Comparative Experiments with Field Crops

G.V. Dyke

COMPARATIVE EXPERIMENTS WITH FIELD CROPS

G.V. DYKE, M.A. (OXON.),
Head of the Field Experiments Section, Rothamsted

LONDON BUTTERWORTHS

THE BUTTERWORTH GROUP

ENGLAND
Butterworth & Co (Publishers) Ltd
London: 88 Kingsway, WC2B 6AB

AUSTRALIA
Butterworths Pty Ltd
Sydney: 586 Pacific Highway, NSW 2067
Melbourne: 343 Little Collins Street, 3000
Brisbane: 240 Queen Street, 4000

CANADA
Butterworth & Co (Canada) Ltd
Toronto: 14 Curity Avenue, 374

NEW ZEALAND
Butterworths of New Zealand Ltd
Wellington: 26–28 Waring Taylor Street, 1

SOUTH AFRICA
Butterworth & Co (South Africa) (Pty) Ltd
Durban: 152–154 Gale Street

First published 1974

ISBN 0 408 70554 X

Text set in 10/11 pt. IBM Press Roman, printed by photolithography, and bound in Great Britain at The Pitman Press, Bath

PREFACE

This book is written for the man (or woman) who is going to do field experiments for a living, or for fun, or both (I am of this last class). Parts of it are written to help him or her in the interpretation of the results of field experiments. None of it, least of all the chapters on statistics, is written for statisticians. In writing these chapters I have in mind the man so far out in the bush that day-to-day statistics is done by him, or not at all; I have also thought of the man who is diffident about asking his learned statistical colleagues 'what is a degree of freedom?'

Publications that deal with the special problems of experiments with grassland and perennial fruit crops are listed in the bibliography.

If the style is at times light-hearted, I offer no apology; I have enjoyed my work in field experiments and if this is evident, so be it.

This book is loaded with my prejudices; that is why I have so much enjoyed writing it. But I have tried to distinguish my own sometimes controversial opinions by introducing them with 'I think' or some such use of the first person singular. Sentences written in the third person contain only what is in the author's opinion more or less generally accepted doctrine.

I do not want to dress up field experimentation as a science (let's keep 'agronomy' out of this argument) but I think it perhaps deserves recognition as a distinct bit of technology. I have not tried to string together a lot of recipes telling you how to concoct the perfect experiment for every set of circumstances – indeed I have sometimes been at pains *not* to do so. I have tried, by giving examples, to stimulate you to think actively about how you do experiments in *your* circumstances – which are almost certainly different in some material respect from mine. On the other hand where it seems appropriate I have summed up and given a few rules of thumb.

There is a glossary of technical terms on page 176; half the battle of creating a new subject is won when a jargon has been established and I want to do my bit. Each technical term is set in italic where it first appears in the text.

Xaghra, Gozo

G. V. Dyke

21 February 1970

ACKNOWLEDGEMENTS

I am grateful to many colleagues and ex-colleagues at Rothamsted for goading me into writing this book and helping me with patience and criticism; Harold Garner and Jim McEwen have suffered most. Mohamed Nour lit the fuse by inviting me to lecture on field experiments in the Faculty of Agriculture of the University of Khartoum. Messrs Hunting Technical Services Ltd. gave me two excellent exposures to hot-weather problems. Finally, I am grateful to Rothamsted – a tolerant, liberal, lively community to which I have been privileged to belong all my working life.

CONTENTS

Part I How to Do and Interpret Field Experiments

Part II Statistics in Field Experiments

Part I

How to Do and Interpret Field Experiments

Chapter One

FIELD EXPERIMENTS IN AGRICULTURAL RESEARCH

1.1 DEFINITION OF A FIELD EXPERIMENT

It is best to consider that whenever two or more parts of a field receive different *treatments** then a field *experiment* has been started. Most experiments have many more than two differently-treated areas ('*plots*') but if very many experiments are done to compare the same treatments useful results can be obtained from experiments with very few (even as few as two) plots in each.

Going back a stage further, we should perhaps define a field. (In parts of England and many other countries a 'field' seems an obvious, more or less permanent entity, bounded by walls, hedges or fences, but elsewhere a 'field' is by no means so simple to define.) For the purpose of this book a reasonable working definition is:

> **A field is a piece of land that has been uniformly cultivated, manured, cropped, etc., in each of the last (say) three years; or failing uniformity, the boundaries between areas with different histories must be known and taken into account in planning an experiment.**

In some circumstances three years are not enough to give even a tolerable semblance of uniformity; for example, on soil naturally poor in available phosphorus (P) at Rothamsted heavy applications of superphosphate made in the past still cause large differences in the YIELD

*Words that are included in the Glossary are printed in italic at their first appearance.

of crops after 70 years. A field where such applications have been made on some areas, not on others, cannot be regarded as an acceptable site for an experiment unless the pattern of the old treatments is known and is taken into account in the *design* and *layout* of the present experiment.

My definition does not exclude 'accidental' experiments. If, for example, a field is being ploughed and the work is interrupted by bad weather, it is possible that the areas ploughed before and after the delay will give different yields of the crop grown in the next (or a later) year. You may get useful information from such an 'experiment', crude and fortuitous as it may be. Looking at the other side of the coin, such an accidental application of two or more treatments should be borne in mind and allowed for when laying out an experiment on the site. Methods of doing this will be discussed in Chapter 2.

1.2 EXPERIMENTS AND SURVEYS

Surveys can provide excellent information about growers' methods and the yields they obtain but they give little information about the effect of differences in methods. Growers who sow better seed may also give better cultivations and a survey will not reveal how much increase in yield is due to each difference separately. A well-planned series of experiments gives much *information* about the changes in yield that will occur if growers change their methods in certain ways. The functions of surveys and experiments are therefore complementary.

1.3 VALIDITY OF EXPERIMENTS

An ideal set of field experiments on one particular subject would be done on many sites and in several seasons. The sites should be a random selection from the whole area of the crop that is being investigated (e.g. all cotton in the Gezira or all wheat on chalk soils in Hampshire). We have to assume that variations of weather between seasons is random. Few people (except perhaps in India) have approached this ideal but we should recognise the limitations of the experiments we do. Growers' methods change as the years pass and the results of a series of experiments done in 1959–1962 may by 1969 be invalid for most of the area of the crop. For instance new varieties of wheat may have stiffer straw and will *respond* to more nitrogen fertiliser than old varieties.

An exception to the above rule: it is legitimate to select special conditions for some experiments. For example, if we wish to compare two forms of fertiliser containing P we do experiments *only* on soils with little available P where P-responses may be expected to be large.

But we *must* state how the sites were chosen in giving the results.

As farming practices improve the experimenter turns his attention from *factors* that have large effects (perhaps doubling the yield) to factors whose effects are relatively small (5% or 10% perhaps). If yields have greatly increased meanwhile the absolute effect (e.g. in kg per hectare) may be about the same. But more refined experiments, with more *replication*, will be needed.

I have deliberately used different units of yield more or less at random in this book – kg per hectare, cwt per acre, and so on. I think an experimenter worth his salt should be ready to use the units most convenient in the immediate circumstances; when in Rome measure yields as the Romans do. I have a respected colleague who records nutrient dressings in grammes per square yard – and why not? (Her balance is graduated in grammes, her measuring tape in feet and inches.)

Field experiments are sometimes done, not to assess the effect on yield of changes in practices, but (for example) to find cheaper ways of achieving the same yields, to compare the efficiency of different systems of draining wet land, or to investigate a matter of theoretical interest such as the availability of P applied in fertiliser many years ago. But in most cases many of the considerations mentioned above still apply.

1.4 THE CHOOSING OF TREATMENTS

Although I am not trying in this book to tell you how to choose the treatments to be included in any particular experiment (this would presume a knowledge of your problems and circumstances that I do not have) a few points are worth making.

Most experiments start from some fairly simple question – 'is variety A better than variety B?' or 'will nitrogen (N) fertiliser give a profitable increase in yield?' – and the experimenter thinks of a correspondingly simple set of treatments. If he thinks no more but does an experiment with these treatments (adequately designed and executed) he will get a simple answer, accurate within the limits set by the intrinsic variability of the site used. But the full truth may be more complicated – variety A may yield more than B in the absence of mildew but, because it is exceptionally susceptible to mildew, it may yield much less than B when there is mildew; if no K is applied a single dressing of N may be profitable but doubling the dressing may lessen the profit whereas if adequate K is applied the double dressing of N may be worth while. In such cases the narrow cross-section of the truth that is given by one simple experiment is not enough; the whole, rounded three- (or more-) dimensional truth is what we need. The answer to the original question is a complex one: 'in such and such circumstances, yes, in other circumstances, no'.

A recent example, taken from the Rothamsted Ley-Arable experiment, may be useful. The simple question 'does wheat after three years of *lucerne* yield more than wheat after three years of arable cropping?' has no simple answer. If no N is applied to the wheat the plots that have been in lucerne give more wheat than those that have been in arable cultivation; if plenty of N (say 0.9 cwt N per acre) is applied 'arable' gives more than 'lucerne'. The maximum yield obtainable by varying the amount of N appropriately for each crop-sequence is greater

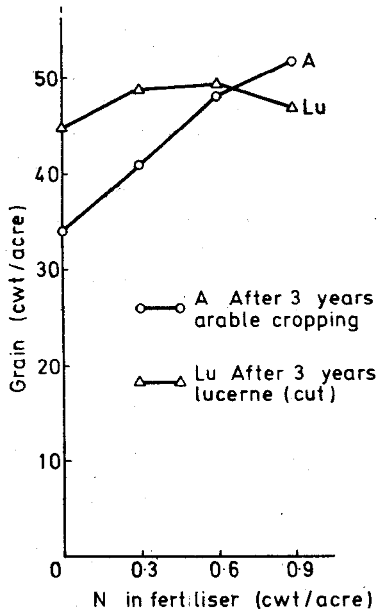

Figure 1.1 Rothamsted Ley-Arable experiment, Highfield. Mean yields of wheat 1961–63

after 'arable' than after 'lucerne'. Similarly (assuming prices of fertiliser and grain as in 1970) if N is adjusted for maximum profit 'arable' gives more wheat than 'lucerne'. All these statements are needed (and perhaps others I haven't thought of) to answer fully the apparently simple question originally put. The relevant *response-curves* (see *Figure 1.1*) are a convenient means of presenting the facts of the case.

1.5 SIMPLE OR COMPLEX EXPERIMENTS?

In general there are two main ways of increasing the usefulness of the answer we get to our question, of increasing the range of circumstances in which the answer will be applicable. The first is to repeat a simple experiment at many sites in several seasons, so getting a more or less fair sample of the variable conditions for which we want the answer (or answers) to the question. This may be appropriate to the comparison of varieties susceptible and resistant to mildew mentioned above. The second is to add more treatments to those originally proposed for the experiment. The change will often (but not always) involve introducing new factors (in the technical sense), that is, including all combinations of the original treatments and some other set of treatments. For example in testing N at different rates of application it may be appropriate to use all combinations of these rates and several rates of K. Or, in the Ley–Arable example, in testing the effects of different preceding crop-sequences we may need all combinations of sequences and several rates of N-fertiliser. For examples where the strictly factorial scheme is not appropriate, see Section 7.5.

This may sound rather daunting and it must be admitted complex experiments are harder to do, to interpret and to report, than simple ones. But if real life is complex so also will be much of the work of investigating and usefully describing it. Some consolation can be found in the fact that an experiment can often be made factorial (or more factorial than it is already) without increasing the number or size of plots.

If you are thinking of using 32 plots to compare four varieties with eight replications each arranged as a *randomised block* of four plots, you can easily put in a second factor, e.g. N-fertiliser at two rates ('*levels*' in the accepted jargon of statistics). You now have four replicates which can be laid down as four blocks of eight (this is not the only possibility – see '*confounding*'). The increase in the size of each block may cause an increase in *error variance* (because this now has to take into account differences in fertility between plots that are further apart) but in most cases that I know this is not serious. But you can go further ('deeper' might be a better word): why not include a third factor (K perhaps, or a systemic fungicide to control mildew) also at two levels? Now you can have two blocks of 16 (each block a replicate) or by using confounding you can still use blocks of eight plots. There are several possible designs available but whichever you use you lose some of the *degrees of freedom* available for estimating the error variance, so lessening the accuracy of the estimate and (with it) the likelihood of detecting as 'significant' an effect of any given magnitude. Put another way, the least significant difference (at a given level of probability) will be larger than in the simpler design, but the increase is relatively small. If you use a probability of 0.05 (1 chance in 20) to define a significant difference the increase is about 3 per cent; if you use 0.01 (1 chance in 100) it is about 5 per cent. The mesh of the net through which you strain your results is slightly coarsened, but the slightly increased risk of missing a 'real' effect, and the slightly decreased precision of the effects estimated from the experiment, are a small price to pay for the widening of the basis of the conclusions due to including the extra factor.

This process can be carried further, to single replicate and fractional-replicate designs. I say no more of these here but end by offering you a rough summary.

By adding in extra factors up to the limit you may lose some accuracy in your estimate of error. But, if there are *interactions* with the added factors you gain in knowledge of a complex mechanism which cannot be simply described; if there are no such interactions, you have most of the value of the degree of replication you would have had with the simpler design.

Chapter Two

PLANNING A FIELD EXPERIMENT

2.1 CHOICE OF DESIGN

If a statistician is available, consult him about the design of the experiment. He may also suggest extra treatments or modifications of the treatments proposed by the experimenter. In this, and in suggesting the number of replicates needed, he uses his experience of other experiments on similar subjects, and of the methods of interpreting their results.

If no statistician is available, pretend for the time being that *you* are one. Try to get detached temporarily from the muddy practicalities of your programme and consider the logical progression from design to interpretation.

A few words on the functions of a statistician (the word 'biometrician' is used by some people for the chap I am thinking of, leaving 'statistician' for the man who calculates standardised death rates and indices of prices and such like 'observational' statistics). Your statistician's job in considering your proposed experiment is to make your field work as effective as possible for the ends you have in mind. (He may, incidentally, help you to define these ends more clearly.) He will try to give you the design that will give you the maximum amount of relevant information from a given number of plots and he will later help you to extract all the useful information from the results of your experiment. He may, from time to time, throw cold water on your hot head – generally by telling you that so many plots have little or no chance of estimating with useful precision an effect of the magnitude you think likely to occur. This is *not* destructive criticism but is meant to save you wasting your time and effort.