalism*. New Haven and London: Yale University Press.
- Kingsland, S. (2009). Frit Went's atomic age greenhouse: The changing labscape on the lab-field border. *Journal of the History of Biology*, 42, 289-324.
- Kohler, R.E. (2002). *Landscapes and Labscapes: Exploring the Lab-Field Border in Biology*. Chicago: The University of Chicago Press.

Kohler, R.E. (2012). Practice and place in twentieth-century field biology: A comment. *Journal of the History of Biology*, 45, 579-586.

Latour, B. (1988). *The Pasteurization of France*. London and Cambridge, MA: Harvard University Press.

Leonelli, S. (2013). Integrating data to acquire new knowledge: Three modes of integration in plant science. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 503-514.

Livingstone, D. (2010). Landscapes of knowledge. In P. Meusburger, D. Livingstone, & J. Jöns (Eds.), *Geographies of Science* (pp. 3-22). New York: Springer.

Maat, H. (2001). *Cultivating Practice: A history of Agricultural Science in the Netherlands and its Colonies 1863-1986*. Dordrecht: Kluwer.

Maat, H. (2011). The history and future of agricultural experiments. *NJAS – Wageningen Journal of Life Sciences*, 57, 187-195.

MacLeod, C. & Radick, G. (2013). Claiming ownership in the technosciences: Patents, priority and productivity. *Studies in History and Philosophy of Science*, 44, 188-201.

McCloskey, D.N. & Ziliak, S.T. (2008). *The Cult of Statistical Significance*. Michigan: University of Michigan Press.

McMullen, L. (1939). "Student" as a Man. *Biometrika*, 30, 205-210.

Mercer, W.B. & Hall, D. (1911). The experimental error of field trials. *The Journal of Agricultural Science*, 4, 107-132.

Morgan, M.S. (2013). Nature's Experiments and Natural Experiments in the Social Sciences. *Philosophy of the Social Sciences*, 43, 341-357.

Morrison, M. (2002). Modelling populations: Pearson and Fisher on Mendelism and Biometry. *The British Journal for the Philosophy of Science*, 53, 39-68.

Müller-Wille, S. (2005). Early Mendelism and the subversion of taxonomy: epistemological obstacles as institutions. *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 36, 465-487.

Müller-Wille, S. (2007). Hybrids, pure cultures, and pure lines: from nineteenth-century biology to twentieth century genetics. *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 38, 796-806.

Naylor, S. (2005). Introduction: Historical Geographies of Science: Places, Contexts, Cartographies. *The British Journal for the History of Science*. 38, 1-12.

- Okasha, S. (2011). Experiment, observation and the confirmation of laws. *Analysis*, 71, 222-232.
- Olby, R. (1991). Social imperialism and state support for agricultural research in Edwardian Britain. *Annals of Science*, 48, 509-526.
- Opitz, D. (2011). Cultivating genetics in the country: Whittingehame Lodge, Cambridge. In D.N. Livingstone & C.W.J. Withers (Eds.), *Geographies of Nineteenth-Century Science* (pp. 73-98). Chicago: The University of Chicago Press.
- Palladino, P. (1994). Wizards and devotees: on the Mendelian theory of inheritance and the professionalization of agricultural science in Great Britain and the United State, 1880-1930. *History of Science*, 32, 409-444.
- Palladino, P. (2002). *Plants, patients and the historian*. Manchester: Manchester University Press.
- Parker, W.H. (1922). *The Testing, Multiplication and Distribution of New Forms of Farm Crops*. Cambridge: The National Institute of Agricultural Botany.
- Parker, W.H. (1931). The methods employed in variety trials by the National Institute of Agricultural Botany. *Journal of the National Institute of Agricultural Botany*, 3, 5-9.
- Parolini, G. (2013). *"Making Sense of Figures": Statistics, Computing and Information Technologies in Agriculture and Biology in Britain, 1920s-1960s*. University of Bologna: Unpublished PhD thesis.
- Parolini, G. (2014). The Emergence of Modern Statistics in Agricultural Science: Analysis of Variance, Experimental Design and the Reshaping of Research at Rothamsted Experimental Station, 1919–1933. *Journal of the History of Biology*, DOI: 10.1007/s10739-014-9394-z.
- Pearson, E.S. (1939). "Student" as a statistician. *Biometrika*, 30, 210-250.
- Pickstone, J.V. (1994). Museological science? The Place of the Analytical/Comparative in Nineteenth-Century Science, Technology and Medicine. *History of Science*, 32, 113-138.
- Pickstone, J.V. (2000). *Ways of Knowing: a new history of science, technology and medicine*. Manchester, Manchester University Press.
- Pickstone, J.V. (2007). Working knowledges before and after circa 1800: Practices and disciplines in the history of science, technology, and medicine. *Isis*, 98, 489-516.
- Pickstone, J.V. (2011). A brief introduction to ways of knowing and ways of working. *History of Science*, 49, 235-245.
- Plackett, R.L. & Barnard, G.A. (Eds.) (1990). *'Student' A Statistical Biography of William Sealy Gosset*. Oxford: Clarendon Press.

- Rheinberger, H.-J. (1997). *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.
- Riggs, W.D. (2003). Balancing our Epistemic Goals. *Noûs*, 37, 342-352.
- Russell, E.J. (1930). Thomas Barlow Wood (1869-1929). *Biochemical Journal*, 24, 1-3.
- Schaffer, S. (2003). Enlightenment brought down to Earth. *History of Science*, 41, 257-268.
- Schwarz, A. E. (2014). *Experiments in Practice*. London: Pickering & Chatto.
- Silvey, V. & Wellington, P.S. (1997). *Crop and Seed Improvement: A history of the National Institute of Agricultural Botany 1919-1996*. Cambridge: National Institute of Agricultural Botany.
- Steel, D. (2010). Epistemic Values and the Argument from Inductive Risk. *Philosophy of Science*, 14-34.
- Strasser, B. (2012). Collecting Nature: Practices, Styles, and Narratives. *Osiris*, 27, 303-340.
- Student. (1923). On Testing Varieties of Cereals. *Biometrika*, 15, 271-293.
- Tufte, E.R. (2001). *The Visual Display of Quantitative Information*. Cheshire, Conn: Graphics Press, 2nd edition.
- Wieland, T. (2006). Scientific theory and agricultural practice: Plant breeding in Germany in the late 19th to the early 20th century. *Journal of the History of Biology*, 39, 309-343.
- Wood, T.B. & Stratton, F.J.M. (1910). The interpretation of experimental results. *The Journal of Agricultural Science*, 3, 417-440.
- Woods, A. (2013). From Practical Men to Scientific Experts: British Veterinary Surgeons and the Development of Government Scientific Expertise, c. 1878–1919. *History of Science*, 51, 457-480.