IICA
A50
269

Biblioteca

PROCEEDINGS OF A WORKSHOP ON RESEARCH METHODOLOGY

IICA
BIBLIOTECA VENEZUELA
31 NOV. 2002
RECIBIDO

Union Agricultural Station
St Lucia
June 1989

PROCEEDINGS OF A WORKSHOP ON RESEARCH METHODOLOGY

Held at Union Agricultural Station,
St. Lucia,

on

Thursday and Friday, 15th and 16th June, 1989.

Sponsored by

INTER-AMERICAN INSTITUTE
FOR CO-OPERATON IN AGRICULTURE
(IICA),

CARIBBEAN AGRICULTURAL RESEARCH
AND DEVELOPMENT INSTITUTE
(CARDI),

and the

MINISTRY OF AGRICULTURE, LANDS, FISHERIES
AND CO-OPERATIVES.

00002463

BIBL
3 ... 2

CONTENTS

FORWARD

by

Dr. Antonio M. Pinchinat, IICA

The Workshop was organised as a co-operative activity between the Ministry of Agriculture of St. Lucia on one hand and CARDI and IICA jointly on the other. It aimed especially at upgrading the technical capability of the Ministry's research and extension personnel in research methodology, for the technological development of agriculture. The organisers however are pleased to record that from the eleven participants, there were two from WINBAN and one from CARDI, along with the remaining eight from the Ministry. Among these, were five of the six leaders of the Core Crop and Livestock Development Projects established then at the Ministry. The participants held at least a diploma up to a post-graduate degree and most had less than five years experience in conducting agricultural research.

The results of the evaluation carried out at the conclusion of the Workshop tended to suggest that the topics covered, and the technical level of their presentation, successfully addressed the practical research methodology needs of the participants and aroused their interest in pursuing this area of professional skill. Consequently as a follow-up they should be provided with hands-on advisory support and technical assistance, to fully achieve the final objective of the workshop. CARDI and IICA commit their cooperation with the national agricultural research system for that undertaking.

CHAPTER 1

OBJECTIVES OF RESEARCH

John L. Hammerton

Understanding and clarifying the reasons for doing research affects our choice of location, our choice of treatments, and our choice of control treatments.

Why do agricultural research?

There are several possible reasons that might motivate us.

* To expand the frontiers of knowledge? *A noble objective. But how realistic or indeed worthwhile is it if we have limited resources?*

* To do some in-depth research and basic research on something or other? *For what purpose--to know more and more about less and less? It's easy to get carried away.*

* To write a paper for a scientific journal and establish a scientific reputation? *A justifiable objective, and somthing we all want--but what sort of priority?*

* To increase national food production? *A worthy and proper objective--but a little vague. How is it to achieved, and what is our role?*

* To develop recommendations for farmers? *Eminently practical, and the recommendations may increase food production. But for what group of farmers, and exactly what crops or livestock?*

* To resolve one or more problems? Problem solving is a good reason for doing research, but we must determine the problem(s). Are they real problems, or those that we would like to research?

Let's assume our objectives are to develop recommendations or to resolve one or more specific problems.

* We must be clear on the target group or groups of farmers. *This means knowing where the farmers are located, what their resources are, what their production systems are, and so on.*

* We must be clear on the precise objectives of our research. *To increase yields and production? That may increase marketing problems, and lead to over-production and depressed prices. So we may need to look for additional markets. To increase output per man? To reduce production costs--by improving technology? To improve quality? In any case we must understand the problems, and the constraints as perceived by the farmers.*

* We must have a framework--a plan, a programme, and a time-table. *We must have an order of priority, and a sequence of technologies to be developed, and of experiments to be done, that matches the resources available. We must think through the objectives, the implications of success, the conditions precedent and the conditions necessary for success. We must also think through the assumptions implicit and explicit in the programme.*

Where to do the research?

* On a field station, or on farmers holdings? The researcher has total control of a field station experiment. How much control does he or she need? At what stage is the technology being developed? Is it ready for farmer participation? Or are there still some problems to be ironed out?

What treatments?

The nature of the most appropriate treatments depends on the objectives: basic research or technology development for farmer adoption? Let's take an example:

> *Diamond-back moth is a major problem in cole crops throughout the Caribbean. What might a research activity on diamond-back moth control attempt to do?*

* Screen all possible insecticides for their efficacy in killing diamond-back moth larvae. *How? Single applications of each insecticide and counts of larvae or assessments of damage at intervals after treatment?*

* Screen those insecticides currently in use for diamond-back moth control in the Region. How? *Single applications of each insecticide and counts of larvae or assessments of damage at intervals?*

* Compare several frequencies of insecticide application for diamond-back moth control. *How? Application of a limited number of insecticides at different intervals over a period of time and counts of larvae or assessments of damage at intervals?*

* Compare several insecticide rotations for diamond-back moth control. *How? Application of a limited number of insecticides in different rotations and counts of larvae or assessments of damage at intervals?*

* Improve on present methods and levels of diamond-back moth control. *How? Application of a limited number of insecticides at intervals based on scouting, use of netting to discourage insects, and counts of larvae or assessments of damage at intervals?*

All these are possible and legitimate sets of treatments. Which is the most likely set to yield meaningful results? Choosing the best set of treatments is not a simple matter.

What control (check) treatment(s) to use?

This depends on the level of technology development, on farmer practice, on the objectives of the experiment, and on the nature of the treatments.

* *Nil control: no fertilizer, no weed control, no pest control, no nothing? Is this realistic?* Do we need to measure what happens with nothing? What do farmers do? Will a nil treatment bias other treatments, causing us to under-estimate the effects of these others?

* *Farmer practice*: his or her pest control practice, fertilizing practice, cultivar, spacing, etc. What if this is variable between farmers? Use an "average" practice?

* *More than one control treatment?* Why not? Why not an "average" farmer practice and the individual farmer's practice as two controls for on-farm experiments? Then compare the "improved treatments" against both the average farmer practice and the individual farmer's practice.

CHAPTER 2

EXPERIMENTAL DESIGN

Bruce Lauckner

There are several assumptions in designing experiments and in blocking that must be understood in order to design an effective experiment.

Suppose an experiment is to be done to compare four corn cultivars (1,2,3 and 4). A novice might suggest a design such as in Fig. 2.1: four plots only, one for each cultivar. There are two types of error variation in experiments: systematic error which can be controlled, and random error which cannot be controlled but is minimised in a good experiment. The use of just four plots cannot control systematic error, and cannot estimate random error as this is done by comparing the response of two or more plots to the same treatment. A second option is to divide the field into 16 plots as shown in Figure 2.2. This design still makes no attempt to control systematic error, which is due in part to environmental variation over the area. In fact the variation between plots treated alike will be very small.

Figure 2.1

1	2
3	4

Figure 2.2

1	1	2	2
1	1	2	2
3	3	4	4
3	3	4	4

Replication

The novice's next suggestion might be that in Figure 2.3. Here the four cultivars are allocated at *random* to the 16 plots so that each cultivar is *replicated* four times. This is the simplest form of experimental design, and is known as a **completely randomized design.** The random error can be validly estimated, but no attempt has been made to control systematic variation. Hence it is not a good design if there is likely to be environmental differences in different parts of the field. The completely randomized design is however a good design for greenhouse and laboratory experiments where there is little or no environmental variation present.

Figure 2.3

4	1	3	1
2	3	2	4
3	4	1	2
4	1	3	2

Figure 2.4

Block 1 Block 2

4	1	3	1
2	3	2	4
2	3	4	3
1	4	1	2

Block 3 Block 4

Blocking

In a flat, apparently uniform field, plots of land close to each other are more likely to have a similar environment than plots distant from each other. Therefore in Figure 2.4 the field is divided into four *blocks*, and each cultivar is allocated at random to one plot within each block. This is the **randomized block design** or the **randomized complete block design.** Note that we can also have incomplete block designs. The differences in the responses between the different blocks is the systematic error which has now been controlled, and differences within the blocks are the random error that should be fairly small. In a design of this type, the treatments are said to be *orthogonal* to the blocks.

If the field is not flat and uniform, the blocking scheme in Figure 2.4 may not be appropriate, and other designs should be used.

It should be noted that the completely randomized design of Figure 2.3 need not have equal replication for each treatment. One of the advantages of the completely randomized design is its simplicity which allows great flexibility in design without introducing a difficult analysis. The randomized block design is not so flexible. Treatments can have different numbers of replicates, but each block must contain the same number of replicates for each treatment as each other block. Hence if there are four blocks the total replicates for each treatment must be a multiple of four. In more complex block designs, the number of replicates for treatments can vary from block to block.

The Concept of Blocking

Blocking and replication are not the same thing: replication can be achieved with blocks, but there can be more replications than blocks where treatments are repeated within blocks. Furthermore, blocks are not necessarily contiguous, nor need they be uniform in shape. An unrestricted allocation of treatments cannot make allowances for known local variation in some environmental factor. If a trend (in some environmental factor) is known to exist, the allocation of treatments must be restricted to take cognisance of this trend. Units or plots known to be as similar as possible are grouped, and each group is a block. The similar units or plots are not necessarily contiguous, though they may often be so. A block is therefore a set of identical, or very similar, units. Examples are: all treatments are compared on a rather stony part of the field, and on a rather wet part of the field; animals of the same litter are allocated to different treatments.

This grouping controls the local variation because it controls the allocation of treatments to units or plots so that no treatment appears only on the stony area, or only on the wet area, or only on the smallest animals. It also reduces the amount of random variation by removing the variation between blocks from the variation involved in comparing treatments. Careful blocking can considerably improve the precision of an experiment. The analysis of variance cannot remove variability due to variation within blocks, which is why it is so important to make sure that plots within blocks are as uniform as possible.

Over-blocking

Blocking is a safety measure. It is like an insurance policy to protect against the effects of variability. So there is a penalty (premium) to pay. By introducing blocks into the design, degrees of freedom are removed from the error term in the analysis of variance. If the blocking has been effective, this loss in degrees of freedom will be more than offset by the variability which the blocks have removed, so the precision will have been increased by blocking. If there is no variability in the field the blocks will not reduce the variability present in the experiment, so the (unnecessary) use of blocks will give lower precision than if the blocks had not been used. A complex design, compared to the simple completely randomized design, should therefore be used only if it is required.

Variability removed by different designs

* Designs that account for variability in one direction or from one source only include the following.

  Completely randomized design: this has no block structure, and should be used only where the environment is homogeneous.

  Randomized block designs: these incude the randomized complete block design, the factorial, and the split-block design.

* Designs that account for variability in two directions or from two sources include the following.

  Latin square: the number of rows and columns is equal to the number of treatments, and the number of plots is therefore the square of the number of treatments.

  Youden square: the number of rows is equal to the number of treatments, but the number of columns is less than the number of treatments.

If the blocking structure is complex, but variability is low, then a blocked structure will give less precision than a less blocked, or unblocked design. Consider an experiment with six treatments and 36 plots available. The analysis of variance for three differently blocked designs would be:

Completely randomized design

Source	df	SS	MSS
Treatments	5		
Residual	30	ss	ss/30
Total	35		

Randomized complete block design, six blocks

Source	df	SS	MSS
Blocks	5		
Treatments	5		
Residual	25	ss	ss/25
Total	35		

Latin square

Source	df	SS	MSS
Rows	5		
Columns	5		
Treatments	5		
Residual	20	ss	ss/20
Total	35		

Clearly, if the residual sum of squares (ss) is the same in the Latin square (LS) as in the completely randomized (CR) design, then the residual mean square for the CR will be two-thirds that of the LS, so giving a more precise analysis. If the rows and columns remove a lot of the total variation, the residual sum of squares will be less than for the CR design and the LS will give a more precise analysis.

Another purpose of blocking

Blocking can also be used to help the practical execution of an experiment, even when there are no gradients in the field. This enables agronomic and crop care operations, and data collection, to be done block by block where it is impossible to do any of these in a single day. So in a large experiment, operations can be spread over two or more days, provided a day's work coincides with the block structure. Any differences due to spraying two days later, for example, will be confounded with blocks. In a completely randomized design all operations must be done the same day, otherwise sources of uncontrolled variation will be introduced and the residual error will increase.

CHAPTER 3

THE SIZE AND SHAPE OF PLOTS

John L. Hammerton

There is no such thing as a "best plot size" for a particular crop or livestock species under any and all conditions. Optimal plot size is a function of the inherent variability of the planting material or of the animal species, of errors of observation and measurement, and of the environmental variation.

Some guidelines on plot size:

* Keep plots as small as possible, and uniform throughout the experiment. *It is better to have more replications of smaller plots than few replications of larger plots on a given area of land.*

* Keep plots big enough to be representative of the population of plants or animals. *This will ensure that the plots give representative values of yield, disease incidence, etc., and ensure that there is a reasonable ratio of border or guard area to "net" plot area.*

Practical factors determining plot size include:

* Ease of treatment application to individual plots: the nature of the treatments and the method of treatment application therefore influence optimal plot size.

* Ease of overall management of the experiment: the methods of applying common or basal levels of fertilizer, of pesticides etc. will also influence optimal plot size.

* Cost of the experiment: formulae have been used to calculate optimal plot size given estimates of soil heterogeneity, costs associated with the number of plots only, costs

per unit area, and costs associated with the borders, but these are difficult to use because of the difficulty of partitioning costs and estimating soil heterogeneity.

Plot Shape

In block designs:

* Where there is *distinct directional variation*, use long narrow plots with the length of the plots in the direction of the variation.

* Where there is *no directional variation*, use long narrow plots in any direction to give compact blocks.

* Where *any directional variation is unknown*, use plots as square as possible to give compact blocks.

* Where *border effects are likely to be large*, use plots as square as possible.

Practical factors determining plot shape include:

* Method of treatment application, especially the width of equipment.

* Method of application of common or basal levels of land preparation, fertilizer, pesticide etc., especially where mechanical equipment is used.

* Need for guard or border rows or areas: rectangular plots have a higher ratio of guard plants or area than square plots

Border rows

Border or guard rows or areas are necessary:

* Where one or more neighbouring plots, or unplanted alleyways, are likely to affect plant growth and yield because of competition. Cultivars, for example, may differ in height and amount of shading; edge rows or plants may poach fertilizer from a neighbouring plot with a higher rate of application, etc.

* Where one or more treatments may increase the spore inoculum level, or the population level of a mobile pest, above the average level because of higher susceptibility, or because of the nature of the treatment (a zero control for example). This may mean very large plots (mostly guard areas), special spraying of guard rows, or special designs (wide spacing of plots in a matrix of a "commercial" planting, for instance). In fact, evaluation of fungicides and insecticides on small plots in a conventional design, may seriously underestimate efficacy by increasing abnormally the inoculum or population level.

CHAPTER 4

EXPERIMENTS WITH LIVESTOCK

Bruce Lauckner

Many of the principles already discussed also apply to livestock experiments, although the latter have special constraints and problems. Some of these are discussed below, under a number of headings.

Pen trials

A randomised block design can be used with many pen trials, with the blocks being chosen according to the geographic location of the pens. The individual pens are the plots. Pens of different sizes can be included in the same trial, but pens of different size should not be included in the same block unless pen size is an experimental treatment. Pens of different sizes must otherwise be in different blocks, so that any effect of pen size is confounded with blocks (i.e. included with the variation due to blocks). Similarly, the number of animals per pen can vary, but, unless this is a treatment factor, animal numbers per pen should be constant within blocks. Uselessly there will be no variation in either pen size or number of animals per pen in an experiment.

Blocks should still be chosen carefully. Some pens may receive more sunlight, or be more exposed to the prevailing wind or the rain. Or there may be differences in the type of watering device, or type of feeder. Blocks should be chosen so that all the pens in a block are as uniform as possible in environmental characteristics.

Blocking may also be necessary according to some physical characteristic of the animals, such as sex, size or breed. However, such physical characteristics may be an experimental treatment (a comparison of sex or breed, for example), in which case it should be balanced orthogonally with blocks if the design is a randomised block.

Thus in a feeding trial in pens, it may be necessary to try to block for animal size, environmental conditions, and other factors, such as type of feeder. If it is not possible to block for all these factors, given the resources of the experiment, than one or more of thes factors should be used as a covariate, and analysis of covariance performed. Initial live weight is one of the most obvious covariates to use, and one of the easiest to handle in the analysis.

Loss of animals

If there is more than one animal per pen, and one dies, the other animals in the pen are likely to begin to perform better than those animals in pens which still have the original number of animals. These problems should be overcome by covariance analysis with number of animals per pen as the covariate. If all the animals in a pen die, or deaths occur in experiments with only one animal per pen, then missing value techniques should be used. The use of reserve animals is not recommended, as such animals have not been subject to the experimental conditions from the outset. If they have been subject to these conditions, then the experiment should probably have been larger from the beginning.

Pen trials are usually subject to severe constraints because of the number of pens or animals available. Furthermore the environment in the animal house may make blocking into blocks of equal size difficult. Other designs may be more appropriate, with additional replication of the control treatment. These designs ensure good comparisons between the control treatment and the others, though the precision of the comparisons between the non-control treatments may be rather poor. Mistakes in the application of treatments to pens, or in the recording of data, can be avoided by clear labelling of the pens.

Grazing trials

In grazing trials single animal plots should be used where possible, to help provide a sufficient number of error degrees of freedom in the analysis of variance. Thus if 40 sheep are available, and there are four grazing treatments, it is better to have 10 blocks with single sheep plots, than five blocks with two-sheep plots. In the former instance there would be 27 degrees of freedom for error compared with 12 with the two-sheep plots. However, some types of trial may not be appropriate for single animal plots.

Where one of the experimental treatments is stocking rate, large areas of land may be involved, since one or two animals grazing small plots cannot simulate the effect of many more animals on much larger areas. Replication of stocking rates may be impossible therefore. However if there is another treatment factor under investigation (which may simply be a time factor), then if each stocking rate is replicated once only for each of the other treatments, than an analysis for the effect of stocking rate is possible.

Replication over time may be possible, either between different seasons or years. Differences between seasons or years, and any interaction with stocking rates, can be investigated by covariance analysis. Such an analysis can be combined with a regression analysis to estimate optimum stocking rates for the different years or seasons. The theory and computations are rather complex, and are not presented here.

Increasing the degrees of freedom

In livestock trials row and column designs are often used to increase the error degrees of freedom in the analysis of variance. The rows are the blocks, chosen as above for pen trials or a group of animals in a grazing trial, and the columns are time periods.

Animals	Jan	Feb	Mar	Apr
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

Figure 4.1
Latin square cross-over design

Consider a sheep trial in which four different feeding regimes (A,B,C and D) are to be compared for their effect on weight gain. Such an experiment can be done with just four animals using the Latin square design shown in Figure 4.1. For the first month (January) animal 1 receives treatment A, for the second month treatment B, for the third month treatment C, and for the fourth month treatment D. The other animals also rotate among the treatments as shown. Of course the rows and columns should be randomised.

During each month the performance (weight gain, feed intake, feed conversion etc.) are measured for each animal. The analysis of variance is as follows:

Source of variation	d.f.
Between animals	3
Between time periods	3
Between feeding regimes	3
Error	6
Total	15

Since the sheep will take about a week to adjust to the new feeding regime (rations, frequency or method of feeding), no observations should be made during this period. If the "settling in" is likely to be longer than a week, then it may be necessary to use experimental time periods longer than a month for the feeding regime to have a measurable effect. This is known as a **cross-over** or **change-over** design. A complete latin square is not necessary (as in Figure 4.1.) if this is difficult to achieve. Other forms of row-and column designs are available and can be used for cross-over designs. One cross-over -- with just two time periods -- can be used if this gives adequate degrees of freedom for error.

An advantage of cross-over designs is that covariance adjustments are not necessary: the row-and-column analysis will include (in the "between animals" term) any differences in animal size, pen size, etc. If the cross-over design is used for an experiment in pens, then animals should in the same pen throughout the trial. In this way pen effects are confounded (i.e. included with) animal effects and not with the treatment effect.

In looking at the data from an experiment in which liveweights have been recorded at intervals over a period of time, a regression approach is to be preferred to analysis of variance. The regression lines for the different treatments are compared. A brief outline of this approach is given in Chapter 9.

Analysis of livestock experiments are often complex because of the need to include one or more covariates, and missing values can occur because of animal deaths. A livestock experiment can generate much more data than an experiment with a short-term crop. In planning the experiment thought should be given to what is to be recorded and how often. A simple pro-forma should be designed to keep track of records, and to ensure that records are taken at the planned time. A computer can also help to keep these records: see Chapter 5).

CHAPTER 5

DATA COLLECTION AND HANDLING

John L. Hammerton

Data collection, handling (and analysis) take time and cost money. So it is important to decide what data are to be measured or recorded, and how these are to be measured, recorded, and analysed.

Types of Data

Data are of three main types:

Counts. These can vary from zero to small numbers to very large numbers. Examples are number of plants per 1 m of row, number of tillers per 1 m of row, number of aphids per plant, number of anthracnose lesions per leaf or per 10 leaves, number of nutgrass shoots per square foot, and so on. These data are whole numbers. It is important to decide exactly what is to be counted. How big must a lesion be to be counted? How much of the seedling must be through the soil to be counted as a seedling? Such data may need transforming before analysis of variance can be done.

Measurement Data or Continuous Data. Such data are specified along a continuous numerical scale including decimal places. They are measured using rulers, tape measures, balances, weigh-bridges, leaf area meters, or by chemical analysis, for example. Weights of grain, live-weights of pigs, heights of plants, height to the shoulder, shoulder back fat thickness, leaf areas, and percentage nitrogen are examples of such data. It is important not to attempt to record to a greater degree of accuracy than the equipment allows. Thus if the balance is marked in tenths of a kilogram, do not record weights as 5.57, 6.15, 4.75 etc. If all the measurements are expected to be between, say, 60 and 100 kg, it may not be worth measuring to a fraction of a kg. On the other hand, if the measurements are likely to range from 4 to 6 kg, then it is important to measure to a tenth of a kg.

Attribute Data. These have finite numbers of discrete classes. They are used when some variable cannot be measured, is difficult to measure, or when a "quick and dirty" means of recording is necessary. Example are anthracnose scores (1 to 5), weediness scores (1 to 10), degree of redness, degree of phytotoxicity, etc. The problem is to maintain consistency, particularly when the recording is to be done at intervals (to study the development of anthracnose over time, for instance). Persons recording should either work in pairs, or score different blocks, but should attempt to agree on the basis of their scoring. It is important to avoid bias by not reading the plot label, if these show treatments, and by not using a plan that shows the treatments! In any case, some notes must be made to describe the top and bottom points on the scale. (Otherwise, 2 months later you may not be able to remember whether 0 was no weeds or no control!). Short scoring scales--with no more than 10 classes--should be used. Attribute data cannot usually be analysed by analysis of variance. An empirical examination of the means--and of the variability may be enough, or a chi-square analysis may be appropriate.

What data to record?

Data are recorded for different purposes. It is important to list all the data (variables or observations) that you hope to collect, so as to prioritise them, and to think about how they are to be collected, and what equipment may be needed. The purposes for which data are collected are as follows:

Primary data. These are ***essential*** data. Crop yield and yield of milk are examples of this type of data.

Substitute primary data. These are used where it is to difficult or too expensive to obtain the primary data. Examples are counts of the number of fruits (instead of the weights in successive harvests), number of bundles of forage (instead of weights), number of weaned lambs (instead of weaning weights), etc.

Explanatory data. These are recorded to attempt to explain treatment effects. Examples are disease and pest incidence, plant nutrient status, internal parasite levels, etc.

Concomitant data (or observations). These are measured to attempt to explain some of the random variability in the primary data. These observations should not be affected by the treatments. Examples are differences in weediness, differences in initial live weight, differences in plant stand (where it is not due to treatments), etc. The precision of an experiment can be increased by using a concomitant observation in an *Analysis of Covariance*. Concomitant observations such as pre-treatment live weights, or the pre-treatment yield of a perennial crop, are useful ways of increasing precision. Other observations may not be planned in advance, but may emerge as ways of explaining some observed variation.

Prior measurement of primary data. If praedial larceny is expected, it may be a useful insurance to make some estimate of yield before the crop matures. This can be used as substitute primary data, or used to estimate yield for those plots damaged.

External records. These include data on rainfall, the dates of weeding, dates of spraying for pests and disease control, and so on.

How to record and store data

* Do not record anything and everything unless it may be of some use in explaining the results.

* Record by blocks in a blocked design, so that any differences in method, date of recording, etc., are confounded with blocks.

* Train technicians in data collection, if necessary, and supervise as far as possible. If necessary, assign different technicians to different blocks. Emphasise the need for care and accuracy in measuring and recording.

* Use a proforma designed for the experiment and for the type of data. These help to avoid errors by providing boxes, columns, or spaces for each bit of data. Thus if plant numbers in five sample lengths of row are to be recorded from each plot, having five boxes or columns will avoid counting six in some and four in other plots. Do not record in diaries, on envelopes, on the back of your hand etc. You may wash your hand before you have transcribed the data!

* Avoid transcription of data: use carbon paper or a photocopier. Guard your data with your life!

* Use a computer to store your data, and to at least do some simple processing. A spreadsheet programme can be used to convert lbs per plot to tonnes per ha, to adjust for moisture content, and so on. The example on page 23 shows how this can be done very simply.

Corn Variety Experiment: Union 1989.

Data: Yield of grain in lb per pl (Plot size: 99 square feet)

	Varieties							
	Bruce	John	Antonio	Dave	Kenny	Henry	Lucius	Earnest
Block 1	7.1	11.5	8	5	6.7	3.6	8.4	4.1
Block 2	5.25	12	7.3	4.5	3.5	5.6	7.1	5.3
Block 3	6.5	13.6	6.9	4	6.2	7.4	5.8	5
Block 4	6.2	11.8	8.3	4.25	5.5	3.5	5.7	4.2
	6.2625	12.225	7.625	4.4375	5.475	5.025	6.75	4.65

Original data from the field.

Formula for average is:-
Average(B8:B11) etc.

Data: Percentage dry matter in grain.

	Varieties							
	Bruce	John	Antonio	Dave	Kenny	Henry	Lucius	Earnest
Block 1	79	86	85	78	78	79	83	82
Block 2	81	85	83	79	83	84	79	82
Block 3	80	84	82	80	80	81	82	79
Block 4	81	85	83	81	81	82	86	83

Grain samples were taken and oven dried to get the dry matter contents. (The data on sample "wet" and dry weights could have been put into the computer and the percentages calculated by the programme, but this has not been shown.

Data: Yield of grain adjusted to 88% dry matter (lb per plot).

	Varieties							
	Bruce	John	Antonio	Dave	Kenny	Henry	Lucius	Earnest
Block 1	6.37	11.24	7.73	4.43	5.94	3.23	7.92	3.82
Block 2	4.83	11.59	6.89	4.04	3.30	5.35	6.37	4.94
Block 3	5.91	12.98	6.43	3.64	5.64	6.81	5.40	4.49
Block 4	5.71	11.40	7.83	3.91	5.06	3.26	5.57	3.96
	5.71	11.80	7.22	4.00	4.98	4.66	6.32	4.30

Grain yields are adjusted to a standard dry matter percentage--usually 88%.

Formula to convert to the standard dry matter percentage is:-
(B8*B19)/88 etc.

Average(B30:B33) etc.

Data: Yield of grain (at 88% dry matter) in metric tonnes per ha.

	Varieties							
	Bruce	John	Antonio	Dave	Kenny	Henry	Lucius	Earnest
Block 1	3.14	5.54	3.81	2.18	2.93	1.59	3.90	1.88
Block 2	2.38	5.71	3.39	1.99	1.63	2.63	3.14	2.43
Block 3	2.91	6.40	3.17	1.79	2.78	3.36	2.66	2.21
Block 4	2.81	5.62	3.86	1.93	2.50	1.61	2.75	1.95
	2.81	5.82	3.56	1.97	2.46	2.30	3.11	2.12

Converted data from lb per plot to metric tonnes per hectare. Could do this in several "bites"--to kg per plot, to kg per acre, and then to tonnes per ha.--or in one "bite" as here.

Formula to convert lb per plot to tonnes per ha. is:-
((B30/2.205)*(43560/99))*(2.47/1000) etc.
Other formulae are possible.

Data:

	Varieties							
	Bruce	John	Antonio	Dave	Kenny	Henry	Lucius	Earnest
Block 1								
Block 2								
Block 3								
Block 4								

An empty "template" that can be copied as many times as needed.

Some of these tables are formatted to show only two places of decimals ("$ format").

CHAPTER 6

SAMPLING

Bruce Lauckner

Many sampling methods have been developed for many different situations. Sampling is not done to make life easier and should be avoided unless it is really the only way to collect data.

Why sample?

We must sample if the method of measurement is destructive, or if the entire experimental unit would be destroyed, as by measuring fresh and dry weights, for example. In such cases sampling is essential. Not every bomb can be tested before use! In other cases measuring the entire experimental unit may be too expensive or take too much time. To measure the entire population, where the population is dynamic, may take so long that by the time the information is tabulated it is out of date. If the measurement is difficult to make, trying to measure all the units may give higher errors than if sampling is done. It is better to collect high quality information from a sample than low grade and dubious information from the entire population.

The purpose of sampling is to estimate some character of an experimental unit from a part of that unit.

How many samples

Generally a large number of small samples is better than a small number of large samples. If the character is fairly uniform a second sample will not add much to what was learnt from the first. If the character is variable, then many samples are needed as there will be differences between them, and no one sample will characterise the experimental unit as a whole. The greater the variability, the larger the number of samples needed.

The question "How many samples do I need?" is difficult to answer without a reliable estimate of variability. The formula below is not helpful unless the level of variability is known or can be assumed. Similar studies in the literature may provide information on variability.

Example: How many samples should be taken to have a 5% (1 in 20) chance of being within 5% of the proportion of the population with that character?. Let p be the proportion with the desired character and q the proportion without it. So 1-p = q. Then n, the number of samples needed, is given by :

$$n = \sqrt{\frac{4pq}{25}} \qquad (6.1)$$

But this equation requires an estimate of the values of p and q before it can be used! If no estimates are available, it is best to take as many samples as possible.

Bias

A problem in not measuring the whole experimental unit is that of selecting parts of the unit in such a way as to avoid bias. Each experimental unit must be fairly assessed: the results are distorted if only the best plots of one treatment are measured and only the worst of another. The samples must reflect the whole unit.

Human beings are prone to bias, and will favour, or otherwise, a unit for many reasons. They may have a preconceived notion of what treatment will give the best results. An objective method of sample selection is therefore needed. The three methods of sampling described below all include an element of randomisation. If the sample is random, then the chosen samples should faithfully represent the population being sampled.

Simple random sampling

This method is used where every item in the population has an equal chance of being observed. If cabbages are to be sampled, every cabbage must have the same chance of being selected. Each cabbage can be numbered and numbers selected from random number tables used to choose the sample.

If Y is the population total for the character being measured, then this can be estimated by $\hat{Y}$,

$$\hat{Y} = N\bar{y} \tag{6.2}$$

where $\bar{y}$ is the sample mean and N the total number of units in the population. The population mean, Y, can be estimated by $\hat{\bar{Y}}$,

$$\hat{\bar{Y}} = \bar{y} \tag{6.3}$$

The standard error of $\hat{Y}$ in (1) is estimated as s.e.($\hat{Y}$),

$$\text{s.e.}(\hat{Y}) = \frac{Ns}{\sqrt{n}} \sqrt{\frac{(N-n)}{N}} \tag{6.4}$$

where n is the number of sample units and s the estimated standard error of the population mean.

s is given by:

$$s = \sqrt{\frac{Sy^2}{n-1}} \tag{6.5}$$

where Sy^2 is the corrected sum of squares of the sample values. The standard error of the sample mean ($\bar{y}$) is estimated as s.e.($\bar{y}$),

$$\text{s.e.}(\bar{y}) = \frac{s}{\sqrt{n}} \sqrt{\frac{(N-n)}{N}} \tag{6.6}$$

As mentioned above, a knowledge of the likely value of s is needed to decide upon the sample size required to estimate the population total with acceptable precision. Equations (6.4) and (6.6) can written as:

$$\text{s.e.}(\hat{Y}) = \frac{Ns}{\sqrt{n}} \quad \sqrt{1-f} \tag{6.7}$$

and

$$\text{s.e.}(\bar{y}) = \frac{s}{\sqrt{n}} \quad \sqrt{1-f} \tag{6.8}$$

where f=n/N, the sampling faction. n/N must be large if N is small, but reduces as N increases. These formulae show some surprising results: for a fixed value of s, a sample of 1000 from a population of 20,000 gives much the same precision as a sample of 500; and a sample of 500 from a population of 20,000 gives much the same precision as a sample of 500 from a population of 10,000.

Stratified random sampling

The method above gives every item an equal chance of selection. The sample may not represent the population however if there is a lot of variation between items in the population. Suppose the population consists of large and small farms, and it is expected that the results from the two sizes of farm will be different. Simple random sampling could result in all the samples coming from one group, giving misleading information about the population. In this situation the sampling method must ensure that both small and large farms are represented. This is done by splitting the population into strata. In this example there would be two strata: large farms and small farms. Samples would be taken from both strata so that the selection of farms represented the whole population.

Generally the number of samples taken from each strata reflects the number of units in the strata. If there were 100 farms, 40 small and 60 large, and a sample of 20 was to be taken, then from the "small" strata (40x20/100) = 8 farms would be taken, and (60x20/100) = 12 farms from the "large" strata.

This method should be used whenever the character being sampled is not evenly distributed in the population. Weeds, insect pests and diseases often appear in clumps, and simple random sampling may give misleading results. While the researcher chooses the strata and assigns units to these strata, the selection of units to be included in the sample is done using random numbers.

The formulae for the estimates of the population mean and its standard error are as follows. The population mean, Y, is estimated by $\bar{y}_{st}$,

$$\bar{y}_{st} = \frac{\Sigma N_i \bar{y}_i}{N} \qquad (6.9)$$

where N_i = the number of units in the strata

$\bar{y}_i$ = the strata mean

N = the total number of units in the population

The standard error of $\bar{y}_{st}$ is calculated as s.e.$(\bar{y}_{st})$,

$$s.e.(y_{st}) = \sqrt{\frac{-W_i^2 s_i^2}{n_i} - \frac{W_i s_i^2}{N}} \tag{6.10}$$

where n_i = the number of unit samples in each strata

$W_i = \dfrac{n_i}{N}$ for each strata

s_i = he estimated standard error for each strata

$$= \frac{S_{yi}^2}{n_i - 1} \tag{6.11}$$

where s_{yi}^2 is the corrected sum of squares of the sample values in the strata.

The population total, $\hat{Y}_{st}$ is estimated by $N\bar{y}_{st}$ (6.12)
with a standard error of N s.e.$(\bar{y}_{st})$

Systematic sampling

A problem with stratified sampling is that sufficient background information is needed to enable the strata to be identified and the units alloted to strata. If there is insufficient background information, this method cannot be used. The time and cost of collecting this background information may be prohibitive, and systematic sampling may provide a practical compromise.

This method is very practical when sampling must ensure that the entire plot is represented by the samples. It will result in sample that shows considerable variation in the character being observed. First the number of units is counted and the number of samples to be taken decided. Suppose there are 49 cabages per plot, and it is decided to sample seven. To cover the entire plot, every seventh cabbage must be taken. The first cabbage to be taken is determined using random numbers, and thereafter every seventh cabbage is taken.

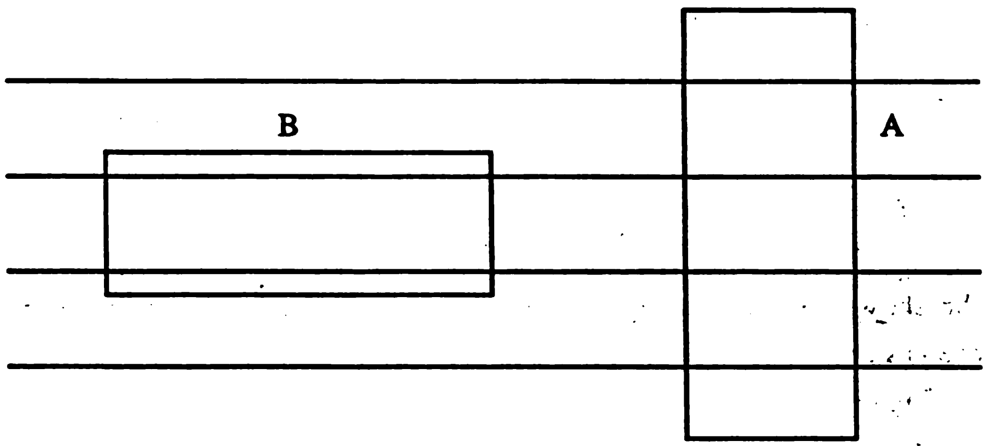
If there are N plants in the plot and n are to be selected, then a random number between 1 and N/n is selected. If this number is r then this gives the first plant to be selected and the second is r+N/n, the third r+2(N/n) and so on. If N/n is not a whole number, then the nearest whole number is used. This method is systematic only after the first sample has been chosen, and this is chosen at random.

The population total, Y, is estimated by $N\bar{y}_{st}$, and the population mean, $\bar{y}$, by $\bar{y}_{st}$, but there are no reliable methods for estimating the standard errors of these values.

Sampling orientation

As pointed out in the Chapter on Plot shape and size, the direction of the plots can influence the precision of the experiment. The same is true for the direction of sampling. If samples are taken in a field they can be across or along the rows, as in Figure 6.1.

Figure 6.1



Samples taken across the rows (A) will possibly result in a lower variability than those taken along the rows (B). This is because if one row has a poor stand the defect will be exaggerated in sample B. If there is a trend in the field, samples should be oriented at right angles to the trend, as in Figure 6.2. In a tall crop it may be necessary to restrict sampling to the edge, so that the rest of the plot is not damaged, as in Figure 6.3.

Figure 6.2

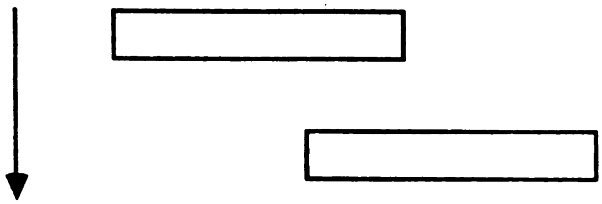
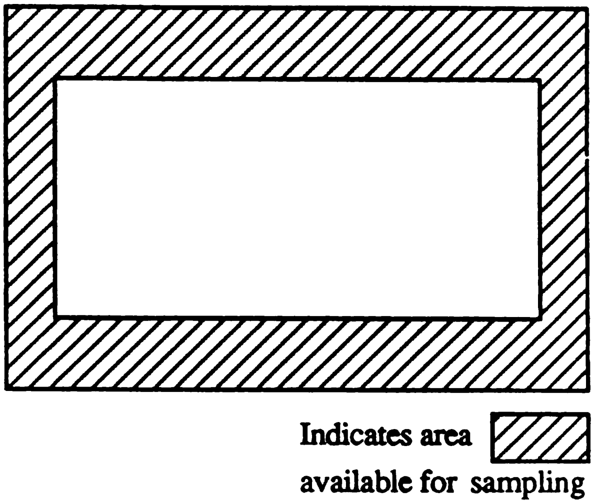


Figure 6.3



Sampling for the estimation of disease incidence

It can be shown mathematically that the "best" method to use to sample for disease incidence is stratified sampling. Often a field is sampled by taking one or two diagonals and sampling units along the line or lines. A variant is the "W" pattern of samples across the field. The advantage of stratified sampling is that it samples from a wider range than the diagonal or "W" method. Every unit has an equal chance of being selected with the stratified sampling method, whereas only units near the designated sampling path are sampled with the other methods. Only stratified sampling provides an unbiased estimate of disease incidence and allows analysis of in-field variance. The correct sampling intensity (the number of samples to be taken) depends on the biology of the disease. Disease estimation is more difficult in fields with low disease incidence or very aggregated distributions. Preliminary sampling may be necessary at a high intensity to establish an incidence threshold defined as the lowest level that must be accurately estimated.

Bias sampling

In some cases it may only be pssible to obtain a biased sample. If observations are to be made over time it may be the trend that is important, not the actual values. If so, than a biased sample may not be a disaster. If the sampling method is likely to be biased, it is important to recognise the fact. If travelling to all parts of the island is difficult because of the poor roads, it is clearly easier to take samples close to the good roads. Such samples could be biased in many ways. However if the same route is followed and the same units (farms, fields, trees, etc.) are sampled over time, the change in the observation should not be biased. This form of sampling is called "touring", and is useful when changes rather than actual values are important.

CHAPTER 7

RESEARCH ON PERENNIAL CROPS

Bruce Lauckner

The same attention to detail at the planning stage, and the same principles of management, as discussed for field crops, also apply to perennial crops. In addition there are certain considerations that apply particularly to perennial crops.

In general in the Caribbean, the researcher will have to use existing trees. It is only rarely that he or she will be able to plant the trees specifically for an experiment. This makes life easier in some ways, but results can be a long time in coming.

The area required is often large because of the wide spacing of the trees, but the number of trees available for experimentation may be much less than the total number of trees on the site. It is therefore essential to map the site and not only count the number of trees, but determine those that, from their spacing and size etc., can be used for experimentation. Trees may need to be discarded because of edge effects, missing neighbour trees, and difference is size and age.

The experiment cannot be designed until until the number of available trees is known. Some trees may have to be used as guard trees, but shared guards can be used. It may be difficult sometimes to fit plots into an odd-shaped area without "wasting" useable trees. Narrow plots must be avoided since a large number of trees will have to be assigned as guard trees. The effect of a missing tree, or of a non-uniform tree, can be very important. Compare the two arrangements in Figure 7.1. In the first arrangement there are no missing trees, and using outside guard trees but none between plots, a total of 16 trees (or four plots of four) can be obtained. Note that of 40 trees only 16 are used for experimentation. Note also that if plots of four trees are to be separated by guards, only three plots can be obtained.

Figure 7.1.
Two arrangements of trees to show the effect of a missing tree.

```
x x x x x x x x x x

x x x x x x x x x x

x x x x x x x x x x

x x x x x x x x x x


x x x x x x x x x x

x x x x x x x x x x

x x x   x x x x x x

x x x x x x x x x x
```

In the second arrangement, with only one missing tree, only three plots of four trees can be obtained, and it could be argued that, because of the lack of competition to the trees on either side of the missing tree, only 8 trees should be used. So of 39 trees, only 12 or fewer can be used for experimentation.

Control of local variation

In long-term experiments, control of local variation is far more important than in a field crop experiment. Some differences may become large when in effect for several years. So special care is needed not only in the layout of blocks and plots, but also in the design of the layout in the field. It is advisable to be more pessimistic about sources of variation in a long-term experiment then with a short-term experiment. A more complex design may be needed therefore: row and column designs or incomplete block designs. Use of guard rows becomes more important than with short-term crops, and selection of plants become very important: small early differences may become large after ten years. There is no alternative to really knowing the crop. Measurements at an early stage of growth may help predict its future performance. This can provide some control of variation that may not have been eliminated in the initial selection.

Use of covariance

This can be very effective with long-term experiments with tree crops. Records are taken prior to the application of the treatments, and used to adjust the data collected after treatment application, so as to reduce the variability due to differences between trees. Parameters such as soil depth, girth of trees, height, or yield in previous years, can be used as covariates. It may well be necessary to have data for more than one year because some tree crops show biennial or cyclic bearing. More than one covariate can be used in analysing data.

Some covariates used in specific crops are:

Crop	Covariate
Cloves	trunk circumference canopy
Cocoa	number of pods trunk circumference
Coffee	stem diameter
Pecan nuts	cross-sectional area of trunk at planting
Rubber	trunk circumference

For many other crops, covariance on previous cropping is recommended.

Statistical design

As pointed out above, extra care must be taken with the selection and use of blocks to control local variation. Small differences may accumulate over time. Blocking in more than one direction or the use of incomplete blocks may be necessary. Latin squares. or extended or incomplete Latin squares (Youden squares) can be useful. Figure 7.2 shows an extended Latin square which could be used for an experiment with six treatments (a to f).

Figure 7.2.
An extended Latin square (needs randomisation).
Letters represent treatments which are applied to individual trees or to plots.

a	b	c	d	e	f	a	b
b	c	d	e	f	a	b	c
c	d	e	f	a	b	c	d
d	e	f	a	b	c	d	e
e	f	a	b	c	d	e	f
f	a	b	c	d	e	f	a

Blocks should be formed in the same way as for short-term crop experiments (see Chapter 2). Forming blocks on the basis of past performance is generally not recommended, because such blocks allow only for past sources of variation. Covariance accounts for all variation in the independent variate, whereas only some variation can be associated with blocks, the rest being within the blocks.

Mishaps

Unforeseen events can occur during the life of a tree crop experiment. Some trees may die, possibly due to causes unrelated to the treatments and to the general crop care. Once a tree has died the site should be replanted, but the replacement tree can play no part in the experiment. Events less extreme than death can also occur. Pruning may be a part of general crop care, but pruning will modify the performance of the trees. All pruning, planned and accidental, must be recorded, as should any event likely to affect performance. A major change which is not uniform -- such as a hurricane or a flood -- poses major problems. Badly damaged trees may have to be dropped from the experiment, or the experiment modified to cope with the enforced changes in the number of trees.

Consider an experiment that is damaged by a hurricane. The tops of some of the trees are snapped off, so that the subsequent growth patterns will be seriously affected. Damage will not be uniform over the site however. The trees can be grouped by level of damage, but there may be only a few trees that will make relatively normal growth and give a reasonable yield. Data from the other trees will refer to the growth and bearing of damaged trees. These data may be of considerable interest to growers in a hurricane area, however.

Changes

During the course of a long-term experiment there may be new treatments that ought to be examined. Modifications to the original design must be very carefully considered, because the current experiment could be adversely affected. If it is considered likely that the treatments may need to be modified during the course of the experiment, it is best to plan for this at the design stage, so that the modifications can be easily made.

As an example consider the fertilizer experiment of Figure 7.3. Plots comprise three trees separated by guards and surrounded by guards, with the treatments t1, t2 and t3. Assume that it is suggested that a trace element deficiency may be limiting growth. How can the current experiment be modified to test this suggestion? One way is as shown: an end tree of each plot is sprayed with the trace element (while continuing to receive the original fertilizer treatment), the middle tree becomes a guard (to avoid spray drift), and the remaining tree receives the original fertilizer treatment. The symbol m indicates the trace element treatment.

Figure 7.3.
The original design of a fertilizer experiment,
and its modification to include an additional factor.

g	g	g	g	g	g	g	g	g	g	g	g	g	g	g
g	t_1	t_1	t_1	g	g	t_2	t_2	t_2	g	g	t_3	t_3	t_3	g
g	g	g	g	g	g	g	g	g	g	g	g	g	g	g

g	g	g	g	g	g	g	g	g	g	g	g	g	g	g
g	t_{m1}	g	t_3	g	g	t_{m2}	g	t_2	g	g	t_{m3}	g	t_3	g
g	g	g	g	g	g	g	g	g	g	g	g	g	g	g

Other modifications are possible, depending on the nature of the treatments.

A modification may also be needed if it becomes clear that one treatment is ineffective, and its continued presence may jeopardise the rest of the experiment. For example, one insecticide may be ineffective, so that plots with that treatment become sources of infestation. This can increase variability if, for instance, other plots down-wind are rapidly re-infested. The researcher should abandon the ineffective treatment and substitute an effective insecticide.

Specific trial types

In fertilizer trials, treatments may continue to be applied, regardless of their effects, or they may be modified if they have no effect or an adverse effect. Initial levels may need to be based on soil analysis, though a base level may be needed, regardless of soil analysis, if the experiment is to be repeated at several sites. Guard rows are necessary to prevent "poaching" of nutrients from neighbouring plots.

In irrigation trials, it is generally easier to use large plots of many trees as the main plot for irrigation treatments, dividing the main-plots into sub-plots for another treatment factor. The need for guard rows will depend on the method of irrigation and the topography.

Spraying experiments generally need large plots with ample guard areas. It may be possible to use screens while spraying, but in any case spraying may need to be done when wind speeds are minimal. A design with fewer "control" plots than treated plots -- including "standard" treatments -- may be useful. Unsprayed plots may need to be put on one side of the experimental area, to avoid having them act as foci of infection or infestation for treated plots. Treatments may need to be defined such that spraying is done based on a threshold level, rather than on a routine cycle basis.

For spacing trials, a fan design may be useful to determine the range of optimal spacings. Having done so, a randomised block experiment can be done, with a limited number of spacings. Such trials take a long time, since they start with planting and must go on through several years of bearing to get reliable results.

Bananas and sugar cane cause problems because plants can "wander" during the course of the experiment. The plant crop will have been planted at a particular spacing, but the ratoon plants will not be in exactly the same place. One way to reduce the effect of this, is to select -- in bananas at least -- followers close to the original position. Plants may need to be marked so that experimental and guard plants can be distinguished after several years.

For economic data, small plots do not give reliable results. Plots of at least 0.2 ha (0.5 ac) should be used. To assess the acceptability as well as the cost of a treatment may require spreading one block over several farms, or the use of incomplete block design.

CHAPTER 8

THE MANAGEMENT OF EXPERIMENTS

John L. Hammerton

Many decisions must be taken before starting any experiment. These decisions include, in crop trials, the variety to be grown, the basal fertilizer level, the methods of pest control and so on. In animal trials decisions on the methods of feeding, watering and littering, and slaughter weights must be taken. Treatment application must be done so as to minimise the chances of introducing error variation into the results.

The Planning Stage

There are some vital questions that need to be asked in planning a crop experiment:

* What basal (overall) management level to use? Low, medium, optimal, or high?

* What cultivar to use? A local "cultivar", or a high-yielding hybrid.

* What methods of land preparation? Manual or mechanical?

* What basal fertilizer level to use? None, a little, optimal?

* What spacing and plant arrangement to use? Ridges, mounds, on the flat? Even spacing, or several plants per "hills"? In rows, on the square, or what pattern?

* What method of planting? By hand, by seed drill. Any intercrop?

* What weed, insect, disease control methods to use? High or low technology? Good, bad or indifferent control?

* What method of harvesting to use? Selective, partial harvesting as the farmer may use, or complete harvesting?

The answers to these questions has implications for the relevance of our results to the farming community. The answers should complement our choice of treatments, and should be guided by our choice of treatments. It may be absurd to use a high level of pest control if this is beyond the resourses of the target farmers.

There are some vital questions that need to asked in planning a livestock experiment:

* What breed and sex or mix of sexes? Local "creole" or exotic breed? Males, castrates, females?

* What type of housing and litter to use? Concrete floors and pens or more primitive housing? Sawdust, straw, or no litter?

* What basic food and supplements to use? Grass plus concentrates or grass alone? Mineral licks?

* What system of feeding and watering? Ad lib or rationed?

* What ration or succession of rations? A standard ration or starter-weaner-finisher rations?

* What slaughter weight?

Again, our answers and decisions have implication for the **relevance** of our results.

From the planning stages should come a list of the inputs needed. It is better to anticipate what *may* be needed, rather than the minimum we hope to get away with. Recognise Murphy's Law, and use it in your planning. What equipment is going to be needed? Is it available, or will it be available? What consumables are needed, in what quantity and when?

AN EXAMPLE. Suppose we have an experiment on white fly and aphid control in hot peppers.

* Do we use a low, medium or high level of management? Something approaching the *average* farmers management level? Do we know what this is? This would ensure that the results would have wide application.

* Or do we use something resembling the *best* farmers level of management? Do we know what this is?

* Or should we aim to develop a moderately high technology complete tech-pack that includes white fly and aphid control?

* Do we use a wide spacing, local cultivar and local seed source, little fertilizer, "partial harvesting", etc. Is there any point in aiming for good insect pest control if other factors remain limiting?

Our input list will include the insecticides we are planning to use, hot pepper seed, fertilizer, bags for fertilizer, bags for harvesting, marker pens, plot pegs and labels, herbicides, etc. Our equipment list will include one or more sprayers, hoes, land preparation equipment, etc.

Our management level must depend on what we are trying to achieve and who our clientele are!

Execution and Management

The management of the experiment must aim at avoiding introducing any variations that are not associated with either treatments or blocks. This will ensure that any differences are confounded with blocks and not with error.

* Obviously plots should be of the same shape and size, and with the same number of plants. Animal pens should be of similar shape, size, and shading.

* If all the blocks cannot be weeded or sprayed with insecticide in one day, then weed or spray one or two *whole* blocks per day. Do not weed one-and-a-half blocks, or

one- and-two-third blocks in a day. Use different workers on different blocks, not on different plots within blocks.

* If the treatments cannot be applied to all the plots in one day, then finish one, two, or more blocks *completely*, in the day.

The Application of Treatments

Application of treatments and of overall management factors must also be uniform within plots. Otherwise the treatment is not tested precisely, and additional inter-plot variation is introduced. Also the same treatment must be applied in the same way to all plots receiving that treatment.

* In a fertilizer trial, the fertilizer must be uniformly applied to each plot regardless of the quantity. Do not attempt to apply it all in one pass over the plot. Take time, practice, and train staff to do it properly.

* In a trial where there is a basal fertilizer application to all plots, how should it be applied? If it is to be hand-applied it is better to do it plot by plot, rather than to try to apply a large amount evenly over alarge area.

* In a spacing trial, or in a hand-planted trial with uniform spacing, it may save time to prepare marked cords, rather than use a tape measure.

* Sprayers, fertilizer applicators, granular applicators, sprayers, etc., must be calibrated to be sure that they do apply the planned quantity.

CHAPTER 9

REGRESSION, CORRELATION AND COVARIANCE

Bruce Lauckner

Regression can be used as a way of analysing data, but the regression should be analysed to obtain the standard errors of the coefficients. Correlation is useful to indicate the relationship between two variates. Covariance is used to adjust a measured variate for another that is not a consequence of treatment.

Linear regression

One use of regression has already been mentioned in Chapter 4: to investigate the growth rate of animals over a period of time. Other uses are to look at plant growth over time, the spread of diseases or insect pest infestation against time, and to investigate the relationship between height and weight of a plant.

Table 9.1 shows the weight of an animal at weekly intervals over a period of six weeks. We can try to fit a straight line to represent this relationship. The equation representing this line is:

$$y = a + bx \tag{9.1}$$

where a and b must be calculated from the data values.

Table 9.1.
The live weight of an animal at weekly intervals.

Time (weeks)	Weight of animal (kg)
1	14.3
2	15.4
3	16.7
4	17.8
5	19.1
6	20.5

The time, in weeks, is known as the independent variable, and is usually represented by x. The dependent variable, in this instance weight, is usually represented by y. It is important to establish which variable is the independent and which is the dependent one.

a and b are estimated from the following formulae:

$$a = y - bx \qquad b = \frac{S_{xy}}{S_x^2} \tag{9.2}$$

where S_x^2 is the corrected sum of squares for x, givenby the uncorrected sum of square minus the correction factor (similar to analysis of variance calculations), and

S_{xy} is the corrected sum of products for x and y, given by the uncorrected sum of products minus the correction factor (see the worked example).

x and y are the means of x and y.

From the example:

the uncorrected sum of squares for x is

$$1^2+2^2+3^2+4^2+5^2+6^2 = 91$$

and the correction factor = $\dfrac{(1+2+3+4+5+6)^2}{6} = 21^2 = 73.5$

Thus $S_x^2 = 91 - 73.5 = 17.5$

The uncorrected sum of products of x and y is

$$(1x14.3)+(2x15.4)+(3x16.7)+(4x17.8)+(5x19.1)+(6x20.5)=384.9$$

with the correction factor =

$$\frac{(1+2+3+4+5+6) \text{ x } (14.3+15.4+16.7+17.8+19.1+20.5)}{6} = \frac{21\text{x}103.8}{6} = 363.3$$

Thus $S_{xy} = 384.9 - 363.3 = 21.6$

$$x = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$$

$$y = \frac{14.3+15.4+16.7+17.8+19.1+20.5}{6} = \frac{103.8}{6} = 17.3$$

Thus we obtain

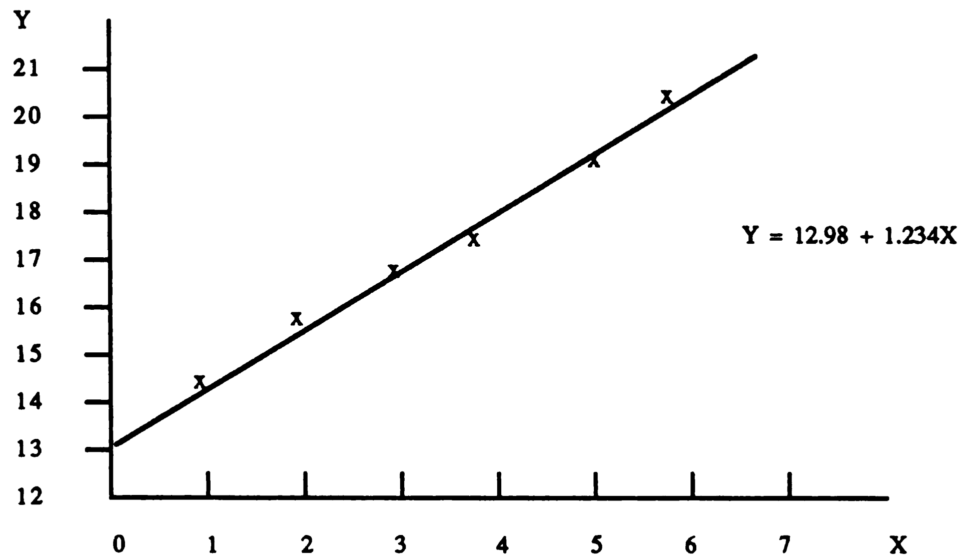
$$b = \frac{21.6}{17.5} = 1.2348 \qquad a = 17.3-(1.2348\text{x}3.5) = 12.98$$

Inserting these values in equation (0), we obtain the regression equation

$$y = 12.98 + 1.2348x \qquad (9.3)$$

The regression line and the observed values are shown in Figure 9.1. Note that it is always a good idea to plot data, such as these, on a graph to see what the relationship is.

Figure 9.1. The observed values (from Table 9.1.) and the fitted regression line.



The values of a and b are not very useful unless we know their standard errors. To obtain these we must complete an analysis of variance of the regression line. First we must calculate the total variation due to y. This is given by Sy2, the corrected sum of squares.

$$S_y^2 = 14.32+15.42+16.72+17.82+19.12+20.52$$

$$- \frac{(14.3+15.4+16.7+17.8+19.1+20.5)}{6}$$

$$= 1822.44-1795.74 = 26.70 \qquad (9.4)$$

Next we need to calculate the amount of variation "explained" by the regression, and this is given by the regression sum of squares, which is calculated as bS_{xy}. These quantities have already been calculated.

Regression sum of squares, $bS_{xy} = 1.2348 x 21.6 = 26.661$ (9.5)

The difference between the values of (0) and (0) is the amount of variation in y not explained by the regression, or the error sum of squares (the uncontrolled variation). The total sum of squares (S.S.) has n-1 degrees of freedom (d.f.), where n is the number of observations. The regression sum of squares has 1 degree of freedom, and the error sum of squares has n-2 degrees of freedom, and is given by the difference between the total and regression sums of squares.

The analysis of variance table is therefore as follows:

Source of variation	**S.S.**	**d.f.**	**M.S.**	**F**
Regression	26.661	1	26.661	2734
Error (Residual)	0.039	4	0.00975	
Total	26.700	5		

The F (Variance ratio) test, with 1 and 4 degrees of freedom, shows that the regression is very highly significant. This only indicates that the regression is significantly different from a random scatter: it does not indicate that the regression equation is the best model describing the relationship between x and y.

The standard deviation (s.d.) is calculated as the square root of the error mean square, thus:

$$s.d. = 0.00975 = 0.0987$$

The standard errors (s.e.) for a and b are given by:

$$s.e.(a) = s.d. \frac{x^2}{nS_x^2} \quad \text{and} \quad s.e.(b) = \frac{s.d.}{S_x^2} \qquad (9.6)$$

where x^2 is the uncorrected sun of squares for x, and, as before $S_x{}^2$ is the corrected sum of squares for x, and n is the number of observations.

Thus s.e.(a) = $0.0987 \quad \frac{1^2+2^2+3^2+4^2+5^2+6^2}{6 \times 17.5} \qquad 0.092$

and s.e.(b) = $\frac{0.0987}{17.5} \qquad = 0.024$

From these estimates of a and b and their standard errors, t tests can be done to test the signifacance of both a and b, and confidence intervals can also be obtained for a and b and for the regression line.

Thus the value of t for a = $\frac{12.98}{0.092} \qquad = 141.1,$

and the value of t for b = $\frac{1.2348}{0.024} \qquad = 51.45.$

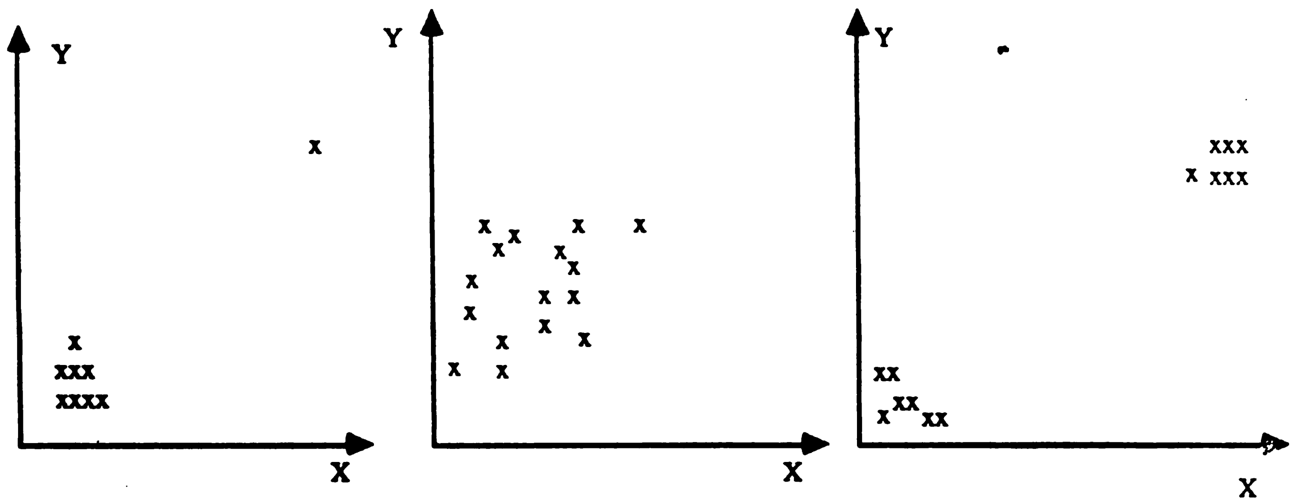
Both these t values, with 4 degrees of freedom, are highly significant.

Note that a represents the value of y when x = 0 (i.e. the intercept on the y-axis) and b represents the slope or gradient of the regression line. If b = 0, the line is parellel to the x-axis; if b = 1 the line is at 45o to the axes; and if b = oo, it is parellel to the y-axis.

Graphing

The advisability of plotting data to look at the relationship has already been mentioned. Figure 9.2. shows some of the problems. The data set for the first graph would give a good, and significant, regression line, but the line is dependent on the single outlying data point. Without it there is no regression! The data set in the second graph will not give a significant regression, and there is clearly no relationship between the two variables. The third data set have something missing: the data are in two groups. A linear regression could be fitted, but it assumes that nothing unusual happens between the two groups of data points.

Figure 9.2. Three contrasting data sets graphed.



Given a satisfactory regression line, it must not be misused. The regression equation must not be used to predict a response outside the range over which the data were collected. To do so assumes that the relationship holds outside the range over which the data were collected. Confidence limits can be calculated, and these will show that predicted values at the extremes of the line, and beyond, are not very accurate.

Correlation

The correlation coefficient (r) between x and y is given by the value:

$$r = \frac{S_{xy}}{\sqrt{S_x^2 S_y^2}} \qquad (9.7)$$

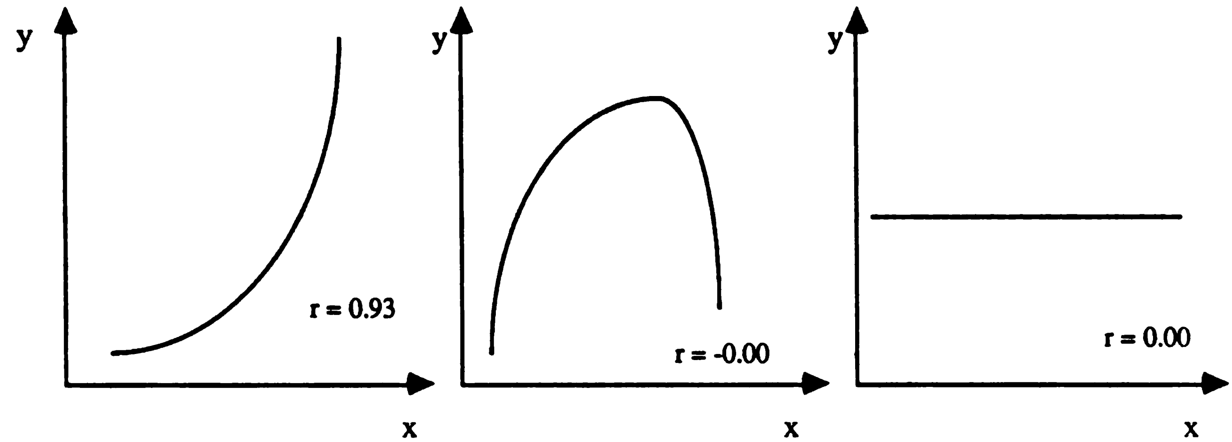
The correlation coefficient is a measure of linear association. The value lies between -1 and +1. A value of 0 indicates no relationship between x and y: that one variable changes regardless of the value of the other. Positive values indicate that x and y tend to increase together. Negative values indicate that small values of x are associated with large values of y, and vice versa.

In our numerical example $r = \dfrac{21.6}{17.5 \times 26.7} = 0.9993,$

indicating a very close relationship of x and y. r does not prove that the relationship is causal: x and y may have no direct causal relationship, but both may be casually related to a third variable z. For example, banana sales in the U.K. in the first few years after World War II, and the number of road accidents over the same period, are positively correlated. Does banana consumption cause road accidents then? Or do road accidents somehow stimulate a demand for bananas? The mind boggles with possible explanations. In fact, banana imports, the number of cars on the roads, the amount of petrol available (and lots of other things) steadily increased after World War II.

Data can be strongly associated but have a low correlation because the relationship between the two variables is not linear. Figure 9.3. shows various associations and their correlation coefficients. Note that a high correlation coefficient does not necessarily mean that the relationship is linear.

Figure 9.3. Digrammatic representation of various associations and the corresponding correlation coefficients.



The value r^2, the square of the correlation coefficient, is often calculated and expressed as a percentage. In our example 0.9932 = 0.9985 or 99.85%. This can be shown to be equivalent:

$$\frac{\text{Regression sum of squares (Sxy)}}{\text{Total sum of squares (Sy}^2\text{)}}$$

It indicates the percentage of variation in y accounted for by the regression line. The goodness of the regression model should not be judged on the value of r2, but on tests of significance of a and b.

Multiple regression

Regression equations of the form

$$Y = a + bx + cx^2$$

can also be fitted. Such equations gives a curve (a parabola) and not a straight line. Other multiple regressions might be of the form

$Y = a + b_1x_1 + b_2x_2$, or $Y = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots$

where x_1 might be rainfall, x_2 might be mean temperature, x_3 might be daylight hours, and so on. It is important to test the significance of each coefficient, rather than rely on looking at the value of r^2, the square of the multiple correlation coefficient. There is a danger of "over-parameterising" the model by including variables that contribute very little information to the model.

Covariance

Where a variable that is not a consequence of the applied treatments is observed to vary in value over the experimental area, or from plot to plot, analysis of covariance can be used to adjust the means of a primary data variable. For example, the number of plants per plot, or the number of weeds, or an index of disease infection, can be used as a covariate to adjust yield, some other primary variable. In animal experiments initial liveweight, or the number of animals surviving, can be used to adjust final liveweight. Covariance is not a remedy for poor blocking. The statistical procedure is not difficult, but the assistance and advice of a biometrician should be sought in instances where covariance seems indicated.

CHAPTER 10

INTERPRETATION OF RESULTS

Bruce Lauckner

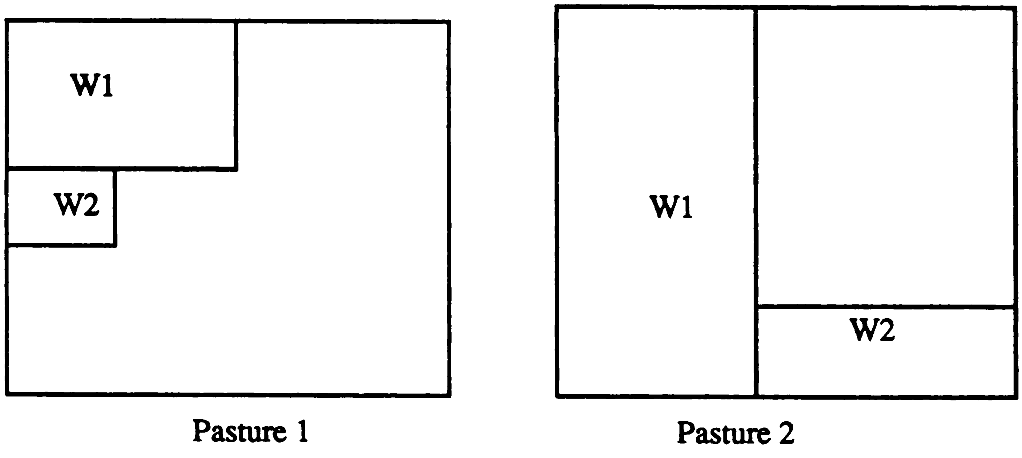
Experiments must be well designed if a valid interpretation of the data. Interpretation also depends on good methods of data measurement, and on reliable data. In addition, interpretation of data may deprnd on background data, or general observations.

The valid interpretation of any data set, or of any analysis, depends on the reliability of the data collected. It is not difficult to collect poor data from which no valid conclusions can be dawn: care must be taken therefore in deciding what data will be collected, how they will be collected, and how they will be analysed.

What sort of data?

To illustrate the importance of collecting appropriate data, assume that some data were collected on the weed populations of two pastures (Figure 10.1). The purpose of the observations was to assess the effect of different pasture management systems on weediness.

Figure 10.1. Distribution of weeds in two pastures.



The Figure shows the distribution of two weeds in the two pastures. On Pasture 1, W1 is the most common weed, so was given a score of 100%. W2 occurs on only one quarter of the area covered by W1, so was given a score of 25%. On Pasture 2, W1 is again the most common weed, and again is given a score of 100%, and W2 again occupies only one quarter of the area occupied by W1 and is agained scored at 25%.

What do these data tell us? W1 clearly occupies larger areas than W2 in both pastures. But is the average value of W2 25%? That does not make much sense. And the data do not allow us to compare the two pastures. Pasture 1 has far fewer weeds than Pasture 2, but the data do not reflect this. Obviously these data are incapable of giving the information required.

Another type of data that are often difficult to interpret come from "eye" estimates, frequently of disease damage. Such data are often so inaccurate and meaningless as to be useless. Compare the "leaves" in Figure 10.2., and estimate the percentage of each leaf covered by the black "lesions". The answers are on page 00. Few people can accurately estimate the percentage area affected by disease. This is because the eye is not good at assessments when a lot of damage has been done. Experience suggests that only one researcher in a hundred can do this eye assessment accurately. Data of this type become even less reliable when more than one person is making the observations.

Experiments must be well-designed

If an experiment has not been well-designed, no analysis, however complex or sophisticated, can remove doubts concerning its interpretation. Even if adjustments are made to remedy some of these deficiencies, the price paid will be a loss of precision, compared with a well-designed experiment. Figure 10.3 is the field plan of an experiment to test the effect of intra-row and inter-row spacing on two varieties of soyabean (UFV-1 and PR140(11)). Note that the layout lacks randomisation: UFV-1 is always to the left of PR140(11), and beds with two rows are always to the left of the beds with three rows. In addition, spacing treatment A does not occur in the left-hand section of the experiment. If there is a gradient across the field, interpretation of the data collected would be very difficult.

Figure 10.2. Brown spot of sugar cane: different percentages of leaf area infected.

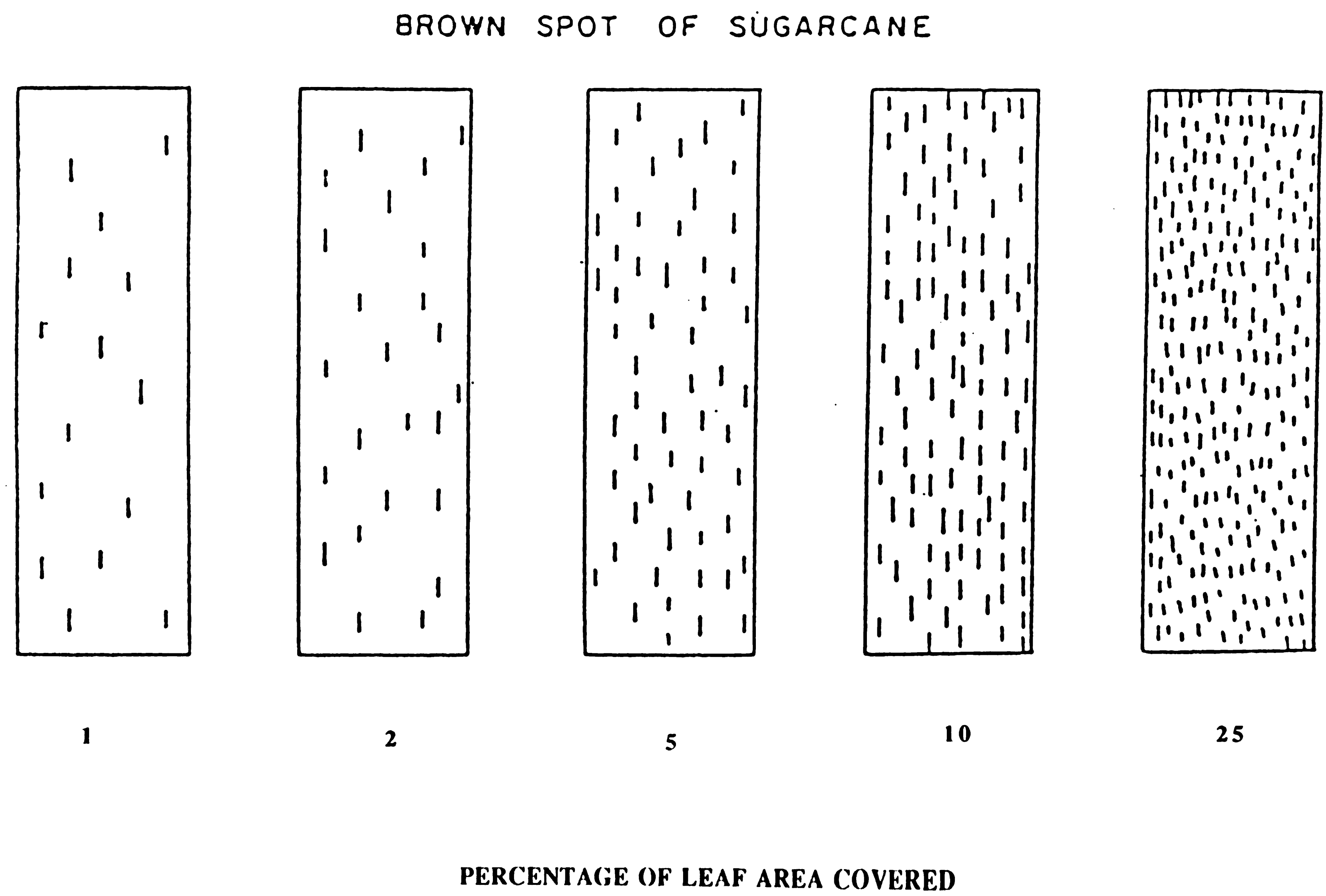
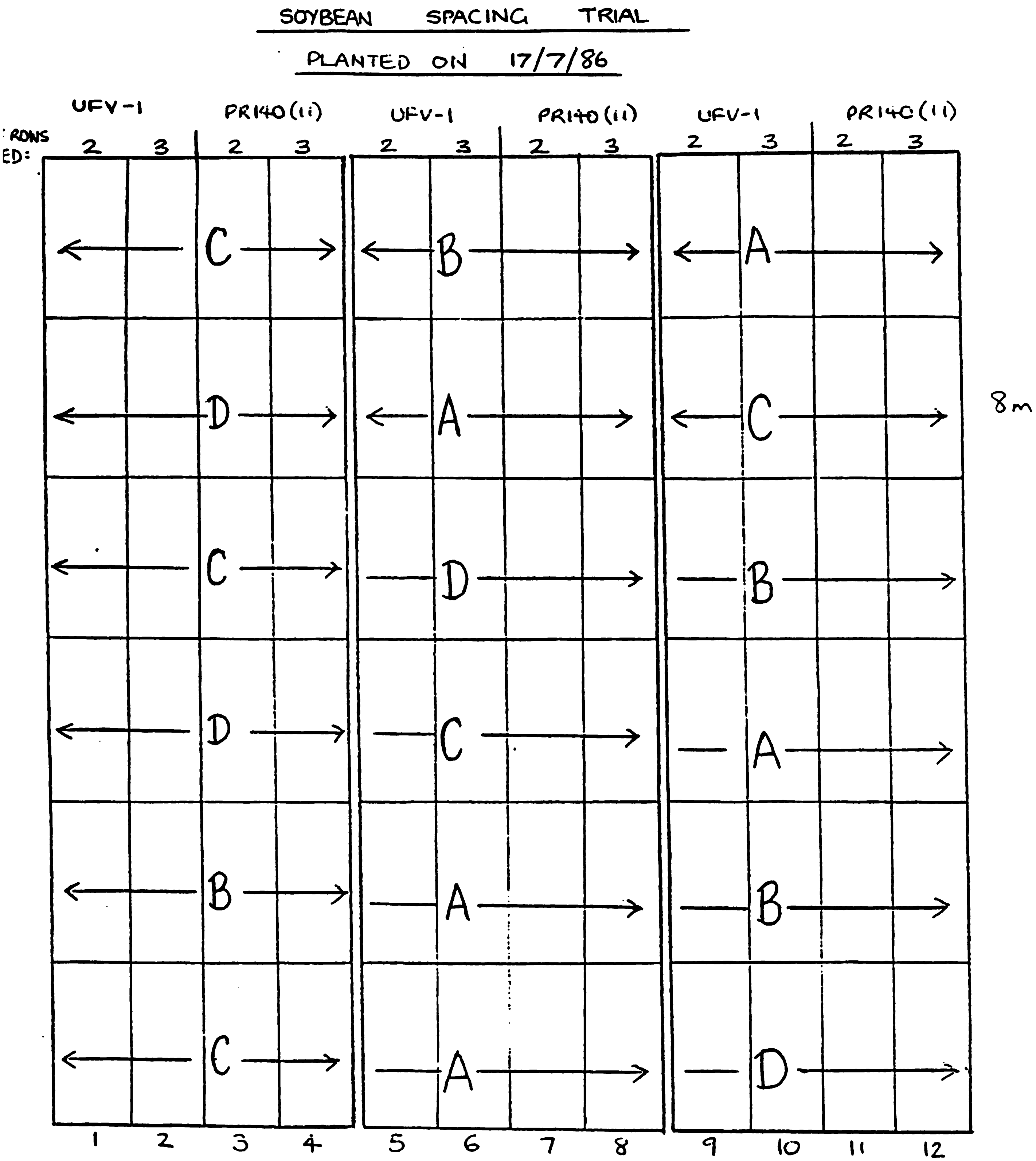


Figure 10.3. Field Plan for a spacing trial on soyabeans.



A = 26 plants / m
B = 22 plants / m
C = 18 plants / m

Need for measure of variability

No interpretation of means, nor any real use of those means, can be done unless some measure of variability is given. It is essential that the standard error of the means (s.e.m), standard error of the difference between any two means (s.e.d), standard deviation (s.d.) or variance be given whenever a mean is quoted. Consider the following example:

Mean of treatment A	60
Mean of treatment B	100

It is not possible to know or to state with any confidence, that B is greater than A, unless the s.e.m or some error value is known. If the s.e.d = 10, then it is fairly safe to assume that B is greater than A (because (100-60)/10 give a value of t greater than tabulated values for an acceptable probabability level). If s.e.d = 25, such a conclusion would not be drawn (because (100-60)/25 is less than 2 which is not significant at an acceptable level of probability). The conclusion depends therefore on the variability.

Use of other variables

There is often a need for background information, not only to help in interpretation, but to ensure that the most appropriate analysis is done. To only record the number of live insects after spraying with an insecticide cannot determine the "best" insecticide. The number of insects before spraying must be recorded. Then analysis of covariance could be used to give results that are readily interpreted.

Sometimes unexpected results may be obtained. This could be due to weather factors, which may vary sufficiently to cause variation in the results. Information on relevant weather factors may therefore be needed.

A change of personnel can affect how data is collected, or in animal experiments, how the animals behave or how thoroughly the feed is mixed. Such seemingly unimportant facts need to be recorded and taken into account in interpreting any unexpected irregularities in the results.

If germination is poor or uneven, final yields may not be directly comparable. The number of plants per plot should be counted, and these data used as a covariate in an analysis of covariance. Some care may be necessary to ensure that the use of a covariate does not render real differences apparently non-significant. For example, in a variety trial, adjusting (by covariance analysis) for differences in plant stand or in insect damage, may render varietal differences non-significant. Yet there may be real and important differences between varieties in germination percentage or susceptibility to disease. Careful examination of the data are necessary before using a covariate to adjust means.

Treatment structure

A full understanding of of the treatment structure and the purpose of the experiment are essential to an understanding of the results. With a factorial structure the analysis of variance table must be read from the bottom up: interactions must take precedence over main effects. An interaction occurs when no uniform recommendation can be made for main effects. For example, a table of means might look as follows:

	Spacing			
Variety	30x30	30x20	30x10	Mean
A	50	60	70	60
B	70	80	90	80
Mean	60	70	80	

Here there is no interaction. Both varieties show the same response to spacing. Only the main effects (plus a measure of variability) need be presented. The 30x10 spacing and variety B could be recommended. Compare with the following table:

	Spacing			
Variety	30x30	30x20	30x10	Mean
A	50	60	70	60
B	70	70	70	70
Mean	60	65	70	

Here there is interaction. Variety B is unaffected by spacing, but A is affected. No one spacing is best: it depends on the variety. Compare again with the following table:

Spacing

Variety	30x30	30x20	30x10	Mean
A	50	60	70	60
B	70	60	50	60
Mean	60	60	60	

Here also there is interaction. The two variety both respond to spacing but in opposite ways. No general recommendation for spacing can be given. Note that their are no main effects.

Further, if no use is made of the factorial structure, paired comparisons might be made, which is incorrect. The data should not be squeezed to do more tests than allowed by the design or than initially planned, otherwise it is possible, particularly if an inappropriate test is done, to obtain misleading "significant differences".

Loss of a treatment

One particular treatment may sometimes "fail" so badly that it gives nothing, or very low yields compared with the other varieties. This may happen in a variety trial, or perhaps with one herbicide that proves to be phytotoxic. The distribution of the data will then be far from normal, and the treatment in question should be dropped from the analysis. The report should not ignore this treatment, but should record what happened, and state of course that that treatment was dropped from the analysis of variance.

Other pointers to proper interpretation, and the dangers of misinterpretation are given in Chapter 9 (Regression, Correlation and Covariance).

CHAPTER 11

REPORTING AND WRITING-UP

John L. Hammerton

No piece of research is complete until the results have been analysed, evaluated, and reported or communicated to someone, or more usually, to some group or groups of people.

The individuals or groups to whom results should be communicated can be one or more of the following:

* The (Worldwide) Scientific Community.
* The General Public.
* One's Immediate Colleagues.
* A Funding Agency.
* The Minister of Agriculture.
* The Chief Technical Officer, the Director of Research, the Deputy Executive Director or some similar person.
* The (entire) National Farming Community.
* A specific group of Farmers.
* The members of a Commodity Organisation.
* The Farmers in a specific location.
* Extension Officers.

Obviously, the type of report must be appropriate to the individual or group, and if reports must be made to more than individual or group, then more than one type of report may be needed.

Reports can be written, verbal, visual, or a combination of more than one of these. We will consider only written reports These are in any case essential as they provide a "permanent" record.

Purpose

* ***What is the intended target readership of the report?*** This must be determined, since different readerships have different requirements, different reading skills, different vocabularies, etc.

* ***What should the report accomplish?*** Inform: to simply report what you did, and what the results were? Change attitudes: to persuade people to think differently? Change ways of doing something: to stir the reader to action? Justify: to account for the expenditure of time and cash. Impress: to show how clever you are?

The approach, the language, the emphasis will vary with what the report is intendedy to accomplish. And more than one thing can be accomplished by a single report if it is written to do so.

Some important questions

* ***What do the target readers already know? Do they know (for example) the following.*** The location you are writing about: its physical, agricultural and socio-economic characteristics? The importance of the crop, livestock, crop/livestock combination in that location? The farming systems, and the technology used? The importance of the crop(s), livestock, or commodities in the national economy? The basic agronomy of the crop, or the husbandry system of the livestock?

  What the reader already knows will vary widely with the target readership. Targetting the report will avoid over- or under-shooting: insulting the readers' intelligence, or going way over their heads.

* ***What do the targetted readers need to know? Do they need to know one or more of the folowing.*** Why it was done. What was done, when, and how. What was found out. What these findings means. Why these findings are important. The next steps.

  The emphasis will vary with the readership. A farmer group does not need to know all the treatment details, the design, the method of data analysis, etc. Scientists may not be so interested in the economics, or the likelihood of adoption.

Multiple reporting

Suppose a three-year study was done on the synchronisation of fruiting in pineapples using several "inducers". This could be reported in several different ways at several different levels:

* A *scientific article* on the physiological basis for the effect of the inducers used (enzymes, sites of action, sugar contents, biochemical analyses, chromatograhy, etc.).

* An agronomic report on rates, timings, yields and fruit quality (yields, grading, rates of application, timing of application and harvests, analyses of variance, etc.).

* An economic report on costs and returns and profitability (partial budgets, dominance analysis, returns to labour and management, etc.).

* A Factsheet for farmers on why, what, and how to use inducers (what the benefits are in terms of cash, the costs, how and when to apply, etc.).

A framework

Almost regardless of the target readership and the level of reporting, the following framework is useful as a checklist.

* **Introduction:** why the work was done, its objectives, its relevance, the problem in perspective, the justification, etc.

* **Methods:** where and how the work was done, what cultivars where used, what spacing, what crop care, the timing, etc.

* **Results:** what was found out, the yields, the quality, the interactions, the analyses and statistical significance, economic analysis, tables, figures, pictures, etc.

* **Discussion:** the agronomic, economic, practical, and social significance, the relevance of these findings, the conclusions, etc.

* **Acknowledgements:** who helped in any way (beyond the call of duty?), who provided land or animals, who analysed the data, etc.

* **References:** where appropriate--but in a conventional form!

For a scientific paper or for a technical report, these headings might well be used as they are. Tables and figures would be included and there would be statements of significance. References would be listed and cited in the text. For a popular article these headings are inappropriate, and more popular headings may be used. There would be less attention to detail and to statistics, and more on economics and relevance to practical farming.

> ***What goes in and what stays out depends on who you are writing for. There are no options on honesty, or on proper and careful conduct of the research, or on proper and thorough analysis of the data, however.***

CHAPTER 12

WRITING: COMMON ERRORS

John L. Hammerton

Even experienced scientists make mistakes in their reports, or write poor, badly-expressed, and sometimes misleading reports. There are some at least 20 error that are common in scientific writing.

The following are common errors in written reports. The list is based on a number of editing exercises done on papers prepared for the Annual Reports of a number of Ministries of Agriculture.

Poor titles: Titles of papers are often misleading, and fail to convey the subject of the paper.

Poor abstracts: Some abstracts are loaded with data, and almost unreadable. There is no attempt in most abstracts to summarise trends, major findings, etc.

Poor writing: Some authors fail to express themselves clearly. They fail to say what they need to say--or what they, presumably, meant to say. The result is that they write something misleading or patently not true. "Objectives"--when specified--sometimes bear little relationship to the work reported on. Many opening sentences lack impact. Punctuation is poor: in particular, semi-colons (;) are often used where a comma (,) would be more appropriate. There is also confusion in the use of "may", "might", and "could"; and between "will", "would", and "should".

Padding: A few authors introduce extraneous material, or make passing reference to something of doubtful relevance. Circumlocutions are common: "at this point in time", "weighing was done on a fortnightly basis", "in the case of", "in relation to", and so on.

Repetitions: Some authors repeat some simple fact, without adding any new insight: a form of padding. Some repeat in their "Discussion", what they have already said in their "Results".

Excessive capitalisation: "1 Gal. of Molasses Weighs 10 lb.", and "Average Liveweight Gains were...", are examples. Not too frequent a mistake.

Confusion of common and proprietary names of pesticides: Common names of pesticides are conventionally written with a lower case first letter (except, of course, at the start of a sentence). Proprietary name are written with an upper case first letter wherever they occur. For example, "oxamyl", but "Vydate"; "carbofuran", but "Furadan"; and "diphenamid", but "Enide". In fact it would be better still to give the full proprietary names, and indicate that these are registered trademarks, thus: Vydate(R) L, Furadan(R) 10G, Enide(R) 90W. Authors often do not observe this convention, and also mix up common and proprietary names in any list. This mixing leads to confusion as to whether application rates are expressed as lb ac^{-1} a.i. or lb ac^{-1} product. It is preferable to give both proprietary and common names, and to indicate clearly how the rates are expressed thus: "Furadan(R) 10G (carbofuran) at 2.5 kg ha^{-1} a.i.....".

Tense: Inconsistent use of tenses is not common, but does occur: "...fertilizer is applied every two months, and irrigation was given weekly", and "the data showed....and the data also show".

Passive voice: Some authors lose impact by using the passive voice. "Fertilizer application increased yields..." (active), compared with "Yields were increased by the application of fertilizer..." (passive).

Pluralisation: "Data was", or "data is", are rare, but they occur occasionally. There is some confusion of "fertilizer" and "fertilizers", "seed" and "seeds", sometimes with the wrong verb form.

Poor proof-reading: Few authors seem to have proof-read their papers--or if they did, they did a very sloppy job--so they have failed to pick up non-sequitors, missing lines, missing data, etc.

Unspecified experimental designs and no details: The experimental design is not always clearly stated. It can sometimes be deduced from "clues" given in the text, but not always. Plot sizes and basal management details are also often ommitted or fudged over.

Lack of correlation between text and tables: The text refers to the wrong table--or to no table at all--leaving the reader to find the relevant table! This is not helped by poor table headings--or no heading at all!

Missing data: This occurs particularly where lists of data are given. For example, "..average live-weight gains for the five years, respectively, were 123, 134, 129, 133." Probably indicative of poor proof-reading. No units given.

Overkill: A few writers tend to make too much of very little data.

Suspect data: Data in tables does not always support statements made in the text--or vice-versa. Are the data correct--and the statements wrong? Or are the statements correct--and the data wrong? Can be resolved only by reference back to the author.

Too many decimal places: There is often an obsession with decimal places--even with 235.0, 233.0, 245.0 etc. And 89.65%, 91.26% and so on. A reluctance to round-off, or the assumption that if a computer printout gives you so many places of decimals, these are inviolate?

"Indefinite" statistics: Some tables have L.S.D.'s, LSD's, or l.s.d.'s or "SIG. DIFF.'s"; others have nothing at all. Some of the L.S.D.'s given are suspiciously low. They may be the S.E.'s, or simply incorrectly calculated! Some authors write "the differences were significantly different at P=0.01"--but the table has no L.S.D.'s or S.E.'s.

Units: kg/ha, kg per ha, or KG/HECTARE, or what? A lack of standardisation! Clumsy units include "kg a.i./ha", and "kg N/ha", and uncertain units, such as "tons/ha". The former would be better expressed as "kg/ha a.i." and "kg/ha N", and the latter should be clarified: tons (2240 lb) or tonnes (1000 kg)? The abbreviation for tonnes is "t". "Per year" and "per annum" are both used. Some authors use only metric units, others use metric units and give the non-metric equivalent. "GM" or "gm" is used--incorectly--for gram or grams: the recognised (SI) abbreviation is "g".

Poor tables: Many tables are badly designed; column and line headings are often poorly phrased, or abbreviated in such a way as to confuse. Table headings may be missing, inadequate, and even too long! Some writers proliferate tables--they have three or four where one consolidated table would do. Other writers include individual plot values, totals, and means--but have usually done no analyses. The order of the treatments sometimes varies between tables--which can confuse the reader.

Inconsistency in citing and listing references: Different methods of citation are used in the text: *"Smith (1976) found that.."*, or *"...gave an increase in yield (8)"*. References are sometimes not listed in alphabetical order, may or may not give the title of the paper or article, and sometimes use non-standard abbreviations for journals etc.

EDITING IS QUALITY CONTROL.

CHAPTER 13

MEASURING READABILITY

John L. Hammerton

Reports are meant to be read: otherwise why write them? There is sometimes a temptation to write in "heavy", "traditional" or "formal" styles, that are often difficult to read. "Readability" can be fairly easily be measured, however.

Do you write conservatively laboured long-winded multi-faceted sentences utilising intricate structural complexities, resplendent with abundances of lengthy poly-syllabic peripheral referential adjectival expressions, redolent of turgidity and liberally replete with jargon, in compositions parsimonious in succinctness, purposefully to impress your readership with your erudition?

Or do you write short, simple sentences? Do you use simple, familiar words? Do you avoid long words and jargon, unless there is no alternative? Do you write for specific target audiences? Do you try to gain and hold your readers' interest? Do you think through what you want to say, and try to find the best way to say it?

Whichever of these is your style, you need to be able to estimate the **readability** of you writings. Then you can adjust your style to your target readership.

MEASURING READABILITY

Two indices of readability are described: **Flesch's Readability Formula (R)**, and the **Gunning Fog Index (FI)**.

1. For a short piece of writing, count all the words. For a longer piece, take one or more 100-word samples, depending on the length. Count as a word all numbers, letters or symbols, and all groups of numbers, letters or symbols *surrounded by space*. Contractions, acronyms, abbreviations, initials, numbers and hyphenated words should all be counted as words. Thus *CARDI, C.D.B., 1989, EC$678, can't, Sept., 100-word, etc.*, each count as one word.

2. Calculate the **average sentence length (ASL)**. Count the number of sentences in your piece or in your sample(s). If in a sample the 100-word mark falls part-way through a sentence, as it normally will, include that sentence if more than half of it falls within the sample. If less than half falls within the sample, do not count it. Do count as a sentence any independent clauses ending with a semi-colon, a colon, a question mark, or an exclamation mark. You will have to use your judgement in this. The sentence "The treatments were: no Benlate, Benlate once a week, and Benlate once a fortnight.", should

be counted as **one** sentence, not two. ("The treatments were:" cannot stand alone as a sentence). But the sentence : "Annual reports must be submitted by November 30th: this deadline must be adhered to to allow time for editing and printing.", should be counted as **two** sentences. (The colon could be replaced by a full stop).

3. Count the number of syllables in the piece, or in your sample or samples. Count them the way you would pronouce them. For example, "spray" and "sprayed" each have one syllable, "statistics" has three, "CARDI" has two, "C.D.B." has three, "$" has two, 1987 has six ("nine-teen-eight-y- sev-en"), "Sept." has three (since it would be read or pronounced "September"), and so on. Record the number of **syllables per hundred words (SPHW)**. (If necessary multiply the average number of syllables per word by 100).

Flesch's Readability Formula (R) is calculated from:

$$R = 0.0778(ASL) + 0.0455(SPHW) - 2.209$$

Interpret **R** as follows:

4.0 to 4.5	Very easy
4.5 to 5.5	Fairly easy
5.5 to 6.5	Standard
6.5 to 7.5	Difficult
7.5 +	Very difficult

The **Gunning Fog Index (FI)** uses the **ASL**, and the number of **words with three or more syllables (WTMS)**. Exclude from the **WTMS** count combination words and capitalised proper names. For example, exclude "kilograms", "radio-active" (*combination words*), "Australia", "Cincinnati", and "Dr. Hammerton" (*capitalised proper names*). See paragraph 3 above for advice on syllable counting.

The Gunning Fog Index (FI) is calculated from:

$$FI = 0.4(ASL + WTMS)$$

Interpret **FI** as follows: it approximates the number of years of education needed to read and understand the piece of writing. Popular writing is usually around 7 or 8. *The danger line is about 13.*

Check your readability!

APPENDIX 1: PROBLEMS AND ANSWERS

PROBLEMS 1: OBJECTIVES

Read very carefully and be very critical.

Think.

1. Pa louis mangoes have a quality problem on the United Kingdom market. *Suggest a series of agronomic investigations to remedy this problem.*

2. Sheep in the Bexon-Odsan area have heavy worm burdens. *Suggest the sort of research needed to reduce the worm burdens and increase mutton production.*

3. Hot peppers in St. Lucia are often badly infected with a virus complex that causes leaf mottling, reduces leaf size, and reduces yields. This virus complex may or may not be seed-borne. *Design a programme of research to investigate whether the virus complex is or is not seed-borne.*

4. Single tubers of twelve cultivars of Dioscorea rotundata have been received for testing. Obviously the mini-set technique should be used to get the maximum number of plants. *Design an appropriate experiment, indicating plot size, shape, and number of replicates.*

ANSWERS 1

There are no single correct answers. Those given below are examples of what could be considered correct answers.

1. We cannnot design a research programme of any sort until we know what the problem is. It may not be an agronomic problem at all. We need to find out what exactly is the problem.

2. We do not need to do any research at all: the technology and the dewormers are well-known. What we need to do is check out the farming systems in the area. These are mainly banana-based, and what we would probably find is that the farmers are not particularly interested in deworming, since sheep are a minor part of their system. so we might do better to try deworming somewhere else.

3. Do we need a programme of research? Why not check the literature first? More urgent is to find a source of virus-free seed of good hot pepper type, either a local or overseas source. If we wish we could try growing some peppers from seed taken from plants known to be infected, and from plants known to be healthy, and see whether we get healthy or infected plants, and in what proportions.

4. It would be premature to design an experiment. We should bulk up first from mini-sets until we have enough planting material to really compare the cultivars.

PROBLEMS 2: MANAGEMENT

Read very carefully and be very critical.

Think.

1. We have to apply Fusilade(R) 2000 and Basagran(R) to a herbicide experiment on beans. Our sprayer has a spray swathe of 1 m, but our plots are 2.5 m wide (5 rows at 50 cm). What do we do?

2. Its 4.30 p.m., and we still have 3 plots to spray to finish the experiment. We've been in the field from 10.45 a.m., and we are tired. What do we do.

3. We have three fertilizer levels to apply, each to two cultivars of yam in each block. In other words we have a 2 x 3 factorial treatment array. There are 4 blocks. We have two field assistants to do the application (by hand). How do we organise the work?

4. In a spacing experiment with cowpeas, a full stand at the designated spacing for each treatment is essential. Row width is the same for all treatments: it is inter-row spacing that will vary (7.5, 10.0, 12.5, or 15.0 cm). How can we ensure (as far as humanly possible) that we get a full stand at the designated spacings?

5. We find that some of our plots of sweet potato have sizeable, irregular patches of *Panicum maximum*. Our experiment is a herbicide experiment, but none of the herbicides we are evaluating will control *Panicum*. What do we do?

6. Three of our livestock pens are exposed to the full sun from about 11.30 a.m. onwards each day. This apparently cause considerable heat stress in some of our experimental pigs. What do we do?

ANSWERS 2

There are no single correct answers. Those given below are examples of what could be considered correct answers.

1. It should not have happened: we should have made our plot width an exact multiple of our sprayer swathe width-- either 2 m (4 rows) or 3 m (6 rows). What we must do now is lower our nozzle to give a swathe of 0.83 m (2.5/3). This means recalibrating, and tying a chain or cord plus weight to our nozzle to help us maintain the correct height.

2. We finish, especially if the three plots are a fraction of a complete block! Of course, if the three plots are a complete block (i.e., our experiment has only three treatments), we could leave that block over to the next day, but we should still try to finish.

3. We assign each field assistant to do two blocks, one block at a time. We also train them and supervise them to be sure that they do it right. It might be as well to break the total quantities down into smaller lots, if not into the amounts for one mound then into the amount for four, five or six mounds. If the yams are on ridges then we break it down into the amount for one ridge.

4. If we are planting by hand, we can use marked cords to get the required intra-row spacing, and sow two seeds per hole, to be thinned to one seedling per station later. If we are planting with a seeder, we could plant at a close spacing and thin to the required spacings later, again using marked cords.

5. We control the P. maximum by means of a post-emergence grass weedkiller, spraying the entire experimental area (and not just the patches) just in case there is some effect on the sweet potatoes.

6. We modify the building by extending the roof and improving ventilation. Or if we have sufficient space we avoid using those pens.

PROBLEMS 3: EDITING

Read very carefully and be very critical.

Example.

Original Version: We have undertaken a reinforced protective reaction strike within the island of Grenada.[1]

Edited Version: We have invaded Grenada.

Now work on these:

1. Seeds were sewn in a prepared nursery, and transplanted three weeks later. Practices were performed according to conventional nursery production technologies.

2. The meticulous desing of the experiment was regarded as a necessary prerequisite to the conduct of a meaningful experiment.

3. Live weight gains for the five diferent feeding regimens were, respectively, 0.56, 0.55, 0.7, and 2.49, for the five feeding systems respectively.

4. Pesticide is commonly, at this point in time, used as a sort of generic terminology for any substances, chemical and otherwive, used for the control of pests, this term itself meaning any living organism that adversely affects man or his property.

5. Little onions are grown in Grenada.

6. Carrots were grated, on a scale of 0 to 5, and mean ratings for the four treatments were 0.75, 1.25, 4.50, and 2.00, respectively.

7. Yams were harvested and the roots weighed and separated into marketable and unmarketable tubers before weighing and counting they.

8. Pastures were fertilized and irrigated as required, and as considered necessary for sustained forage production. In practice this worked out at approximately 1.5 onches of water each week, and 125 lb of Nitrochalk each month.

[1] Attributed to a Pentagon Spokesman, late-October, 1983.

9. Stink bugs were counted weekly. The plants used for determining these counts were specially market plants. These were randomly selected, using random number tables. They were marked with colour-fast red plastic labels.

10. Grass weeds are either perrenial or non-perennial: the former are more difficult to control, but may not set seed, whereas non-perennial grasses can spread very rapidly by virtue of their seed production, and may be easier to control.

Now try to improve on this:

"This is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."

ANSWERS 3

There are no single correct answers. Those given below are examples of what could be considered correct answers.

1. Seedlings were raised in a nursery and transplanted at three weeks old. (The rest of the passage is redundant).

   (sewn should be sown)

2. Pure padding. Can probably omit totally and not bother to rewrite.

3. Live weight gains for the five feeding regimes (or syst- ems) were 0.56, 0.55, 0.70 and 2.49, respectively.

   The third value should read 0.70 for consistency. The 2.49 is suspicuously high, and way out of line with the other values: it should be queried. But what are the units?

   (diferent should be diffeerent)

4. Padded, with lots of uncertain and unnecessary phrases, including the abominable "at this point in time".

   Pesticide is now used to mean "any substances used to control pests", the latter term now being used to refer to "any living organism that adversely affects man or his property".

   Maybe the next version is better:

   A pest is now defined as "any living organism that adversely affects man or his property", and a pesticide is therefore "any substances used to control pests".

   (otherwive should be otherwise)

5. Few onions are grown in Grenada.

6. Mean ratings on a scale of 0 to 5, were 0.75, 1.25, 4.50 and 2.00 for the four treatments, respectively.

   The carrots were rated, not grated. Or were they rated for gratability? What property or characteristic was being rated? Was 0 good or bad? Presumably the four treatments are listed somewhere.

7. *Repetitive, and confusion of roots and tubers.*

   Yams were harvested and separated into marketable and unmarketable tubers, before weighing and counting.

   (them, not they)

8. **Pastures were irrigated with 1.5 inches water per week, and were fertilized with 125 lb per acre Nitrochalk each month.**

   Why the claim to some other basis for irrigating and fertilizing, if it came down to standard and regular applications?

   (inches, not onches)

9. Too much detail. We do not need to know what was used to mark the plants, nor how the random selection was done. More important would be how many plants per plot were used.

   Stink bugs were counted weekly on randomly selected marked plants.

   (marked, not market)

10. Confused. Difficult to see what the point is. Would need to refer back to the author. Maybe this is more or less what was intended:

    Perennial grass weeds are generally more difficult to control than annual grass weeds. The latter can sometimes spread very rapidly however, as a result of their heavy production of seed.

    (perennial, not perrenial)

APPENDIX 2: PARTICIPANTS

Lucius Alexander	Ministry of Agriculture
Kenny Daniels	Ministry of Agriculture
Nymphia Edward	Ministry of Agriculture
Joseph Esnard	WINBAN R & D Division
Cathalina Fontenelle	CARDI-St. Lucia
Francis Fredericks	Ministry of Agriculture
Ernest Henry	Ministry of Agriculture
Gillian James	Ministry of Agriculture
Henry Lubin	Ministry of Agriculture
Colin Paul	Ministry of Agriculture

APPENDIX 3: EVALUATION.

Nine participants completed the Evaluation Questionnaire.

1. What are your qualifications?

   M.Sc. or above......2
   First degree...........6
   Diploma...............1

2. How many years experience of research do you have?

   0 - 2....3 3 - 5....4 6 - 10....1

3. Was this Workshop useful to you in improving your skills and understanding?

   Yes.......7 No.......0 Not sure.......2

4. Was the level of presentation of the Subject Matter:

   Too simple....2 Too difficult....0 About right....5
   (No response....1 Too simple and about right....1)

5. Were the handouts:

   Too detailed....0 Too simple....0 Too long....0
   Too short.......0 About right...9

6. Which two handouts did you find most useful?

   Tree crops.........1 Data collection.............4
   Interpretation......5 Correlation etc.............2
   Sampling..........1 Management of experiments...1
   Objectives.........2

7. Was the amount of statistics:

   Too much....0 Too little....5 About right....4

8. Would you have liked more practical examples or exercises?

   Yes.....9 No.....0 Don't know.....0

9. Was the standard of presentation:

   Good....6 Poor....0 Fair....2 No answer....1

10. Which two topics do you think were particularly useful to you?

Design of experiments....4	Interpretation...........5
Data collection..........2	Management experiments...2
Livestock................1	Sampling.................1
Correlation etc..........3	

11. How could the Workshop have been improved?

Group assignments......1	Field visit..................2
Handouts given earlier.1	Practical computing sessions.1
More time and detail...1	More practical examples......2

More time on correlation etc. and data analysis.......1
Examples to be worked by participants.................1

12. What additional topics would you have liked to see in the programme?

Use of computers....1 | Choice of designs........1
Interpretation of computer printout..................1
Practical anova examples and interpretation.........1
(Expanded time for existing topics..................1

13. Would you like other Workshops on similar topics, or in other disciplines? Specify.

Data analysis............1 | Interpretation of data....1
Computer input of data...1 | Use of case studies.......2
More practical details...1 "Pure statistics"..........1
Entomology, nematology, plant pathology.........1
Other workshops on biometry.......................1

14. Did you ENJOY the Workshop?

Yes.....8 | No.....0 | Moderately.....1

15. Any other comments?

Need a 5-day workshop.....2 | Workshop too general....1
Need more specifics, and fewer topics.................1
Starting time difficult....................................1
Need follow up and contact.............................1
Extra day needed for participants problems............1
More work needed on interpretation...................1

FECHA DE DEVOLUCION

IICA
A50-269

Autor

Título Proceedings of a workshop on research methodology

Fecha Devolución	Nombre del solicitante