

# On-Farm Trial Guide

## Reference Book



How to run successful on-farm trials

---

P. J. Stone, A. J. Pearson & M. J. Bendall

LandWISE Inc. - [www.landwise.org.nz](http://www.landwise.org.nz)

*Sustainable cropping through technology*

***On-Farm Trial Guide***

*Reference Book*

*How to run successful on-farm trials*

**Peter J Stone, Andrea J Pearson, Maaike J Bendall**

Crop and Food Research

Hastings

PO Box 85

Hastings

New Zealand

**Bibliographical reference**

***Stone, P.J., A.J. Pearson and M.J. Bendall 2003: On-Farm Trial Guide Reference Book: How to run successful on-farm trials. LandWISE and Foundation for Arable Research, Napier, NZ***

ISBN 1-877174-41-6

Hawke's Bay Regional Council Plan Number 3264

© LandWISE Inc and Foundation for Arable Research 2003

**Acknowledgements**

Funding for this book was made through the Ministry of Agriculture and Forestry's Sustainable Farming Fund, the Foundation for Arable Research and LandWISE.

The support of Crop and Food Research and Hawke's Bay Regional Council is gratefully acknowledged.

## TABLE OF CONTENTS

3		
1	Getting your aims straight.....	6
1.1	Setting goals and objectives – what do you want to find out and why?.....	7
2	Getting the treatments right .....	8
2.1	What type of treatments should you apply? .....	8
2.2	How much of each treatment should you apply? .....	10
2.3	Do you have the tools required to apply the treatments as planned?.....	11
2.4	Do you have the time to apply the treatments as planned? .....	12
3	Getting the measurements sorted out.....	14
3.1	When should measurements be made? .....	14
3.2	What type of measurements should you make? .....	16
3.3	Do you have the tools to make the measurements as planned?.....	19
3.4	Do you have the time to make the measurements as planned? .....	20
4	Getting the design right .....	21
4.1	The first of the four Rs - Relevance .....	21
4.2	The second of the four Rs - Replication.....	22
4.3	The third of the four Rs - Randomisation .....	24
4.4	The fourth of the four Rs – aRrangement of the trial on the site .....	27
4.5	Is bigger better? – choosing plot sizes and shapes .....	30
4.6	Getting more bang for your buck – sharing trials .....	32
5	Getting the trial site right.....	35
5.1	Marking out the trial site .....	35
5.2	Measurements on the site .....	36
5.3	Making notes.....	37
5.4	Making sure.....	38
6	Getting the trial underway.....	39
6.1	Establishing the treatments .....	39
6.2	Monitoring and recording conditions at treatment establishment .....	40
7	Monitoring the trial – enacting your operational plan.....	41
7.1	Observing pre-planting treatment effects .....	41
7.2	Checking that treatments have eventuated as planned .....	42
7.3	Observing in-season treatment effects .....	43
8	Finishing the trial– getting the harvest done .....	45
8.1	Selecting the harvest area .....	45
8.2	Harvest equipment .....	46
8.3	Data recording.....	47
9	After the trial .....	49
10	Analysing the trial results.....	51
10.1	Re-check that data is complete and accurate .....	51
10.2	Entering data.....	53
10.3	Statistically analysing the data.....	54
10.4	Analysing the implications of the data.....	57
10.5	Recording results and conclusions .....	60
11	Avoiding common trial pitfalls .....	61
11.1	Preventing trials becoming a trial - common pitfalls in trial design .....	61
11.2	Applying a single treatment to a whole paddock... ..	62
11.3	Fertiliser trials without accompanying soil tests.....	63
11.4	‘Head-to-head’ trials that aren’t really head-to-head trials... ..	63
11.5	Treatment comparisons where one treatment isn’t managed properly.....	64
11.6	Common design problems.....	64

## FOREWORD

Our business is helping farmers with their business by providing practical information based on good science. We see a lot of on-farm trials being run by farmers and their various supporters as they test out ideas, compare options and determine whether a recommendation 'actually works on their place'.

Unfortunately we also see trials that for a range of reasons don't provide the returns they could. Sometimes people run out of time, or the question no longer seems relevant. But often, for lack of a little forward planning or guidance in correct procedures, really useful business information is lost.

We looked around and found lots of guides with titles about running trials on farms. We found they fitted pretty much into one of two groups: introductory and a bit light on detail (so not really useful), or crammed with detail and statistics that made them scary (so not really useful).

Peter Stone and colleagues from Crop & Food Research and the broader agricultural science community have enormous experience in running trials on research stations and on farms. It's what they do for a living.

Many of them are or have been farmers themselves. And they work with farmers every day, hearing the issues, listening to the questions, and providing practical guidelines based on strong science. They know farming is a business, and have the special ability to 'do science' that lets farmers make better management decisions.

This guide is designed to help farmers make business decisions based on trials done on their own farms. It is on-farm-practical, scientifically sound and business based.

This guide will also be extremely useful to all the other people – technical reps, students, advisors – who, for whatever reasons, run trials on farms. Like the farmers, they need a down to earth, practical guide but they need to make sure their testing is scientifically valid.

So regardless of your link to farming and to running trials, we commend this guide to you. Take a little time to work through it, have a look at the checklists for the different kinds of trial you might be contemplating. And don't forget to read the section, "Avoiding common trial pitfalls". As it says, you might recognise some old 'friends' that would best be avoided. . .

Dan Bloomer & Nick Pyke

June 2003

## PREFACE

*“Can you tell me what will happen on my farm if I...?”*

If we had a dollar for each time we've been asked this most important question in agriculture, we'd probably be living in the Bahamas, rather than working for a living. We suspect, too, that both readers and authors would be bound for sunnier climes if we actually had the answer to that question!

The fact that we have to work for a quid suggests that we can't tell you with certainty what will happen on your farm if you do or don't adopt a given practice.

As practising scientists, we know that even our best experiments can't answer your most important question – “what will happen on my farm if I?”. That's where this guide comes in.

We developed this guide with farmers in mind, because we know how important it is to test new ideas, and reassess old ones, if you want to run a successful farm business. When it comes to the practical business of farming, trials are just about the only way to test those ideas. And just about the only way to be certain that those ideas will work on your farm is to test them on your farm using a reliable on-farm trial.

We wrote this guide to help you run reliable trials on your farm, with the minimum amount of bother and expense. We recognise that you've got a business to run, and that running trials probably isn't at the top of your priority list. That may be because running good trials isn't easy – but we don't think that it has to be that hard, either.

We show that running an on-farm trial is a step by step process – we call it a “10 Point Plan”. This doesn't mean you do exactly the same ten things each time you run a trial. It means you have to consider the same ten issues each time. To run a really good trial you need to have a good answer for each step before you should move on to the next.

Most trials that are run on farms are one or other of half a dozen types. To make this guide as helpful as we can, we've prepared “10 Point Plans” for each of these. If there is a trial that you think is missing, please give us a call and we'll see what we can do.

We hope that you find this guide useful and interesting. Most of all, we hope that it helps you to improve your farm business.

If you are not a farmer, you have not been forgotten. Anyone running trials on farms has to go through the same checklists. If you are running trials on somebody else's farm, we are sure this guide will be just as useful to you.

We are grateful to our many friends and colleagues who helped with the production of this guide. Derek Wilson, Pete Jamieson and Bruce Searle provided the Crop Specific Tips for pea, wheat and squash, respectively. Brian Rogers, Isabelle Sorensen, Scott Shaw and Jeff Reid provided critical comments and feedback. We wish to record our thanks to the many growers with whom we've undertaken on-farm trials – it's been fun learning with you.

This edition of “How To Run Successful On-Farm Trials” was prepared and produced with funding from the MAF Sustainable Farming Fund, the Foundation for Arable Research and LandWISE. We greatly appreciate their support.

The authors: Peter Stone, Andrea Pearson, Maaiké Bendall, Crop & Food Research.

LandWISE Inc. - [www.landwise.org.nz](http://www.landwise.org.nz)

*Sustainable cropping through technology*

## **1 GETTING YOUR AIMS STRAIGHT**

*By the end of this section you will be able to formulate the aims of your trial. You should aim to be able to make a clear statement of what you want to know and why you want to know it.*

## 1.1 Setting goals and objectives – what do you want to find out and why?

***Know what you want to know, and why you want to know it. If you're not clear on these, don't start on your trial.***

You need to have a clear and simple statement of what you want to know (objective) and why you want to know it (goal). Goals are statements of the overall target that you want to reach. Objectives are the specific questions that need answering for you to achieve those goals.

Let's start with goals. A common goal would be to increase gross margin per hectare. There are obviously lots of ways that you could try to achieve this, but for the purposes of a single trial you need to focus on a single approach. Your goal might be to see whether increased plant population increases your gross margin per hectare.

Once you've got your goal clear, you can move onto your specific objective. In this case, a common objective would be to compare the gross margin from your current population with that from a higher population. Gross margin is made up of expenditure and receipts, so fulfilling the objective will require you to measure the amounts and values of inputs and outputs. Objectives should always be measurable.

You can see that setting a clear goal led naturally to a simple objective which, in turn, suggested the sorts of treatments and measurements that will need to be made for the trial to succeed. This gets you through the worst of the planning stage!

Setting goals and objectives isn't usually a difficult task, but it does require some care. Fuzzy goals usually lead to poorly defined objectives which, in turn, produce unfocused trials and inconclusive results – an all round waste of time and effort.

If you find that your objective doesn't relate adequately to your goal, have another go at it. By the same token, if your objective requires tricky treatments or more data collection than you can handle, modify it. The key is to be clear and simple.

If your goal and objectives can't fit into a simple sentence, you're probably biting off more than you can chew. Write them down and see if they pass muster.

If they're not quite right, have another go at it. You'll never regret spending time getting this part of the trial planning process just right. Mistakes made here almost always lead to trouble down the line.

Having established an achievable goal and objectives, it's time to think about the treatments that you'll apply and the measurements that you'll make on your trial.

### **Summary:**

**A simple statement of what you want to know and why you want to know it is the essential first step in running successful on-farm trials.**

### ***End of section critical decision point***

*Can you write a clear statement of what you want to know from trial, and why you want to know it? This is important if you are to get the best from the next section – Getting the treatments right.*

## 2 GETTING THE TREATMENTS RIGHT

*By the end of this section you'll be able to work out the details of two alternative treatments that you'll compare in your head-to-head trial. You should aim to have a set of instructions that outlines how many treatments there will be, how much of each treatment there will be, and how and when each treatment will be applied.*

**Good treatment application is as simple as following the 4 Ts – Two treatments, 20% difference, Tools for the job, Time to do it properly.**

### 2.1 What type of treatments should you apply?

*We strongly recommend simple head-to-head trials that compare something old with something new - 2 treatments is usually plenty*

*Armed with your simple statement of goals and objectives, selecting the right treatments shouldn't be difficult.*

Using the example from Section 1 "Getting your aims straight", we decided that we'd compare the gross margin from "normal" and "high" plant populations (objective) in order to see whether it increased gross margin (goal).

Clearly, this suggests that our treatments will be the "normal" population and a higher than normal population.

In most cases you will be comparing an existing practice or technique with a simple alternative, so you'll need only two treatments. Until you become really familiar with running experiments it's probably best to stick with this type of trial.

If you're comparing rates of herbicide application, a head-to-head comparison of label rate and either half or three-quarter rate might be a good start.

For comparing different types of herbicide, we'd recommend using the label rate for both.

If you're comparing two methods of herbicide application (eg. surface sprayed vs. sprayed and incorporated), use only one rate and type of herbicide.

Head-to-head comparisons of this sort are likely to be the types of trials that you perform most often - you know what you currently do, and want to see whether a simple alternative is any better.

The key to success in head-to-head trials is to change only one part of each treatment and nothing else

If your treatments are normal and high populations, then the only difference between your trial plots should be the population. You must sow each population on the same day, with the same equipment, using the same variety in the same paddock, using the same amount of fertiliser and herbicide.

If you've heard that higher populations need more fertiliser, or compete better with weeds, don't be tempted to give the higher population more fertiliser or less herbicide than the normal population. If you do, you won't be able to tell whether treatment differences were caused by population, fertiliser or herbicide. You won't have achieved your simple objective and, by generating a false impression of the effects of population, you could be worse off than if you'd not done a trial at all.

Similarly, if you're comparing broadcast and incorporated fertiliser with that applied down the spout, use the same rate and type of fertiliser and keep everything else the same. Don't be tempted to apply more fertiliser to the broadcast treatment to help the seedlings off to a good start. Once you add in "extras" you lose the ability to fulfil your objective.

Likewise, in comparing treated and untreated seed, don't "top up" the untreated area with spray to take care of extra pests.

In a head-to-head comparison you should take changes one step at a time

This doesn't always seem to fit best with a management change that will require more than one adjustment in your system, but it's still often the best way to go. Let's look at how this works, using the plant population example above. We'll then look at a complex example involving lots of changes – minimal tillage.

### **2.1.1 Example: A Plant Population Trial**

You know that higher populations may require more fertiliser and may require less herbicide than your normal population. Each of these factors affects the gross margin - which is why you want to do the trial – but you still shouldn't look at them all at once.

The best approach here is to first compare the populations using the same levels of management inputs. Apply the upper level of each input so that neither treatment has an unfair advantage. In this case, the rate of fertiliser required for the high population should be applied to both treatments. By the same token, the rate of herbicide required for the lower population should be applied to both treatments.

If the higher population doesn't show an advantage under these conditions, the effects of herbicide savings or fertiliser costs probably becomes irrelevant, and no further trials may be required.

If, on the other hand, the higher population does show benefits, the next step would be to see whether these can be improved upon by reducing the herbicide inputs or altering the fertiliser application.

**Only by taking a step at a time can you reliably work out where the costs or benefits are coming from.**

### **2.1.2 Example: A Minimum Tillage Trial**

Of course, there are going to be exceptions to every rule! Moving from conventional to minimum tillage, for instance, is not simply a matter of cultivating less. It's likely to involve a whole raft of changes in your production system. These could include altered sowing times, methods of dealing with pests, weeds and diseases, new approaches to residue management and possibly altered rates, timings and methods of fertiliser application.

It wouldn't be a good idea to take each of these changes a step at a time. Not only because it would take a lifetime to look at them all individually, but because looking at them in isolation ignores the fact that they interact with each other. So what's the solution?

**If you can't change a treatment in single steps, monitoring each ingredient of change is the next best thing to controlling them.**

You can't or shouldn't look at the effects of each of tillage, pest, weed, disease, and fertiliser treatment individually. So, because you know that they're likely to be important, you should try to quantify their impacts.

Careful, planned, monitoring with the aim of (at best) quantifying or (at least) describing the changes in inputs and outputs will help to show which parts of the system are contributing to benefits and costs.

So a reliable head-to-head comparison of conventional and minimal tillage might involve using 'best management practise' for each system. The trick is to try to work out which parts of the different systems are having meaningful effects on costs and receipts. You can only do this by measuring them.

The art of measurement is covered in Section 3. "Getting the measurements sorted out".

The statistical spreadsheet supplied with this guide will only handle simple A versus B comparisons such as those outlined above. Have a word with your friendly neighbourhood scientist if you'd like to do something with a bit more complexity.

**Summary:**

**Head-to-head comparisons of two treatments, involving only one change at a time, generally make the best sort of on-farm trials.**

**2.2 How much of each treatment should you apply?**

*Treatments should be different enough to have a readily identifiable practical or financial effect*

Given that you'll only be applying two treatments, you need to carefully consider the 'amount' or level of each treatment that's applied.

Your main aim will be to see whether treatments make any difference to yield or profit, so you'll want to choose levels that are likely to make a difference.

By the same token, the treatments need to differ by enough for you to be certain that the levels of application actually differed.

If you're comparing rates of herbicide application, for instance, it makes little sense to compare the label rate with 95% of the label rate – you'd be lucky to pick the difference, in either practical or economic terms. You'd probably also be lucky if you could actually apply a 5% difference in rate.

If you're comparing time of fertiliser application, a day or even a week between treatments is probably too similar, because both are small compared with the life of the crop.

Similarly, you're unlikely to pick the difference in yield between populations of 88,000 and 90,000 or even 95,000 plants/ha.

**As a rule of thumb, choose levels of treatment that differ by at least 20%**

It's not that treatments that differ by less than 20% won't make a difference. It's just that the difference will probably be difficult to pick up. This occurs because every experiment and every measurement has errors and variability associated with it. This variability is unavoidable – the trick is to know that it exists and design your treatments around it. Choosing treatments that differ by at least 20% is a rough-enough way to achieve this.

Scientists and statisticians go to great lengths to quantify these errors and variability. They often design experiments so that it's possible to work out whether the difference between treatments was caused by natural variability or the treatments themselves. The spreadsheet in the Data Analysis Toolboxes section does all of this for you, provided you follow the experimental design guidelines provided in the trial templates.

Of course, just because you might experiment with a minimum 20% difference between treatments doesn't mean that you need to implement that level of difference when you grow your commercial crops.

For example, if a population trial shows a yield and profit increase as you move from 88,000 to 110,000 plants/ha, you needn't switch your farm over to crops with 110,000 plants/ha. Having found a difference at 110,000 you can be pretty certain that 95,000 plants/ha will also yield more than 88,000 plants/ha – it would just be difficult to show it in a trial without lots of replication.

So there's a subtle difference between there being “no difference between treatments” and “not being able to *show* a difference between treatments”. By choosing appropriate treatment levels, you should be able to ensure that you can both generate differences and show that they exist.

#### **Summary:**

**To ensure relevant and successful head-to-head trials, treatments should generally differ by at least 20%.**

### **2.3 Do you have the tools required to apply the treatments as planned?**

*Make certain that you have the equipment to apply realistic treatments to your trial*

By this stage you should have a precise (but flexible) plan of the treatments that you'll apply and the levels that you'll use.

Take a couple of minutes to see how easy it will be to implement the plan.

If you plan to compare populations of 88,000 and 110,000 plants/ha, can you adjust your planter to actually achieve these populations? If you've never sown at the higher population, how can you be sure?

If you plan to compare spray rates of 600 and 800 L/ha, can you apply these rates accurately? Are you certain?

If you plan to compare fertiliser rates of 250 and 300 kg/ha, can you apply these rates accurately? How would you know?

It's not likely that you can rely on the 'notches' on adjustment levers, the pressures on gauges, or the numbers given in user's manuals when applying treatments. They are usually just a guide for ease of operation, and will rarely provide the level of accuracy required to apply treatments with any certainty.

By the same token, even high-tech devices that monitor spray rates, seeding rates and fertiliser output should be checked by calibration. The precision of a digital read-out often obscures the uncertainty behind the numbers that go into it!

Calibrating and checking the equipment that you're using is probably already part of your commercial crop management. Keep in mind that when you're experimenting, you're likely to be using new rates, techniques or methods of application. As a result, you will need to re-check that these new things are going to work according to your experimental plan.

**If you're going to the trouble of doing a trial, you might as well go to the trouble of ensuring that treatments are being applied exactly as planned. Check and calibrate all the necessary equipment in advance.**

Obviously, altering and calibrating equipment to apply different treatments can be time-consuming, and in the rush of a busy planting season best intentions can get left behind. Use the 10-point checklist to help calendarise your trial tasks.

**Summary:**

**Don't assume that you can apply treatments as planned. Do a test-run to be certain that you can.**

## **2.4 Do you have the time to apply the treatments as planned?**

*Good trials require quality time for both planning and execution. Set this time aside, because rushed trials usually give dodgy results.*

Calibrating your gear in a quieter time will make it easier to get the job done right.

You don't need to calibrate your equipment the day that you're going to use it, although you should try to ensure that conditions during calibration are similar to those occurring during operation.

This means that if an experiment is being established on cultivated ground, calibration should be done on cultivated ground. This is important because cultivation changes ground speed relative to engine speed, which changes the rate of application for many sorts of equipment.

In addition, the shaking caused by surface roughness causes big changes in performance of planter and fertiliser delivery equipment.

**Calibrate your gear under the conditions in which it will be operated**

Use the 10-point checklist provided in the trial Type Booklets to help calendarise your trial tasks.

**Summary:**

**Leave plenty of time to establish your treatments. Your trial is only as reliable as the treatments it compares.**

**End of section critical decision point**

*Can you put together a set of instructions that outlines how many treatments there will be, how much of each treatment there will be, and how and when each treatment will be applied? This is important if you are to get the best from the next section – “Getting the measurements sorted out”.*

### 3 GETTING THE MEASUREMENTS SORTED OUT

*By the end of this section you will be able to work out what type of measurements are required to get the most value from your trial, and when you should make them. You should aim to have a trial measurement plan in your diary.*

Having established goals and objectives, and the treatments required to accomplish them, the next step is to plan the measurements that will be required.

Your guiding principle should be “what do I need to know to fulfil my objective?”

This probably seems like an overly simple question, but to really get the most from your trials you'll probably want to think about measuring more than just the obvious variables (such as yield or cost) that relate directly to your treatments.

This is largely because cropping systems respond to factors such as climate, soil type and conditions, and management factors such as paddock history and timeliness of operation.

So there is a whole raft of factors that will influence the results of your trial, quite apart from the treatments that you apply. For this reason, and to enable you to interpret the applicability of your trial results to different circumstances (eg. paddocks, seasons or crop types), you'll need to know what you're dealing with. The only option here is measurement.

#### 3.1 When should measurements be made?

*Plan well in advance the type and time of measurements to be made – you can't turn the clock back once the season's underway.*

Measurements are probably going to be required at several stages of a trial, not just at the final harvest.

A simple but effective aid to planning measurements is to think of them as occurring at 4 times - before, during, at the end, and after the trial.

This can help to avoid missing crucial measurements and, if you attach estimated dates to each, will help to identify whether you're going to hit trouble at pinch times. It's a good idea to mark estimated measuring times in your diary so that you don't miss important ones.

Data sheets to help you plan and make measurements at each of the 4 stages of your trial can be found in the Data Analysis Toolboxes.

'Before' measurements can be done when you're selecting and establishing your trial site, and might typically include:

- General notes on the trial site (crop history for previous 5 years, any problems that have occurred)
- Soil fertility tests (especially if you're doing a fertiliser trial). This would usually include the major nutrients (nitrogen, phosphorus and potassium) and might include micronutrients if these can be an issue. Soil organic matter and pH levels can also help to interpret trial results. All these test figures are good to have as

part of your farm plan in any case. You can use them to see whether you're getting changes (good and bad) in your soil over time.

- Weed, pest or disease type and severity (especially if you're doing a spray trial)
- Soil physical conditions (soil tilth, depth of compacted layers, etc.), especially if you're doing a tillage trial.

'During' measurements can form part of your crop scouting programme, and might typically include:

- General observations of conditions at sowing and subsequent crop health, vigour and development
- Emergence and population counts
- Weed, pest or disease type and severity
- Time of key crop events such as jointing, tillering, flowering, silking, maturity, etc.
- Weather events (specific phenomena such as high winds, rainfall amounts, duration of surface ponding, etc).

**Never be tempted to discontinue measurements mid-season because you can already see a difference between treatments.**

Carry your trial through to the bitter end and make measurements at the end, no matter what you've seen during the season. You don't sell your crop halfway to harvest so don't make final trial measurements at this stage!

We commonly participate in trials where you'd be certain only a month or two into the season that one treatment was going to yield more than the other at final harvest. Despite this 'certainty' we often find no difference between treatments at the season's end, or even that early treatment 'differences' have been reversed.

The most common example of this occurs with fertiliser trials, particularly when comparing starter fertiliser with nil starter fertiliser treatments. It is very common for the starter fertiliser treatment to have an 'obvious' growth advantage early in the season. We've all seen it – bigger crop, greener leaves, more tillers. This need not translate into bigger yield at the season's end.

We've also seen rotation trials where some treatments get off to a wonderful start (due to good nitrogen supplies) only to fall over before the finish (due to a lack of water).

**The lesson is simple. If you want to know the effect of treatment on yield, measure yield. Good mid-season performance doesn't guarantee end of season success.**

'End' measurements are likely to occur around the time of final harvest and would typically include:

- Harvest date
- Yield
- Biomass
- Lodging

'After' measurements may or may not be required, and will vary with the type of trial that you do. They might include:

- Soil conditions
- Residue levels and ease of residue handling
- Quality analyses (eg. moisture, protein content, metabolisable energy, etc.)

Planning and sticking to your trial measurement plan is integral for the success of your on-farm trial.

#### **Summary:**

**Plan measurements that you need to make and write them in your diary. Stick to your measurement plan even if treatment differences appear to be so obvious that they don't need confirmation.**

### **3.2 What type of measurements should you make?**

*As a minimum, you need to make measurements that satisfy the aims of your trial. If your aim is to determine the effect of treatment on yield, then yield must be measured. If your aim is to determine the effect of treatment on disease incidence, then disease incidence must be measured. If you aim to determine the effect of treatment on soil 'health' then relevant indicators of soil health must be measured.*

The key point underlying the statement above is that visual observations are fine, but that they cannot under any circumstances replace the 'hard' numbers gathered by objective measurement.

'Hard' results are quantitative measurements that are intended to give you numbers for comparison and analysis. Typical examples include emergence counts, plant populations, grain and biomass yield, grain moisture, thousand kernel mass, bulk density, or the dates at which given events occur (eg. emergence, silking, anthesis or harvest maturity).

'Soft' results are qualitative observations that are intended to give you data for which numbers are inappropriate or difficult to obtain or interpret. Typical examples that you may use include soil tilth (blocky, coarse, medium or fine), harvestability (easy, moderately difficult, difficult), paddock history, presence/absence of pests and diseases and your general impressions about the trial.

Scientists often get a bit sniffy about the 'soft' results that you get from 'eyeballing' trials, but they can often provide essential support to the more 'hard' numbers that they're so fond of. A combination of hard and soft measurements is often the perfect approach for a trial, so it's really a case of 'horses for courses'. The trick is to know when to use a given approach...

Soft results are really important for 'setting the scene' of the trial. Records of trial location, general soil and seasonal conditions, prevalence and type of pests and diseases and your general feelings about 'crop health' and its relationship with 'normal' years are all really important for understanding how your particular trial is

likely to fit in the general scheme of things. A sheet for making these measurements is included in the “Data Analysis Toolboxes” section.

However, ‘soft’ methods of measurement are a recipe for disaster when used inappropriately.

**Eyes tells lies – objective measurements are hard to argue with.**

NEVER be tempted to assess yield or biomass using a visual assessment. It simply doesn’t work, because there is no reliable or verifiable relationship between a crop’s appearance and its performance.

For example:

- Crop height has no consistent relationship with yield.
- Crop colour has no consistent relationship with yield.
- Stem diameter has no consistent relationship with yield.
- Tiller or stem number has no consistent relationship with yield.
- Ear length has no consistent relationship with yield.
- Individual yield components (such as ear mass, ear number or kernel mass) are little use in isolation because when one goes up the others usually go down. In isolation, they tell you very little about your crop.

In our experience, it’s better not to do a trial at all than it is to rely on a visual assessment of a key variable such as yield or biomass.

**If you need to know the yield or biomass of a crop, measure it objectively and directly, using quantifiable methods.**

Recipes for these methods are provided for a range of crops in the “Crop Specific Tips” section.

When you aren’t able to control factors that will influence trial outcomes, try to measure them.

As we mentioned in the section on “Getting the Treatments Right”, it’s not always possible to run a simple head-to-head trial where only one variable is changed at a time. The example we used was a comparison of conventional and minimal tillage methods.

In that case, we saw that the trial would involve lots of changes applied at once and that this would make it difficult to follow our ‘one step at a time’ rule.

You can’t look at the effects of each of tillage, pest, weed, disease, and fertiliser treatment individually. But, because you know that they’re likely to be important, you should try to quantify their impacts. Careful, planned, monitoring with the aim of quantifying the changes in inputs and outputs will help to show which parts of the system are contributing to benefits and costs.

For example, if you find that you get a lower yield with minimal tillage and have little by way of supporting measurements, you might be inclined to give up on the technique altogether, even though it improved soil health. This could be throwing out the baby with the bath water, especially if...

...on the same trial you've made measurements of plant emergence and population. These measurements may show that poor emergence is causing all of the yield problems and that fixing them will increase yield and give you better soil health.

In other words, measurements enable you to break your trial down into a series of mini-trials. Each mini-trial focuses on a small part of your cropping system and can be used to optimise each step – a sure-fire way to increase performance.

So a reliable head-to-head comparison of something complex like conventional and minimal tillage might involve using 'best management practise' for each system along with a bunch of 'extra' measurements.

The 'extra' measurements might include:

- Soil conditions. What is the soil type and physical condition? What is the soil fertility? What aspects of these are likely to impact on the relative suitability of conventional and minimal tillage? How much of which fertilisers should you apply?
- Seasonal conditions. Was it a dry or wet year? At which stages of crop growth? How might this have affected the results?
- Tillage. What is the time/cost of cultivating? How might this vary with soil/seasonal conditions?
- Sowing. Did it go OK? Was there any effect of tillage method on seed placement (depth, distance between seeds)?
- Crop emergence. Did time of 50% emergence change? Was emergence even over time? Was emergence variable within the row? What was the effect on the emerged population? If it was different, was it because of soil temperature, wetness or physical restrictions (surface capping, etc.)? How many seeds germinated but did not emerge?
- Pest, weed and disease incidence. Did these differ and did they require different control measures? What did these cost? How much time did they take? How much would these vary with soil or season?
- Crop development. Did jointing/flowering/silking/maturity occur at the same time?
- Crop biomass or yield. Did these differ? By how much?
- Trash/residue handling. How much residue was there? How much time did dealing with it take? How much did it cost? What is it likely to contribute to future crops?
- Soil conditions. What was the effect of each system on soil physical conditions and fertility at the end of the season or the start of the next? What would this be worth in dollars, time or flexibility?

Because you can't isolate the effects of each of these factors experimentally, it's important that they are monitored or, ideally, measured so that you can see how each affected costs and receipts. These may be measured in terms of timeliness, management time, inputs required, outputs received and effects on your system's future value.

If you can't measure them all, then pick the ones that are most important to you and concentrate on them.

Only by measuring the different parts of the system will you be able to identify which bits work and which don't. This is the key to repeating your successes and avoiding repetition of failures – surely the reason you did the trial in the first place?

**Summary:**

**There's no substitute for good hard numbers. One direct measurement of yield or quality is worth more than almost any number of visual observations.**

### **3.3 Do you have the tools to make the measurements as planned?**

*Make certain that you have the knowledge and equipment required to undertake the measurements that you've planned. If you need help or equipment, organise it in advance.*

By this stage you should have a precise (but flexible) plan of the measurements that you'll make

Take a couple of minutes to see how easy it will be to implement the plan.

Specific information about methods and equipment required for crop measurements can be found in the "Crop Specific Tips" section.

If you plan to measure soil physical conditions, have you decided on the best indicators to use? Measuring the depth of compacted layers is pretty simple, but how about severity of compaction? Will a bucket and spade do the job or is specialist equipment such as a penetrometer required?

If you plan to measure soil chemical properties, have you decided which nutrients are most important? What depth will you measure to? Do you have a soil sampler?

If seasonal conditions are important, do these need to be measured on site, or will data from the local newspaper do the trick?

Pest, weed and disease incidence can have a big influence on crop performance and trial results. Do you recognise the important species of each?

Will you use a weigh wagon to measure yield or will hand harvesting be required? Is a weigh wagon going to be available? If you're hand harvesting, do you have access to a drying oven and a reliable set of scales?

Do you have facilities (refrigerator, oven, etc.) for safely storing samples that need to be analysed at a later date?

If you're using specialised equipment make sure that it's been calibrated properly. This is really important for indirect measurement devices, such as grain moisture meters and yield monitors. Because these don't measure grain moisture or yield directly, it's important that the calibration factors are up to date and correct.

Be especially wary of equipment that comes with a digital read-out. The precision of a digital read-out often obscures the uncertainty behind the numbers that go into it!

If you're going to the trouble of doing a trial, you might as well go to the trouble of ensuring that measurements can be made exactly as planned. Check the availability and calibration of the necessary equipment in advance.

**Summary:**

**Think about the measurements that you want to make and ensure that you have access to the right equipment at the right time.**

**3.4 Do you have the time to make the measurements as planned?**

*Reliable measurements require quality time for both planning and execution. Set this time aside, because good measurements are the foundation of good trials*

Planning the measurements required and calibrating the necessary gear in a quieter time will make it easier to get the job done right.

You don't need to calibrate your equipment the day that you're going to use it, although you should try to ensure that conditions during calibration are similar to those occurring during operation.

Use the 10-point trial templates to help calendarise your trial measurements.

**Summary:**

**Your trial is only as reliable as the measurements that you make on it. Give them the time and care that they deserve.**

***End of section critical decision point***

*Can you put a trial measurement plan into your diary, outlining what measurements need to be made, and the equipment needed to make them effectively? This is important if you are to get the best from the next section – Getting the design right.*

## 4 GETTING THE DESIGN RIGHT

*By the end of this section you'll know most of the tricks of trial design used by the pros. You'll be able to replicate, randomise and arrange your trial on the paddock so that you get water-tight results. You should aim to have in your hand a map of your trial design, ready to be pegged out in the field.*

Good trial design is as simple as following the 4 Rs – Relevance, Replication, Randomisation and aRrangement. Once you've got these right it's time to think about optimising plot size and shape. This section covers the technical detail on these factors and then looks at how you can get more bang for your buck by sharing trials with friends and neighbours.

### 4.1 The first of the four Rs - Relevance

**A good trial site has low variability, is similar to the rest of your farm, and is easy to get to.**

Finding this combination of features requires some planning, so select your trial site well before the rush of the cropping season begins. This will enable you to avoid some common pitfalls...

More often than not, on-farm trials are located in funny little paddocks that aren't good enough for 'real' crops. If a site isn't good enough to farm profitably, why use it to make decisions about farm profit?

Ideally your trial site would be representative of your property as a whole. That way the results will be most relevant to the majority of your business.

**Consider giving your trial site priority over your commercial crops. A commercial crop pays you only once, whereas the results from a reliable trial could pay dividends for years.**

Early site selection will give you time to be certain that you've picked the best place for your trial. It'll give some lead time for soil testing - an essential ingredient of fertiliser and many other trial types - and gives you time for 'second thoughts'

Failure to recognise and minimise variability in trial paddocks is a common problem. Variable sites can ruin trials by making it impossible to distinguish effects of treatment from effects of variability caused by other factors. Have a look at your commercial crops and avoid putting trials in areas that have had variable establishment, growth or yield (unless you're wanting to investigate the sources of this variability).

You can minimise variability by avoiding corners of paddocks, which often receive multiple applications of fertiliser or herbicide and are less likely to represent the property than other parts of a paddock.

Stock camping areas, gateways, water troughs, headlands, areas where water gathers, steep slopes, trees, weedy patches, old fence lines or tracks, sheds or other irregularities are best given a wide berth. Hopefully, they're not like most of your property, so you're unlikely to get relevant trial results near them.

You want easy access so that it's not a big effort to establish and monitor the trial. Somewhere that you pass on a regular basis would be ideal, provided it fits other site selection criteria.

You may need to get equipment in at unusual times and so it might be necessary to have a site that can be reached without destroying other crops.

In addition, select a site that best matches the purpose of your trial, by paying particular attention to things that may influence your treatments...

Fertiliser trials should be conducted where soil has known and relatively uniform nutrient levels. The lower the soil test for a given nutrient the greater the response to fertiliser containing that nutrient, so don't expect measurable responses in highly fertile paddocks. Other factors affecting the response to fertiliser, such as drainage, organic matter levels and soil depth should be considered in terms of variability and similarity to the farm as a whole.

Herbicide trials are best placed where there is a weed problem of known type and severity. If the herbicide is to be soil-applied, factors that will affect response should be considered. Measures should be taken to minimise variability and maximise relevance to the rest of your farm. Factors affecting response to soil applied herbicide include soil tilth and texture, organic matter content and pH.

Tillage trials are tricky because the results are often greatly affected by the timeliness of operation, but soil factors such as soil type, texture, depth and crop and tillage history should all be taken into account when selecting the site.

**You've all seen yield maps and been surprised by the huge variability in what seemed to be uniform paddocks. That variability is likely to ruin your on-farm trial if the '4 Rs' of good trial design aren't applied – Relevance, Replication, Randomisation and aRrangement.**

#### **Summary:**

**You need to select a site that represents your farming system and is relevant to the type of trial that you're running.**

### **4.2 The second of the four Rs - Replication**

**The need for replication in trials is similar to the need to play more than one rugby match to see who's the best team.**

You wouldn't bet your house on the All Blacks to win the World Cup after seeing them against one opponent on one ground. It's the same with trials – you need to see more than one comparison before you can make a sensible bet on the results.

To extend the rugby analogy... The outcome of any given match isn't certain partly because combinations of opponents and grounds affect the outcome. Some teams play better against a particular style of play and some teams have a marked home ground advantage.

In trials, replication aims to increase the certainty of betting. By having more than one head-to-head comparison, it enables each treatment to play each opponent more

than once. In addition, by having treatments repeated at different times or parts of the paddock, it includes at least one 'home' game and one 'away' game.

Replication refers simply to the number of times that you repeat a treatment using separate plots. Two replicates is where you have two separate plots for each treatment. Three replicates is three separate plots for each treatment, and so on.

Four replicates usually provides the best balance between precision (the ability to separate effects of treatment from other effects, such as site variability) and the amount of work involved. Three replicates might give 25% less work, but it gives about 35% less precision. Five replicates, on the other hand gives 25% more work but only about 15% more precision than four replicates. So four replicates is a good compromise.

**Trials without adequate replication are highly unreliable and could harm your business. Four replicates provides a good balance between effort and reliability.**

You can achieve replication using space or time, or a combination of the two. Each method works by giving head-to-head comparisons under slightly different conditions.

An example of replication over time is where two treatments are compared head-to-head in the same paddock for four years.

An example of replication over space is where two treatments are compared head-to-head in four different paddocks in one year.

The best method of replication depends partly on the type of business that you run. If your yields (per hectare) in a given year are pretty similar amongst paddocks then replicating over time may be the best option. If your paddocks vary widely in yield (indicating different soil conditions) then replicating over space may be the best bet. Combining the two is often even better because it enables you to compare treatments under the widest possible range of conditions.

If you have one or several fairly uniform blocks of land, then replication over time will probably be the best option. This is because on relatively uniform sites your biggest source of variation will be that which occurs due to season. By replicating over time you'll find out how well two treatments compare 'on average, over time'. This is a sound basis for decision making where soil and other conditions are fairly consistent.

With a standard two treatment (A & B) trial with four replicates (giving a total of eight plots [2 treatments x 4 replicates]) you could replicate in time by having two plots in the same paddock for four years. Alternatively, you could have four plots in one year and four plots in the next. Another option would be two in the first year and six in the next. It doesn't really matter how you arrange it, as long as you have head-to-head comparisons in each year.

If you run a variable block, or lease different types of land, it's possible that replicating over space may be the best bet. That way you'll get to compare treatments under your different conditions and will be able to see how they perform 'on average, in different locations'. This is a sound basis for decision making where you may want to apply a change across your whole operation.

With a standard two treatment (A & B) trial with four replicates (giving a total of eight plots [2 treatments x 4 replicates]) you could replicate in space by having all eight plots in one paddock. Alternatively, you could have two plots in each of four paddocks. Another option would be two plots in one paddock and six in another. By the same token, you could have two plots in each of four different properties. There's practically no limit to the combinations that you can use to get your replication, as long as you maintain simple head-to-head comparisons.

Also keep in mind that you want to be able to apply the results to certain types of conditions, so it's best to put your replicates where those conditions are most likely to exist.

Of course, just because you get variation over space doesn't mean that you won't also get variation over time. For this reason, if your trial results have far-reaching consequences, it's often a good idea to have a look at running a trial more than one season, even if you've replicated using several different places.

There's no reason why you can't run trials with your friends or neighbours and treat their plots as replicates of the same trial. That way you can pool resources, cut down on the work involved for each of you and extend the relevance of your findings across a wider range of conditions. Some pointers for running multi-farm trials can be found in Section 4.6 Getting more bang for your trial buck. Of course, if trial partners have conditions that are too different they may be reducing the relevance of the trial for both parties.

If we've convinced you of the need for replication, it's now time to look at some specific examples of how to replicate trials in space or time. But first we need to have a few words about randomisation....

#### **Summary:**

**You need to use four replicates of each treatment if you are to get a reliable comparison between them.**

### **4.3 The third of the four Rs - Randomisation**

**Randomisation serves two basic functions. First, it enables you to repeat a head-to-head comparison under a variety of slightly different conditions. Second, it ensures that no treatment has an unfair advantage by always receiving the best spot(s) in the paddock.**

Playing the trial game without randomised plots is like playing cards with a stacked deck – it's designed to be unfair and you're likely to do your dough!

It's easiest to show how important randomisation is by using our standard two treatment (A & B) trial with four replicates (giving a total of eight plots [2 treatments x 4 replicates]) as an example.

If you were to compare treatment A and treatment B without randomising you'd end up with the layout shown in Fig. 1 below.

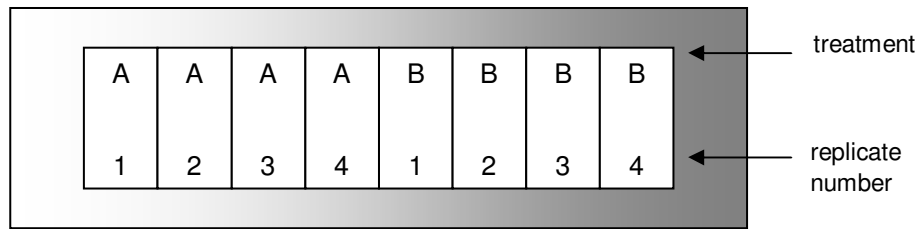


Fig. 1. An example of poor trial design - poor replication and no randomisation

There are a few problems here. First, there's not really any replication – you could argue that there's actually just one big plot for each treatment rather than four little plots. Secondly, if one end of the paddock is better than the other, one treatment is going to have an unfair advantage. This is demonstrated in Fig. 2, below.

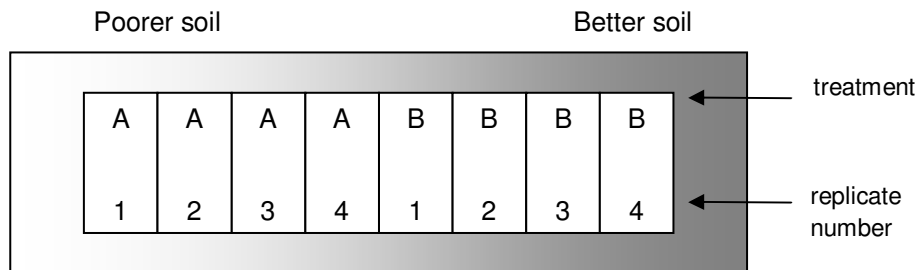


Fig. 2. An example of poor trial design - poor replication, no randomisation and built-in bias

Because treatment A has been assigned to the poorer soil and treatment B to the better soil, it's highly likely that treatment B will come out on top, even if it's not the better treatment. Obviously, a design like this can't give reliable results. An improved approach is shown in Fig. 3, below.

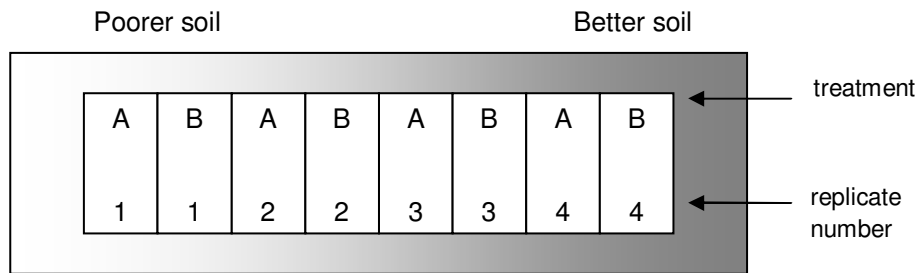


Fig. 3. An example of better (but still poor) trial design – adequate replication, but lack of randomisation in blocks does not minimise chance of bias

You'll notice that we've re-arranged the replicate plots so that treatments are paired (replicate 1 of treatment A next to replicate 1 of treatment B, 2 next to 2, 3 next to 3, and 4 next to 4). This is called blocking, because each combination of treatments occurs in a 'block' that is part of a treatment 'path'. There are 4 blocks in this treatment path.

We'll return to the image of blocks in a treatment path later, because it helps to visualise simple and effective trial designs. For the moment, you'll see that by re-arranging the replicates into blocks we instantly achieved two major improvements. First, we got direct head-to-head comparisons of each treatment. Second, we gave each treatment a better chance of being assigned to an area of good or poor soil.

**Effective randomisation ensures that there's a valid head-to-head comparison within each replicate.**

Even though we've made 2 big improvements, there's still a problem with this trial design. The lack of randomisation in Fig. 3 means that in each of the 4 head-to-head comparisons, treatment B always gets the better end of the paddock. So even though we've improved the bias situation we've not reduced it to a minimum. The next step to establishing the optimum trial design is achieved by randomising treatments within blocks, as shown in Fig. 4, below.

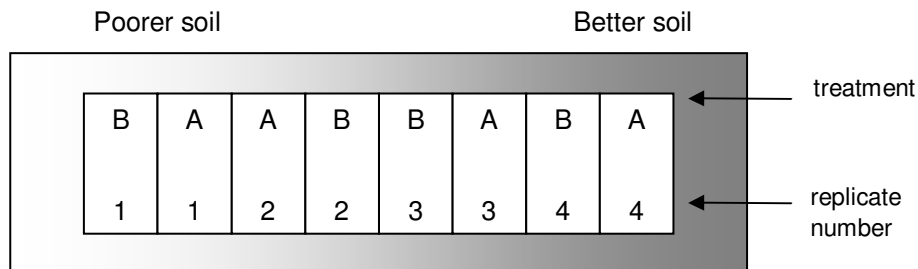


Fig. 4. An example of very good (but not perfect) trial design – adequate replication, randomisation in blocks reduces but does not eliminate influence of bias

The trial design in Fig. 4 improves on that shown in Fig. 3 by having randomised the treatments within each block. You'll notice that instead of treatments reading a regular A-B, A-B, A-B & A-B as you move from left to right, they are now randomly assorted: B-A, A-B, B-A & B-A.

You can randomise treatments within blocks using the Excel worksheet in the "Data Analysis Toolboxes" section.

This random assortment has given each treatment within each head-to-head block an equal chance of being assigned to a good patch of soil. As a result, we've eliminated most, but not all of the bias in this experiment.

You can see that the thoughtful use of replication in blocks and randomisation of treatments within blocks has vastly improved the pretty lousy trial design that we started with in Fig. 1.

The improvements that we've made didn't involve much work. All we had to do was (1) think a little about the characteristics of the trial site, (2) put blocks along the trial 'path' and (3) randomise the treatments within the blocks.

This little bit of 'paper' effort has paid big dividends. Previously, the trial was effectively 'rigged' to show that treatment B was best. Now, we're in a position to more reliably assess whether treatment A or B is best.

Application of the basic principles of replication and randomisation has given us a near-perfect trial design. A little more thought about how the trial could be arranged in the paddock will result in a trial that would withstand scrutiny from an international panel of experts! It's not hard, so read how in the next section.

**Applying the basic principles of replication and randomisation is a very quick, very cheap and very easy way of converting a trial from a meaningless mess into a jewel of reliability. It's 10 minutes of planning that could save you hours of effort and make you pots of money**

#### **Summary:**

**Randomisation is needed to ensure that each treatment gets a fair chance of winning.**

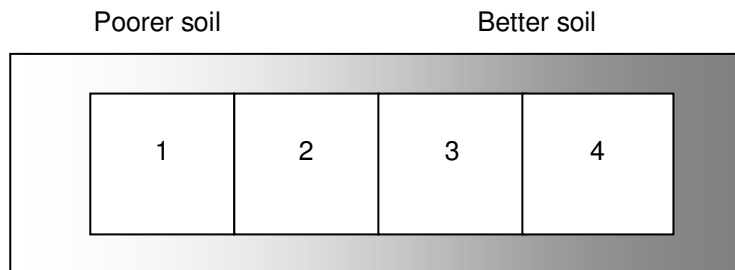
#### **4.4 The fourth of the four Rs – aRrangement of the trial on the site**

**The arrangement of treatments, plots and blocks on the site is as important as the site itself**

The clever use of replication and randomisation has massively reduced the bias that would have occurred had we used the design outlined in Fig. 1. By the time we'd reached the latest design (Fig. 4), we'd blocked the treatments to provide 4 head-to-head comparisons instead of just 1, and we'd randomised the treatments within the blocks to reduce the unfair advantage that consistent allocation to better soil had given to treatment B.

But our trial design (Fig. 4) isn't perfect yet. The arrangement of the treatment plots within each block still means that one treatment in each pair is assigned (by chance) to better soil conditions than the other. Even though random favouritism of this kind is better than the inherent bias shown in Fig. 2, it's still not ideal.

A simple trick for banishing this 'random' bias is to re-align the plots to better match the trial site. As a first step, we'll return to the image of the trial as a path. When variation can't be avoided, blocks should be arranged so that they form a trial path that progresses from block 1 to 4 along the variation in the paddock rather than across it. By this means, each block is most likely to contain a uniform set of properties. Figures 5 and 6 below show how this works.

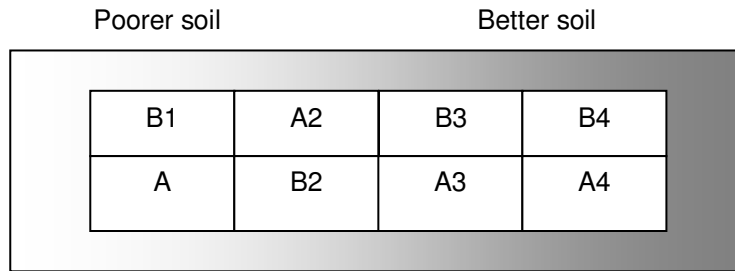


*Fig. 5. An example of good blocking – the blocks are arranged to make a path that leads along the site variation, from poorer to better soil*

The blocking used in Fig. 5 is correct because it makes a path that goes along the variation of the site, from poorer to better soil. Block 1 contains mostly poorer soil, block 2 is mostly medium-poor, block 3 is mostly medium-good and block 4 is mainly good soil. This arrangement allows each head-to-head comparison (within a block) to occur in reasonably uniform conditions.

**Arrange blocks to form a path that runs along the trends in the trial site. This will increase the head-to-head value of comparisons within each block.**

This is the same form of blocking that we used in our latest trial design (Fig. 4). But, as we mentioned earlier, that design wasn't perfect because it didn't eliminate all of the bias in soil type. The last problem in our trial design can be eliminated by simply rotating the plots within each block. When this is done, the plots run across the trend in the paddock (at right angles to the blocks). This is shown in Fig. 6, below.



*Fig. 6. An example of very good trial design – adequate replication, blocks that run along the trends in the paddock, plots that run across the trends in the paddock, and full randomisation of treatments within blocks*

Our final trial design is an improvement on the Fig. 4 model. Each treatment within a head-to-head block not only has an equal chance of being assigned to poor or good soil, but each treatment within a block actually receives the same soil type. We can now be fully confident that any bias in our trial design has been minimised.

**Arrange plots to form a path that runs across the trends in the trial site. This maximises the head-to-head value of comparisons within blocks.**

Changes in soil type or quality aren't the only things that will introduce trends or random variability in a trial site. The principles outlined above should also be applied to trials established on slopes, near shelter belts or any factor that could result in a trend in the trial site. Of course, if you can, it's always lovely to pick a site without any obvious trends.

**Even when you can't identify trends in the trial site, it's essential that you assume that the site is variable. This means that you must block the trial and randomise treatments within each block.**

Plot orientation within each block becomes less crucial when a site doesn't have any obvious trends in it. When this happens you can orient plots within each block so that the trial can be established, monitored and harvested most simply.

**Summary:**

**Arrange treatments to minimise the effects of site variability. Place blocks along the variation and then arrange plots across the variation.**

#### 4.5 Is bigger better? – choosing plot sizes and shapes

**Most growers (and many scientists) believe that, for on-farm trials, bigger plots are better than smaller plots. This is simply not true.**

Statistically speaking, there is no reason to believe that big plots give better answers than small plots. In fact, the reverse is often the case.

Bigger plots are more likely to contain more variation in soil and other conditions. Unless you can readily define what that variability is, bigger plots may just be creating more uncertainty about the results of the trial. Smaller plots, simply because they occur in a smaller area, tend to be less variable within themselves and, if blocked correctly, create less variability within a block. In addition, being smaller, it's easier to characterise the variability within or between plots, should any exist.

We often see on-farm trials with plots that extend the full length of a paddock. This can incorporate some of the best elements of design, but it can lead to problems if it's not done carefully.

Long narrow plots are usually better than short wide plots, because they enable head-to-head treatments to be close to each other. This can minimise some of the influence of site variability.

However, as we've mentioned earlier (in the "Relevance" section), most paddocks aren't very uniform, and one end is often quite different from the other. In addition, the ends are usually quite different from the middle section. This all adds up to the sort of variability that you want to avoid. How can you most easily avoid the problems, whilst maintaining the convenience of big long strips?

Just because it's convenient to sow plots that are full paddock length doesn't mean that you need to make measurements on the full paddock length.

There's no doubt that it's much easier to sow a strip the full length of a paddock than to stop and start all over the place. It's also easier to find a full paddock length strip than a plot buried in the middle of a heavy crop. The simple compromise between simplicity and reliability is to sow as long a plot as you find convenient, but to make measurements on only those parts of the plots where there aren't variability problems.

**Establish huge plots if you feel like it, but avoid making measurements in areas of the plot that are not typical of the paddock**

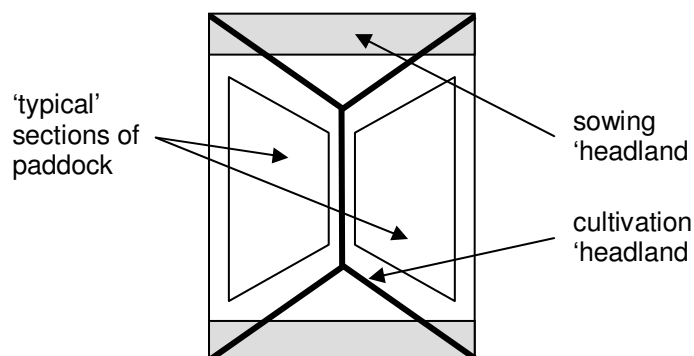
We've outlined common sources of variability in the section devoted to maximising the relevance of your trial. Additional problems related to the ends of paddocks include:

- Gateways mean traffic, which can introduce compaction and different weed and pest populations
- Troughs, stock feeding and camping areas
- Drains provide different drainage characteristics, either because of the drains themselves or the subsoil dumped to the side of them

- Headlands often have different amounts of herbicide, pesticide, fertiliser, seed, cultivation and traffic on them

Headlands are a common problem because, depending on the patterns of cultivation and sowing, they can cover irregular parts of the paddock. A lot of paddocks are cultivated using ever-decreasing circles, giving the characteristic envelope pattern shown in Fig. 7.

When this is added to by one of the common headland sowing patterns (also shown in Fig. 7), the area 'typical' of the paddock as a whole is broken up into two big and separate chunks. To ensure that your trial results are representative of the paddock, it's a good idea to confine your measurements to areas found within 'typical' chunks of land. The precise location of these areas will depend on the methods of cultivation and sowing that have been used, as well as the general characteristics of the paddock.



*Fig. 7. Headland pattern resulting from cultivation using ever-decreasing circles, and 'across the top and bottom' sowing of headlands. Typical sections of paddock are the best spots for making measurements.*

Clearly, field size is an important influence on plot size – you've got to be able to fit all of the required plots in! Don't be tempted to cut down on the number of replicates because the plots are taking up too much space. When this happens, make the plots smaller, don't cut back the number of replicates.

**Replication is much more important than plot size in determining the reliability of trial results, so more replication of smaller plots is far better than less replication of larger plots.**

Having gone through the philosophical niceties involved in selecting plot sizes, we mustn't lose sight of the fact that the equipment that you have access to often determines the plot dimensions that you can most easily use.

Ideally, plots should be wider than the equipment used to harvest them. This reduces the influence of errors in treatment application and harvesting, and edge effects.

Commercial equipment is usually designed to provide a good 'on-average' application of seed, spray or fertiliser. Their precision in small areas isn't generally very good so you'll want to use plots big enough to enable the small area errors to be averaged out over a bigger area.

By the same token, commercial harvesting equipment loses yield between the paddock and the field bin. Some of this loss is fixed, which means that the error arising from the loss decreases as the mass of harvested material increases. From this point of view, plots should be large enough to reduce measurement errors. Note that it is not usually the plot size that matters when we're talking about harvesting errors – it's crop mass. The higher yields of NZ crops compared with those of USA or Australia means that we can use much smaller plot sizes than you might have heard from these places.

*Details on harvest sizes required for reliable measurement are outlined in the Crop Specific Tips section.*

Some equipment - such as fertiliser spreaders or sprayers – give naturally diffuse edges, where the sides are markedly different from the centre of the strips. Where this occurs the application of treatments should be wide enough to enable the measurements to be taken in the more precise plot centres. So, when planter widths are the same as harvest widths, it may be an idea to consider making plots two planter widths across.

**Harvest widths should be smaller than plot widths, to avoid or minimise edge effects.**

If you're harvesting a trial by hand, you can easily use the relatively small plots often used by scientists (about 2 x 10 m for cereals or 5 x 10 m for row crops). Most commercial planters in use today give reliable populations within about 3 m of operation, so plots don't need to be more than about 10-15 m long for there to be enough material for a reliable harvest.

**Summary:**

**Plot size, in itself, isn't usually that important. You end up with the ideal plot size after you've ensured that (in order of importance): 1) you have 4 replicate blocks of each treatment, 2) the area within each block is uniform, 3) you can minimise edge effects and, 4) you can minimise measurement errors.**

#### **4.6 Getting more bang for your buck – sharing trials with friends and neighbours**

**Sharing trials with friends and neighbours can achieve the twin advantages of cutting the cost of running trials and increasing the range of conditions under which treatments are compared.**

Achieving this without compromising the credibility of the trial isn't hard – you just need to apply the methods of good trial design outlined above.

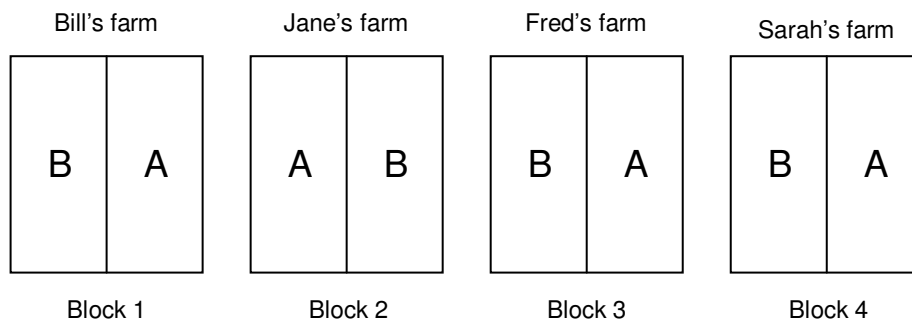
The main difference between single-farm and multi-farm trials is the arrangement of blocks. Just about everything else is the same.

You still use the same number replicates in a multi-farm trial as you would in a single farm trial. Instead of having all 4 replicates in one paddock, you may have one replicate in each of 4 farms. Alternatively, you could have 2 replicates in each of 2 farms or any other combination that gives you 4 replicates of each treatment (subject to the condition that they are head-to-head comparisons of treatment).

You should still randomise the treatments within each trial, just as you would for a trial occurring in a single paddock.

Each different farm in a multi-farm trial is considered to be a block. This is because each farm contains a head-to-head treatment comparison that occurs under conditions that are similar within each farm. On each farm, the blocks and plots should be arranged to minimise the effects of site variability and bias, just as if the trial was being conducted on a single farm. Follow the guidelines given elsewhere in the Guide for Relevance, Replication, Randomisation and aRrangement.

A typical multi-farm (4-farm) trial design is shown in Fig. 8 below:



*Fig. 8. An example of good design in a multi-farm trial. Each farm is a block that contains a head-to-head with treatments randomised to reduce effects of variability and bias.*

The results from a multi-farm trial like this can be analysed in exactly the same way as for the same trial conducted in a single paddock. The spreadsheet in the “Data Analysis Toolboxes” will spit out the right answers for you.

Of course, there's probably little point in sharing a trial with friends or neighbours if their conditions are completely different from yours. It might save time and money in conducting the trial, but how relevant will the results be to you (or them)? The ideal situation occurs when you have access to a range of sites that differ enough to make a difference but not so much that your results aren't broadly relevant to all the trial participants.

#### **Summary:**

**Sharing trials with friends and neighbours can be a good way to get more bang for your trial buck. As long as you treat each farm as a block and follow the usual design guidelines, it's a great way to pool resources and experiences.**

**End of section critical decision point**

*Do you have a map of your trial design in your hand? Does it incorporate the best elements of replication (4 replicates), randomisation (treatments fully randomised within each block) and arrangement (blocks along the site variability and plots across the site variability)? Is the site where you want to enact this plan representative of your farm operation? Is it relevant to the treatments being applied? It is important that you can satisfy these questions if you are to get the best from the next section – Getting the trial site right...*

## 5 GETTING THE TRIAL SITE RIGHT

*By the end of this section you will be able to characterise the relevant starting conditions of your trial site. You will be able to use that knowledge to assess whether your trial plan and trial site are well matched. You should aim to have a marked-out trial of known dimensions and a plot-by-plot knowledge of relevant conditions and potential problem areas.*

Getting the trial site right is as simple as following the 4 Ms – Marking, Measuring, Making notes and Making sure. Before embarking on this task, make certain that you've satisfied all previous 'end of section critical decision points'. There's not much point in getting the perfect site for an imperfect trial design.

### 5.1 Marking out the trial site

Enter the trial site with the plan of your trial layout in one hand and a tape measure in the other. Measure the maximum dimensions of your trial to see whether it will all fit into the space selected. Assess whether there's room to comfortably establish, maintain and measure the trial.

Remember that replication is much more important than plot size in determining the reliability of trial results, as outlined in Section 4.5 *Choosing plot shapes and sizes*

*Can you put together a set of instructions that outlines how many treatments there will be, how much of each treatment there will be, and how and when each treatment will be applied? This is important if you are to get the best from the next section – Getting the trial underway.*

. So if it looks like a tight fit cut down on plot size rather than dropping replicates.

If you're OK so far, fill in the trial space by marking out the individual blocks and plots in accordance with your trial plan.

Once the basic trial layout has been marked with pegs or flags it's time to re-consider whether the design matches the site. Have the blocks been arranged to minimise variability within them? Do they form a path *along* the site variability? Have the plots been arranged to minimise variability between them? Do they form a path *across* the site variability? Are there any plots that are likely to cause problems in future due to weeds, obstructions or other irregularities?

If the match between site and layout isn't quite right, now is the time to change it. Screwing up a piece of paper and re-planning the layout will cost you a lot less than establishing a problematic trial.

You may be able to re-arrange the trial to better match the site. The principles outlined in Section 4 "Getting the design right" should help, particularly those relating to "Is bigger better".

Alternatively, it may be better to move all or part of the trial to a different paddock. There's nothing wrong with splitting a trial into separate parts, as long as you move whole blocks. This is outlined in Section 4.6 "Getting mre bang for your buck".

Don't implement a good trial design on a poorly matched site.

**Summary:**

**Once you've marked out your site, check that it fits the criteria for arrangement outlined previously. If it doesn't, you should consider re-designing the trial or locating it elsewhere.**

**5.2 Measurements on the site**

Once you're happy with the general layout of the trial in the paddock or paddocks, it's time for a more detailed examination of each plot.

Double check the plot dimensions, especially if you're measuring yield from a weigh wagon. Without an accurate measure of plot area you won't get an accurate measure of yield. Use a tape measure to get plot length. Don't use paddock length because (1) your paddock may not be as square as you think it is and (2) you probably won't harvest the whole plot length in any case, because you'll be trying to avoid headlands and other 'untypical' areas.

Walk the plots after they've been marked out. Check for patches that might need special treatment, such as weeds, hollows, tree stumps, compaction zones, camping grounds. These may not be big enough to prompt you to alter the trial layout, but you should make a note of them. A trial map is good for this purpose. If these patches visibly affect the crop you may wish to exclude them from yield or other important measurements as the trial progresses.

If measurements of 'initial conditions' are required to achieve your aims you should make these now. These kinds of measurements are necessary when (1) treatments are likely to cause an important change from initial conditions (eg. tillage or fungicide trials) and (2) specific site characteristics are likely to influence your results. More details are given in Section 3, "Getting the measurements sorted out". But briefly...

If you're doing population, sowing time or hybrid trials, the treatments with the highest potential yield are usually the most likely to suffer from stresses such as inadequate water and nutrients. The higher the potential yield, the more resources required by the crop. For this reason, it's a good idea to check out the fertility status of your site to see whether there'll be enough nutrients to support the growth of the high potential yield treatments. Soil testing is the way to go here. Similarly, a knowledge of the water holding capacity of your soil will help to interpret trial results.

If you're doing fertiliser trials, a knowledge of soil fertility is required because the response to a given fertiliser is highly dependent on the availability of that (and possibly other) nutrients in the soil.

If you're doing herbicide trials, a knowledge of weed populations and types is a good idea.

If you're doing fungicide trials an understanding of existing disease prevalence is important.

By the same token, if you're doing tillage trials, baseline measurements of soil characteristics that you're expecting to change should be undertaken. These might

include tests of soil organic matter, permeability or depth and severity of compaction layers.

Make soil tests on a plot-by-plot basis if you can afford it. This gives you replication of starting values which, depending on what you want to know, can be just as important as replication of yield or other 'final' results. You needn't get a complete soil test done on each plot sample. If you're doing a nitrogen fertiliser trial, it might be a good idea to get a plot-by-plot analysis of nitrogen and a complete soil test done for the whole site.

If you're not interested in the expense of plot-by-plot soil tests it's still a very good idea to get a soil test for the whole site. This can be done using approved methods (a minimum of 20 samples taken from a zig-zag profile across the site) or by taking a minimum of 5 samples from each plot, bulking and sub-sampling.

Ideally, soil tests should be done at least a month before planting. This gives you time to get the results back before action is required. You can then decide whether or not the site is still suitable for the type of trial that you have planned. See Section 5, "Getting the trial site right" for pointers. For instance, there's probably not much point in doing a phosphorus trial if there's already stacks of phosphorus present.

If you're doing trials on standing crops (eg. foliar fertiliser or fungicide application) it could be a good idea to make plot-by-plot measurements of key crop variables (eg. starting biomass, nutrient content or disease profile). This is the only way that you can be certain that differences between plots were caused by treatment and didn't exist before the trial started.

#### **Summary:**

**Understanding the starting characteristics of your trial site is essential if you are to (1) avoid problem areas, (2) track changes in site conditions and (3) meaningfully interpret your trial results. Plot-by-plot measurements of relevant variables made before the trial has started are the key to understanding your trial site.**

### **5.3 Making notes**

Gathering all of the information outlined above will be of use to you only if you can find and understand it when you need it.

Record keeping is a really important part of conducting on-farm trials. In addition to helping you interpret your current trial, good records can be referred to for years to come. In this way, they can save effort on trials in the future.

It's a good idea to have a special book devoted to your on-farm trial work. Make it the place where you keep every piece of information related to the trial – aims, treatments, design, measurements, everything. A book is probably better than a computer file, because you can cart it round with you in the paddock and enter any measurements or observations on the spot.

Data sheets to help you plan, make and record measurements at each of the 4 stages of your trial can be found in the "Data Analysis Toolboxes".

### **Summary:**

**Keep records of everything relating to your trial in one place, preferably a book. This will help you to most efficiently run your trial and will be a valuable resource for making future farm business decisions.**

### **5.4 Making sure**

If you've followed all of the guidelines to date, you should now be the proud owner of:

- A clear statement of what you want to find out from your trial, and why.
- A pair of treatments that you'll compare in order to generate that information.
- Equipment that has been calibrated to accurately apply those treatments.
- A calendarised list of the measurements that you'll make in order to gather trial results.
- A trial layout for a relevant site that has been replicated, randomised and arranged to give each treatment a fair chance of success.
- A paddock or paddocks loaded with pegs and holes. The pegs show you where the plots are. The holes are a reminder of your completed before-trial measurements.
- A book containing records of the above.

You're now almost ready to establish the treatments in your trial.

Before you hit the paddock, cast an eye over steps 1-5 of the relevant trial template to make sure that you haven't missed anything.

#### ***End of section critical decision point***

*You're just about ready to put your treatments in. Have you made measurements on the important characteristics of your site? If you're asked 'what was the trial site like', will you be able to provide data on the conditions or just general impressions? Both are important if you are to get the most from your trial.*

*Have you quickly double-checked that you've done all the planning necessary for the future of your trial? The relevant trial template will help.*

***It is important that you can satisfy these questions if you are to get the best from the next section – Getting the trial underway.***

## 6 GETTING THE TRIAL UNDERWAY

*By the end of this section you will be able to verify that your treatments have been applied in accordance with your trial plan. You should aim to have entered into your trial book (1) any deviation of treatment application from the trial plan and (2) records of relevant site conditions at the time of treatment establishment.*

On-farm trials are a bit like painting a house – the preparation's the hard part and the job itself is relatively easy. As a reflection of this, the last half of the On-Farm Trial Guide is going to be a lot shorter than the first half (!). From now on, we'll largely be reminding you to undertake the actions planned in the previous sections.

Because many of these actions will vary widely from trial to trial, we suggest that you rely mainly on the trial-specific templates provided in layers 2 & 3.

This section is designed to provide you with background information in the event that you have questions arising from the use of the trial templates.

### 6.1 Establishing the treatments

In the vast majority of cases you'll be establishing the treatments using your own farm equipment. You know how that works better than we do, so we won't spend time teaching you or your grandmother how to suck eggs.

Having said that, however, establishing a trial is a little different to establishing a commercial crop.

Trials usually involve comparing something 'old' with something 'new'. While you can be pretty confident that you're on top of the old technique, the new one will probably require special care. Give it this care by double-checking that treatments are being applied as intended in the trial plan.

This involves 2 steps. First, before establishing any treatments, calibrate your equipment to ensure that it *can* apply the right amount and distribution of material. Second, after establishing each treatment, ensure that it *has* applied the right amount and distribution of material.

The correct calibration method for your equipment should be outlined in the owner's guide. If it isn't, consult an appropriate chemical or fertiliser application handbook.

There are several methods for checking that the right amount of material has been applied. The easiest is probably to measure the amount of material going into your equipment and, at the end of each application, emptying it and measuring what comes out. If the amount doesn't tally with your plans, assess whether this will have compromised your trial and take the appropriate actions. These could include abandoning the trial or making 'touch-up' re-applications.

#### **Summary:**

**Correctly establishing treatments involves 2 steps. (1) Calibrating all necessary equipment before treatment application. (2) After treatment application, checking that treatments have been applied correctly.**

## **6.2 Monitoring and recording conditions at treatment establishment**

The time of treatment establishment is a red-letter day in the life of your trial. What you do and what the site is like on this day can have a big bearing on trial outcomes.

Immediately after you've finished the job, record the conditions prevailing during treatment application. Include in your records your impressions of how it all went.

Make notes of possible problems and things to look out for in future. These will help to interpret the results at the end of the trial. The sheets provided in the "Data Analysis Toolboxes" will help.

Record details of crop type and variety, sowing rate, planting date, row spacing, soil temperature, equipment used and anything else that may help to explain crop establishment and performance.

The sheets provided in the "Data Analysis Toolboxes" will help.

### **Summary:**

**Ensure that you have accurate records of conditions and activities at the time you establish treatments. The sheets provided in the "Data Analysis Toolboxes" will help.**

### ***End of section critical decision point***

*Have you been able to verify that your treatments have been applied as planned? If not, how can you be certain about the reliability of your trial? If the treatments have not been established according to plan, does it significantly reduce the value of the trial? If so, you may be best to keep your trial plan but implement it in a different place or time. There's still quite a bit of work to be done, and it would be best to do it on a well-established trial. If, on the other hand, it all looks OK, then move on to the next section -7 Monitoring the trial – enacting your operational plan..*

## **7 MONITORING THE TRIAL – ENACTING YOUR OPERATIONAL PLAN**

*By the end of this section you'll have revisited the trial measurement plan developed in Section 3, "Getting your measurements sorted out". You will have assessed which of those measurements you will continue with, which you'll drop and which you'll add as the trial progresses. You should aim to have a firm but flexible trial measurement regime to ensure that you don't miss key measurements and that you are able to add in new ones as trends and events become apparent.*

*If you've done a thorough job of planning your trial measurements, as outlined in Section 3, you'll find this section a breeze. We'll be looking mainly at why you might want to deviate from that plan. So we won't be presenting much detail here.*

The type of measurements that you'll make depends very largely on the type of trial that you're running. You'll find specific advice on these in the "Type of Trial" Booklets. You may want to refer to either the appropriate booklet or your trial measurement diary as you read this section.

Details on how you might actually make those measurements are given in the "Crop Specific Tips".

*If you've arrived here without first having passed the critical decision point at the end of the "Measurements" section, you'd better go back. What follows won't have a meaningful context.*

### **7.1 Observing pre-planting treatment effects**

Observation of pre-planting treatment effects is likely to be most important in situations where different starting conditions form part of the treatments.

This will occur in trials that compare tillage treatments, crop rotations, drainage, liming or a range of other treatments designed to alter the crop environment.

Clearly, in these situations, paddock conditions before planting are a major part of the treatments being applied. For this reason, it is important that starting conditions are monitored closely.

Details of the types of pre-planting observations that might be made in a tillage trial are given in the tillage trial template. Additional or different factors may be important to you. If that is the case, devise measurements that enable you to quantify how those factors differ and what impact that might have on crop performance.

If you are devising pre-planting measurements that aren't covered in a trial template these general guidelines might help...

Look at how your two treatments differ. Which differences are likely to have an impact on crop establishment or performance? Can you measure those differences directly by counting or weighing? If not, can you measure them indirectly using a scoring system or proxy variable?

Proxy variables can be really useful if they integrate a range of properties. The number of days of surface ponding after heavy rain (for instance) can be a good

proxy for soil structure. It incorporates elements of aggregate stability, infiltration and permeability that, on their own, aren't the best indicators of soil structure.

**Summary:**

**If your treatments are likely to respond differentially to pre-planting conditions it is important that any differences in these conditions be measured and recorded.**

## **7.2 Checking that treatments have eventuated as planned**

When establishing the trial as outlined in Section 6, "Getting the trial underway", you checked whether material was applied in accordance with the trial plan. This was an important step in verifying that your trial was comparing what you expected it to compare.

The next step is to check whether the different application of material has actually given you the treatment differences that you'd planned.

For example, if a population trial was undertaken, you will have checked that the correct number of *seeds* were sown per unit area. That's a useful piece of information, but it's not nearly as important as knowing whether the number of *plants* has eventuated as planned. After all, it is the population of plants, not seeds, that matters when it comes to determining yield.

**While it is important to know whether the treatments were *applied* as planned, it is even more important to know whether the treatments *eventuated* as planned.**

Because lots of things can happen to a crop between sowing and harvest it's important that you check that you've actually achieved the treatment differences that you set out to achieve. Wherever possible, make direct measurements of fundamental treatment differences.

This is a simple task where you can readily measure the treatments themselves. In population trials, for instance, it's simply a matter of counting the emerged and final plant population. Furthermore, if treatments don't eventuate as planned, you can investigate why that has occurred by looking at seed populations (digging) to see whether seed placement, germination or survival is to blame.

Only by quantifying the fundamental trial variable (population) are you able to (a) see that the trial is running according to plan, and (b) find out which part of the production system may be limiting.

Unfortunately, it is not always possible to directly check that the treatment has eventuated as planned. It isn't possible, for instance, to measure the amount of fertiliser or herbicide that has been added to a paddock once it's been applied! For this (very common) type of treatment, it's important that close observation of other factors is made during the season. This is covered in the next section...

**Summary:**

**Whenever possible make measurements to check that your treatments have turned out just as they were planned.**

### 7.3 Observing in-season treatment effects

Lots of things happen to a crop between sowing and harvest, as you'll have noticed when doing your regular crop scouting. Some of them are unexpected, but many of them can be anticipated and planned for. Monitoring your trial regularly will enable you to catch and record both kinds of crop events, which will help to interpret the final trial results.

**It's a good idea to have a look at your trial at least once a month – once a week is much better.**

The key to really successful trial monitoring is to keep your treatments in mind while you're looking, and seek things that may or should differ between them.

If your visual observations alert you to interesting or useful treatment differences it may well be worth making objective measurements of them. If you think there have been growth differences early in the season, why not harvest some plants and weigh them to see how different they really are? This will enable you to confirm or deny your suspicions and will provide 'hard' numbers upon which sound practical and financial decisions can be based.

It's also important to anticipate and look for both direct and indirect effects of the treatments applied. Direct effects are those that are a straight-line consequence of the treatment that you imposed (eg. different plant populations in a plant population trial). Indirect effects are those that are a secondary response to the treatments (eg. increased lodging due to high plant population).

Examples of important indirect treatment effects include things like...

- The simple act of cultivating soil differently can have significant effects on fertiliser requirements, crop emergence, crop development and pest and disease incidence, quite apart from its effect on soil tilth.
- Increased plant populations can reduce weed severity and bird damage but can increase the risk of lodging.
- Broadcasting fertiliser can increase weed vigour compared with banding, by enhancing weed access to nutrients.
- Side-dressing fertiliser can reduce plant populations ('Sheffield canker') and growth (root pruning) if placement is too close to the row.

**It's worth devoting some time to anticipating indirect treatment effects because, even though they aren't always obvious, they can often swing the balance for or against a given treatment.**

For example, a trial examining the effect of fertiliser placement on yield may have shown that broadcast fertiliser increased yield more than fertiliser applied 'down the spout'. That might be the real result, or it could be an artefact of the trial method. It's possible, for instance, that 'down the spout' application had a lower yield because it had a lower population. Without your knowing it, the high concentration of banded fertiliser generated ammonia gas and, because it was close to the seed, it killed quite a few seedlings. Only by measuring emergence and final plant stand would you have found out that this had occurred.

The benefits from having made the measurements on population are several. First, you won't have drawn the wrong conclusion from the trial. Second, you'll have learned a trick or two about applying fertiliser. Third, next time you do the trial you will hopefully avoid the toxicity problem and will get a reliable trial result that you can use to optimise your farm operations.

*The trial templates give examples of measurements of both direct and indirect treatment effects that should be considered when running a trial.*

**Summary:**

**Only by measuring key – and sometimes apparently unrelated – variables during the season, are you able to fully and reliably interpret final trial results.**

***End of section critical decision point***

*Have you been monitoring your trial regularly? If not, will you have learned anything about why your final results turned out as they did? Have you made measurements that show that your treatments actually turned out the way that you planned them? If not, can you be sure that your trial is comparing what you think it is comparing? Have you made measurements to verify that the application of treatments hasn't had unintentional or indirect effects on crop performance? If not, can you be certain that you'll draw the correct conclusions from the trial?*

*If you can't give positive answers to the above questions, you may want to treat your trial results with some caution – you just can't be certain that they are reliable. If, on the other hand, it all looks OK, then move on to the next section – “Finishing the trial – getting the harvest done”.*

## **8 FINISHING THE TRIAL– GETTING THE HARVEST DONE**

*By the end of this section you'll be able to plan your trial harvesting operations so that they run like clockwork. You should aim to have a harvest plan written out and ready for action. Following that, you should have a fistful of data sheets and a cupboard full of samples.*

By this stage, you'll have done most of the hard work on your trial. The vast majority of the planning and measurements are behind you. The harvesting operation itself is generally a fairly mechanical operation. You have, after all, harvested plenty of crops before as part of your farm business operation.

We've already provided a reasonably detailed description of the trial harvesting operation in each trial template.

We've also provided a detailed description of how to make yield measurements for individual crops in the "Crop Specific Tips" section.

For these reasons, this section will be short and sweet. We just want to point out that...

Harvesting a trial is much like harvesting a commercial crop. You want to gather all of the crop product without loss and you want to know how much you got.

Additional considerations when harvesting trials are outlined below.

### **8.1 Selecting the harvest area**

Ensure that you have accurately located plot boundaries. You don't want to be including areas that didn't receive any treatment in your plot harvest. Nor do you want to be mixing treatments in the one harvest.

Ensure that you do not include parts of the plot that have non-treatment effects in them. These might include areas that have been trampled by rampant stock, have not been established properly, have suffered from flooding or have copped some spray drift.

- Exercise careful judgement when excluding parts of a plot from harvest. You don't want to exclude them just because they don't look any good. You need to assess whether the 'problem zone' is related to the treatment. If the problem zone is related to the treatment in any way – either directly or indirectly – then it should probably remain in the harvest area. The reasoning here is that if the problem is related to the treatment it could easily recur if conditions allow it. You need to know this before adopting the practice in your commercial crops.
- Your records of in-season observations and measurements will prove invaluable in helping you to sort out whether trends or events that you see in the plots at maturity are likely to be related systematically to the treatments.
- An example. When conducting a plant population trial, we noticed that birds paid particular attention to some plots and ignored others. At first these observations were made near a shelter belt, and we assumed that the birds were arriving from the trees and grazing on plots based on their proximity to perching sites. We considered eliminating from the trial plots next to the shelter belt. Closer

examination indicated that, while the damage was indeed greater near the shelter belt, it was spread throughout the entire trial. Further examination indicated that damage tended to be much greater in the low populations than in the high populations. It seemed that birds were either better able to see or perch on cobs in low population treatments. For this reason, we decided to keep all of the plots in the trial. Bird damage was related to treatment and so should not be eliminated from the trial. Furthermore, because the trial was designed properly – taking the location of the shelter belt into account – the interference caused by proximity to perching sites could be eliminated using the statistical procedures outlined in the “Data Analysis Toolboxes”.

Ensure that you have accurately determined the harvest area. Yield is always expressed on a mass per unit area basis - tonnes per hectare, for example. If your assessment of harvest area isn't accurate then no amount of precision in measuring the mass of harvested material will give you an accurate yield estimate. Use a tape measure for this purpose. Nothing else will give as reliable a result.

Try to minimise edge effects. You'll already have given thought to this when designing your trial, as outlined in Section 4.5 “Error! Not a valid link.”. Basically, you want to minimise the harvesting of plot edges because they usually differ from the plot centres. In addition, most of your commercial crop is like the plot centres so you'll get the best match between trial and commercial crop results by harvesting only from plot centres.

**Summary:**

**Accurate yield estimates require accurate determination of harvest area. Measure this with a tape measure. You can only reliably assign yield differences to treatments if harvest areas fairly represent treatment effects. Minimise edge effects and don't harvest areas that have clear non-treatment problems associated with them.**

## **8.2 Harvest equipment**

Ensure that measuring equipment has been calibrated – recently and reliably. You can never assume that a measuring device works as intended. No matter how expensive or high-tech, it needs to be calibrated and checked against standards. Some details on this are provided in “Crop Specific Tips”.

- Weigh wagons need to be calibrated using standards of known mass. This should be done on site, once it's in place for measurement. Moving a measuring device, especially on rough tracks, can put them out of calibration.
- Yield monitors need to be calibrated against weigh wagons. Details are provided in “Crop Specific Tips”.

Ensure that harvesting and measuring equipment are clean and empty. You need to measure each plot individually, without interference from other plots or crops.

**Summary:**

**Accurate yield estimates require accurate determination of crop mass. This can only be achieved with correctly calibrated equipment.**

### 8.3 Data recording

It goes without saying that you'll want to keep complete and accurate records of trial outcomes. Surely there'd be little point in doing the trial if you weren't going to keep good results?

Good record keeping has several advantages. It will enable you to more fully and reliably interpret the data. This can be important when new information comes to light at some future occasion. Without good records you'd have to rely on your memory for the details (we wish you luck!). Good records enable you to share results and can provide the basis for discussion with people with specialist skills or experience. Most importantly, good records provide the basis for you to repeat your wins and avoid repeating your losses.

*Data sheets are provided in the "Data Analysis Toolboxes".*

While it is our hope that the data sheets are comprehensive enough to help you record essential results, don't treat them as all-inclusive. The unique combination of conditions and treatments that make up a trial mean that you may well want to record measurements that we haven't anticipated. In the likely event that this occurs make up your own data sheets.

The primary requirement of a data sheet is that it enables you to record:

- primary data that you require (eg. mass, number)
- appropriate units of measurement (eg. kg, number)
- the area from which data were obtained (eg. 1 m<sup>2</sup>, or 3 plants from each of 2 rows)
- the unique plot number or name from which the data were obtained
- the date that data were obtained

It's almost certain that you'll make some mistakes when making or recording measurements. You can most easily fix these errors if you record as much primary information as possible.

For example, if you suspect that your plot area measurement doesn't seem right you can spot the error most easily if you've entered both the length and width on the data sheet, rather than just the calculated area. By the same token, you'll most easily spot a mistake in yield if you have each of the harvest mass, moisture content and corrected harvest mass (at a standard moisture %) entered separately.

**Do not dispose of the material harvested from a plot until you are certain that you've gathered all of the data required from that plot.**

A well-designed data sheet will enable to most readily check that you have filled all of the necessary data 'holes'.

Some of your data will come from subsamples of harvested material. It is important that subsamples from each plot are kept separate. If they are pooled you will lose most of the benefits that you gained from implementing a good trial design.

Subsamples will be required for common 'quality' measurements such as moisture content and rejects. If you're taking subsamples for these purposes you might like to

consider whether you want to keep them for further tests. Some of these are outlined in the next section – “After the trial”.

**Summary:**

**Record keeping isn't the most exciting part of the job, but is essential if you are to extract the maximum value from your trial. Keep accurate and reliable records of your measurements and observations on purpose-built data sheets. Data must be recorded on a plot-by-plot basis. Store your data sheets carefully for future reference.**

***End of section critical decision point***

*Have you harvested the trial from a part of the plot with known dimensions? Was the harvest area free of unrepresentative 'problem zones'? Did you measure harvested masses with calibrated equipment? Did you record the results on a plot-by-plot basis in a form that will enable you to retrieve them for later analysis? If not, can you be confident that you've got accurate, reliable and retrievable data to go on with? If, on the other hand, it all looks OK, then move on to the next section – “After the trial”.*

## 9 AFTER THE TRIAL

*By the end of this section you should aim to have a clear idea of whether you have completed all of the physical and data collection requirements of your trial or if further procedures and measurements will be required.*

### **A trial isn't over just because the 'final' harvest has been done.**

If you're interested in measurements of quality or storage characteristics you'll need to do more measurements or have someone do them for you.

- In most cases, this work will be done by an accredited lab or testing facility.
- The most common problem facing testing labs is that samples haven't been stored correctly. It is essential that your samples of plant, soil or other material are stored in such a way that their important characteristics are preserved.
- Samples that have degraded through infection by mould or bacteria, inappropriate dehydration or re-hydration, insect or vertebrate infestation or other undesirable '-ations' are not much use to you or a testing lab.
- Store and transport samples according to the instructions provided by your test provider.

In some trials, the after-harvest effects of treatments will have an important bearing on their relative profitability. Examples of this are given in each trial template, but can include...

- Treatments that influence the amount or type of crop residues. These will impose costs of incorporation or disposal that need to be taken into account if your trial objective was to investigate effects on gross margin. Residues may help to preserve soil structure but may also help to harbour pests and diseases. If these are likely to impact on the gross margin from the different treatments then they should be taken into account as part of your trial.
- Treatments that influence soil conditions. The extent to which these carry over to other seasons or crops will impact on the relative productivity and profitability of the treatments that you have just considered as part of your trial. You may want to know whether the treatments have influenced soil fertility, structure, moisture or some other factor. If so, there is no substitute for measurement of either direct or proxy variables.
- Treatments that influence the timing of key events, such as sowing and harvest. These sorts of changes can have a big impact on the timing of subsequent management events (such as tillage) and on the types of crops and productivity that can occur after the trial crop is finished with. The economic and practical value of these phenomena is likely to influence the real value of the treatments just tested.

### **Summary:**

**If your treatments are likely to have influenced the quality of harvested products or the capacity of your paddock to profitably support future crops**

**then it's probably a good idea to consider a range of after-harvest measurements as part of the current trial.**

***End of section critical decision point***

*Would measurements of additional variables such as product quality, soil properties or productivity of subsequent crops provide a more complete picture of the value and implications of the treatments that you've just tested? If so, you may want to give thought to additional after-trial measurements or even a post-trial trial. If this isn't required, now's the time to get into the numbers in the next section – "Analysing the trial results".*

## 10 ANALYSING THE TRIAL RESULTS

*By the end of this section you should have converted the numbers contained in your data sheets into hard information. That information should enable you to choose which of the two treatments examined in your trial will form part of your future crop production system. You should aim to have a clear statement of what happened in the trial and what this might mean for future trials and commercial crops.*

One of the problems that continually dog do-it-yourself experimenters is the issue of statistical analysis.

Let's not kid ourselves. Most people don't enjoy statistics. That's largely because most people want to get an answer, not a long explanation about how they can get it.

With this in mind, we'll explain *what* you need to do to convert your trial data into actionable information and knowledge. The discussion will be based around the use of the ANOVA tool in your "Data Analysis Toolbox". It's easy.

However...don't get the idea that you're going to get off scott-free! To get the most from your trial, you need to know a little bit about what statistics do. We'll give you that information, but as briefly and painlessly as possible. Don't worry - there won't be any incomprehensible jargon or funny formulae!

Whatever the purpose of your trial you will need to do some basic statistics to be sure that you're interpreting the data correctly. Simply comparing the averages from each treatment only gives you a small part of the story. A little bit of simple statistics will enable you wring every last piece of information from your trial. You've come this far – don't fall over at the last hurdle!

**Statistics are just a tool to help you to make sense of trial results. Their use conforms to the usual rules for tools. First, use the right tool for the job. Second, if you know how to use the tool properly, it'll make the job quicker, easier and better done.**

**The following section is a simple user's guide for your trial data analysis tool.**

### 10.1 Re-check that data is complete and accurate

Statistics conform to the old adage – 'rubbish in, rubbish out'. Make sure that you keep the rubbish out by checking over your data sheets before you start playing statistician.

The first thing to check for is missing data. If you've used the pro-forma data sheets in the "Data Analysis Toolbox" this should show up fairly readily as gaps in your results.

If you can, it's best to fill the gaps made by missing data. There are several approaches that you can use.

- The best approach is to try to find the missing data. Was it written somewhere else?

- Second best is to see whether you can reconstruct the missing data using other measurements that you've made. For example, if plot area is missing it can be calculated from recorded measurements of plot length and width.
- As a very last resort, you can consider filling the gap with the average value obtained from the other replicates of that treatment. For example, if yield data are missing from the third replicate of treatment A then this data can be reconstructed using the average value for treatment A from replicates 1, 2 & 4. Use this method only in an emergency, as it undermines the reliability of your trial by using data that didn't actually occur. Do not under any circumstances do this for more than 1 out of 4 replicates. Having 25% of your data 'fake' is bad enough!

**No data is better than wrong data. If you come across data that is clearly incorrect it is better to delete it than to include it in your analysis.**

So how do you pick 'wrong' data? This can be a bit of a slippery slope. You have to be very careful not to delete data just because you don't like the look of it. Always keep in mind that plots do vary and that even 'obvious' treatment effects don't always show up in the cold hard numbers.

Data is probably only 'wrong' if:

- You know that a mistake was made when the data was being taken. This shouldn't occur very often. If you know that a mistake's been made you can usually correct it on the spot, often by repeating the measurement.
- The data in question isn't consistent with related data. Yield, for example, is often calculated from fresh mass after correction for moisture content. If you have similar fresh masses and moisture contents, but widely different yields then the yield data may be wrong. Check it and see. If you can't fix the problem, then the data may well be wrong and should be deleted.
- The data is simply not possible. If you do enough trials, you'll end up with some of these. Wonder where that 100 t/ha grain yield came from? So do we. Sometimes you can't tell where it went wrong, but you know that it has. When this happens, there's only one answer – delete it.

Accidents, omissions and errors *will* occur! The following brief list highlights the most common sources of error. We hope that this helps you to track and fix yours.

- Data are transposed. You wrote '71' when you meant to write '17'.
- Data was entered in the wrong box. The yield was 12 (t/ha) and the moisture content was 24 (%), and not vice versa, unfortunately.
- The weigh wagon wasn't emptied between plots. Look for numbers that are roughly double the expected size and see whether they 'fit' with corresponding missing values.
- The balance wasn't tared or zeroed properly. Any 'extra' mass might well correspond with that of the container that was used for weighing.
- You wrote the right piece of data in the wrong place. Look for data that are wrong in 'pairs'. This often occurs when one measurement was recorded in the

wrong spot and the other error was made because data were entered into the only space that was left.

These are the errors that I usually make, but you're sure to come up with some of your own.

#### **Summary:**

**Start your statistical analysis only after you've checked that all data are present and correct. If in doubt, throw it out.**

## **10.2 Entering data**

**Data entry is an easy 2-step process.**

**First you have to decide which measurements need to be statistically analysed and which don't. Second is typing the numbers in.**

Quantitative or numerical data such as yield, biomass, test weight, moisture content, bulk density and lodging should be analysed statistically using the ANOVA tool in the "Data Analysis Toolbox". You can do statistical analysis on these because they are numbers. You should do analysis on them because it will help you to get the right answer from your trial.

Qualitative data can't be statistically analysed using standard techniques, because they aren't numbers. Scores such as 'poor', 'good' and 'very good' are qualitative data that you can't enter and analyse statistically.

**You should enter and statistically analyse all relevant qualitative (number) data. You should evaluate and analyse the implications of qualitative (observations or impressions) data.**

Whether or not the data requires statistical analysis, you should analyse it. For observations, just think about what the data means and record your thoughts.

*Data that is to be statistically analysed needs to be entered into the ANOVA tool of the "Data Analysis Toolbox".*

*First of all you'll need to download and open this tool. Do that by saving the ANOVA tool on a place where you can find it easily on your hard disc.*

*Now that you've got a copy of the ANOVA tool, you can use it whenever you want, without having to access the cd, disc or website that you've been given. You can simply open it using the Excel spreadsheet programme. If you have any problems, please contact the author for help.*

#### **How do you use the ANOVA tool?**

*We'll cover data entry here, and will discuss data analysis in the next section – "Statistically analysing the data".*

When the ANOVA tool is open, you'll see that there are two columns with four yellow rows (eight cells total) into which you can enter data. The two columns represent

your two treatments. The four yellow rows represent the four replicates that you had of each treatment.

Additional green rows have been included in case you decide to go for more than four replicates.

*Never be tempted to double your number of replicates by just entering the same data twice (by turning 4 replicates into 8 replicates, for example). That's cheating, and you'll end up with the wrong idea about your results. Have a play and see what we mean – entering the same data twice often makes your treatment differences seem more certain than they really are.*

***It's important that data is entered into the right places. Only data from treatment 'A' should go into the first column and only data from treatment 'B' should go into the second column. Mixing these up will give the wrong result.***

It's also important that the data are entered into the correct replicate row. Both pieces of data from replicate 1 need to go into the first row, replicate 2 into the second row, and so on. This enables the ANOVA tool to work out whether differences among replicates in location or time had any effect on the final result. This is partly what makes the ANOVA tool such a powerful data analysis device. Unlike simple averages, it's able to sift through other information to help show what the real treatment effects were, independent of differences between paddocks or seasons.

Once your numbers are entered, check to see that they're all correct.

Once you're satisfied, hit the 'produce report' button, and you're ready to get on to the next section.

*If you make an error simply change the numbers and delete the report using the 'delete report' button.*

**Summary:**

**Numerical data should be entered into the ANOVA tool for statistical analysis. Put all of treatment A into column 1 and treatment B into column 2. Put data for replicate 1 into row 1, data for replicate 2 into row 2, and so on, until all 8 cells are full (for a 4 replicate experiment).**

### **10.3 Statistically analysing the data**

Once you hit the 'produce report' button, the ANOVA tool does all of the number crunching and interpretation of statistics for you.

It will help you to work out:

- The difference in 'size' (eg. yield, biomass) between your two treatments
- Whether or not this difference was caused by a real treatment effect or simply by chance
- The extent to which you can trust the results as the basis for decisions about your farm system

***The ANOVA tool does this by giving you 3 pieces of information:***

### 10.3.1 P-value.

Statisticians are cagey creatures and believe that anything can happen due to chance. The P-value is the probability that the trial results occurred due to chance, rather than as a result of the treatments that were applied.

- In other words, a P-value of 0.05 means that there's a 5% chance that the observed difference between treatment A and B was just luck. Looking at it from another viewpoint – there's a 95% chance that the observed difference between your treatments was caused by the treatments.
- By the same token, a P-value of 0.10 means that there was a 10% chance that the observed difference between treatments A and B was caused by luck. You can be 90% certain that they were caused by treatment.
- Because it's a guide to the odds, the P-value gives you some information about the likelihood of repeating the benefits (or otherwise) of a given treatment. So if the P-value is 0.10, as above, you can be 90% certain that the results were caused by treatment. This also means that you could expect a similar sort of result 9 years in 10, if the trial was repeated under similar conditions.
- Clearly, the P-value gives you useful information that you can use to estimate the reliability of your trial results, and the relative risks and benefits of a given pair of treatments.

### 10.3.2 Least significant difference, or LSD.

The LSD is a statistical term that you can use to see whether two numbers are really or significantly different from one another. The value of the LSD changes in response to all sorts of statistical considerations that you're unlikely to care about. No need to worry - the ANOVA tool calculates the LSD by assuming that you'll be happy with a real treatment difference 9.5 years in 10. That means that it calculates the LSD using a P-value of 0.05.

- The size of the LSD tells you how different your two treatments really are. This is most easily shown by example.
- Let's say that the averages of treatments A and B are 10 and 15 t/ha, respectively, and that the LSD for them is 2 t/ha. Take the lowest treatment average from the highest:  $15 - 10 = 5$ . This 5 t/ha difference is more than the LSD (least significant difference) of 2 t/ha, so you can be confident that treatments A and B really are different.
- It's called the least significant difference (LSD) because your 2 treatments need to differ by at least the value of the LSD in order to be considered truly different.
- Let's look at another example. Let's say that the averages of treatments A and B are 10 and 15 t/ha, respectively, and that the LSD for them is 6 t/ha. Again, take the lowest treatment average from the highest:  $15 - 10 = 5$ . With an LSD of 6 t/ha, our treatments need to differ by at least 6 t/ha in order to be truly different. Our treatments differed by only 5 t/ha. That's less than the 6 t/ha needed to be truly different, so we conclude that treatments A and B were not really, or significantly different.

***How is it that 10 and 15 t/ha can be considered not different?***

This is where the subtlety of statistics comes into play. Statisticians wouldn't say that they aren't different, just that there's a good chance that they aren't different.

In their language, they say that they are not significantly different. It all boils down to how certain you can be about what you've observed. Another example can best illustrate how this works....

Let's look at a rugby example.

- If the Wallabies beat the All Blacks 15-10 in a single game, does that mean for certain that the Wallabies are the better side?
- Could you be certain that if the same sides played again the score would be the same?

No. You'd have to say that, despite the 5 point difference in scores, there's no conclusive evidence to prove that one side is definitely better than the other.

A statistician would say that there is not enough evidence to conclude that there is a significant difference between the two sides.

*If you type these scores into the ANOVA Tool (Treatment A for Wallabies and Treatment B for All Blacks), you'll see that it says that 'a statistical analysis is not possible' with these scores. In other words, one game doesn't tell you anything much.*

If they played another game and the score was 16-10, the evidence would be mounting, and you could be more certain that the Wallabies were the better side, but you still couldn't be 100% confident that that was the case, could you? If you add these scores into the ANOVA Tool (the second row, with Treatment A for Wallabies and Treatment B for All Blacks), you'll see that a statistician would think that there was a 5-7.5% chance that the Wallaby win was a fluke. They'd only concede that it is 'quite likely' that the Wallabies are the better side.

If the results from the Third Test were 9 vs. 4 (another Wallaby win), even a statistician would have to concede that the Wallabies are probably the better side, and that there is only a 0.1-0.5% chance that their winning streak was a fluke. They'd even go as far as saying that 'you can be confident' that the Wallabies are the better side.

Run through these scores again, and look at the LSDs that the ANOVA tool gives you. You'll notice that as you increase the number of replicates the value of the LSD drops. This occurs because the 'hurdle' that your treatment differences have to cross to be 'significantly different' gets smaller as you get more confident about the results.

**You can only say that there's a *significant difference* between two teams or treatments when you can be reasonably confident that you've got a reliable comparison between them that shows consistent differences.**

The more comparisons that you have, and the more similar the results from each, the more confident you can be of the reliability of your comparison. In other words, more replicates and more similarity of results among replicates makes it easier to pick the

difference between two treatments. These also give a smaller LSD and a lower P-value.

Statisticians tend to be cautious people and are usually only reasonably confident when there's at least a 95% chance that their assessment is reliable (P-value 0.05). The ANOVA tool in the Data Analysis Toolbox will calculate an LSD that tests whether or not treatments are different with 95% certainty (P-value 0.05).

### **10.3.3 An interpretation of the P value and LSD.**

*Working out what the P-value and LSD mean for your trial is an important part of the data analysis. We urge you to look at them and interpret them for yourself, using the guidelines provided above.*

*To help you get the most from them, we'll also make comments on the reliability of the results that you've analysed. These will show up whenever you enter data and push the 'produce report' button of the ANOVA tool. These are self-explanatory.*

While they are essential for understanding your trial, the results from the ANOVA tool have a number of limitations:

- They can only look into the past, not the future. They tell you about what has happened in your trial. They cannot tell you what would happen if you did the trial again.
- They apply only to the conditions in which the trial was conducted.
- They do not give information about why results occurred, only what results occurred.

Understanding these limitations will help you to get the best out of your trial. Statistical analysis is not enough. Once you've done the stats, you need to think about what they mean for your operation. The real analysis of your trial results begins when the stats have been finished.

**Statistics and data analysis are not the same thing. Data analysis is the art of turning statistics into knowledge that you can use to improve your farm system.**

#### **Summary:**

**The ANOVA tool is an essential part of your trial data analysis, because it will help to show how much faith you can have in the results. Differences between treatments can't be relied upon without P-values and LSDs to support them. If you can't be at least 95% certain that treatment differences are real (with P-value less than or equal to 0.05), then you should take the view that the treatments may have had no effect, no matter how different they appear to be.**

### **10.4 Analysing the implications of the data**

As we mentioned in the previous section statistical analysis is the start, not the finish, of the data analysis process. Your task now is to turn the statistical information that you have into knowledge that you can use to improve your farm system.

**Where statistical analysis tends to be cut-and-dried, analysing the implications of your data relies on judgement. You'll need to make some big calls about what the trial results mean for you and your business.**

Your statistical analysis will provide you with 3 basic outcomes:

- The results indicate that there was a significant difference between the treatments. You can be confident that treatment A really was different from treatment B.
- The results indicate that some indecision is in order. The treatments weren't different enough for you to be really confident (eg. 80%) about a significant treatment difference, but nor are they so similar that you'd treat them as alike.
- The results indicate that there was no reliable treatment difference (P-value 0.30 or more).

Each alternative suggests a different set of possible action plans. We'll cover them in turn, below.

**If you found significant differences between your treatments:**

You can be confident that your treatment differences were real.

You now need to decide what the implications of this result are for your farm operation. Just because one treatment is 'more' than another doesn't mean that it's 'better'.

For example, if you did a population trial and found that yield was significantly increased by a higher population, you might want to consider a wide range of other factors before adopting the practice on part or all of your farm.

- What was the effect on profit? You'll need to do some sums to be certain that higher yield equals higher profit.
- How much consideration should be given to other variables? Did lodging increase? Was this an acceptable increase or is it likely to enlarge your exposure to risk? What happened to quality?
- What is the seasonal context of the results? For instance, if lodging did go up, but just a little, was it because the season wasn't wet and windy? Would lodging become a problem in those circumstances?
- Scale. If it looks like the new treatment increases both profit and risk, should it be adopted on all or part of the farm, or not at all?
- What is the locational context of the results? Are the results likely to occur to this extent in different paddocks? Why? Why not?
- Capital requirements. Does the change in population require any change to your equipment, storage or transport needs?
- Inconvenience. Is the new treatment a hassle?

No doubt you already make these sorts of decisions whenever you look at new possibilities. We mention them here to reinforce the fact that a big statistical difference between two treatments doesn't necessarily mean that one is better for your system than another.

### **If you found that some indecision is in order**

You cannot be confident that your treatment differences were real.

- You needn't reject a new treatment just because you can't be confident that it is statistically different from an alternative. As outlined above, the new treatment may have other advantages that make it attractive, despite the lack of significant yield or other benefits.

It's possible that two treatments had a big numerical difference but only a small or no statistical difference.

- Again you needn't reject the new treatment, but you might want to proceed with caution.

Large but inconclusive (not significant) differences between treatments usually mean that the data is variable. That is, one or both of the treatments is a bit hit-and-miss.

- When this occurs it is worth examining why the treatments were variable. This will give you some insights into whether the treatments are inherently risky, or whether they varied in response to difficulties with the site, season or trial execution. If you can understand the factors that led to variability you may be able to reduce or avoid them and increase the certainty of obtaining a favourable outcome.

How you proceed with large but not-statistically-significant treatment effects also depends on your risk profile.

- If you found yields with a 5 t/ha difference (eg. 10 and 15 t/ha) that weren't significantly different, you might want to go ahead with the larger alternative if it poses no extra risk or cost. It could be that you have the possibility of gain with no possibility of loss.

Keep in mind, though, that variability is just that. It can go up or down, unpredictably.

- If you get numerical differences, large or small, and you can't be confident that they are statistically different, can you be confident that they'll always be up or always down?

If you have these sorts of doubts about a result, reject the new treatment or give the trial another go next year.

### **If you found that there was no reliable difference between treatments**

This is usually pretty conclusive. If there's no reliable statistical difference your preference between treatments will have to be based factors other than those for which you've done statistical analysis.

For instance, if you find that treatments A and B do not have a statistically reliable difference in yield, then preference will come down to factors such as cost, risk, convenience and 'gut feel'.

A 'no difference' result may be exactly what you're after. This will occur particularly where you're looking at treatments designed to cut costs, such as reduced rates of pesticide or fertiliser application. If you can reduce input costs with no significant effect on yield, you stand to increase profit – in the short term at least.

### **Summary:**

LandWISE Inc. - [www.landwise.org.nz](http://www.landwise.org.nz)

*Sustainable cropping through technology*

**Statistical analysis is intended to provide you with reliable numbers that you can use as a tool for decision making. Statistics will not make decisions for you. When analysing trial data you need to make judgements about the broader implications of adopting or not adopting a given treatment. Trial results are only relevant when they help to ask and answer questions that relate to factors such as profit, risk, season and scale.**

## **10.5 Recording results and conclusions**

By the time you've reached here, it will be obvious that you've had to put quite a lot of effort into analysing the results from your trial!

Whether it's all gone well or turned to custard you should record your major trial results, the decisions that you've made based on those results, and the reasons that you made them. This information will prove invaluable in the future.

Without detailed records, you're unlikely to remember the details of your trial in a week's, month's or year's time (depending on your memory). Clearly, forgetting the results and conclusions based on them will severely limit the value of your trial.

Having ready access to good trial data can help to solve new problems as they pop up, often without the need to conduct a separate trial. By this means, good trial records keep paying for themselves.

Good records make it much simpler to repeat successes and to avoid repeating failures. Give this task the time that it deserves – you'll never regret it.

### **Summary:**

**So much effort goes into establishing, running and analysing a trial that it would be a real pity to fall over at the last hurdle. Finish off your trial by writing a brief statement of the important results, conclusions and decisions coming from the trial. You'll find it a valuable business tool.**

### ***End of section critical decision point***

*Have you learned something new about how your crop ticks? Ultimately, this is more valuable than the specific results from your trial. While an individual trial is, by necessity, narrow in focus, the benefits of conducting trials are usually wide-ranging. Hopefully, the knowledge that you've gained will help you to improve some part of your cropping system.*

*Check out the next section – "Avoiding common trial pitfalls" to sharpen up your trialling skills.*

## 11 COMMON TRIAL PITFALLS

*This section is designed to help you avoid the mistakes most commonly made by On Farm Triallers. Even if you've read the other sections and think you've got it all under control, have a browse to see whether you recognise some old 'friends' that would be best avoided...*

### 11.1 Preventing trials becoming a trial - common pitfalls in trial design

Having demonstrated the steps involved in designing the 'perfect' trial, we'll now reinforce the principles that we've established by showing a few common pitfalls, and the reasons why they cause problems. First we'll look at common 'operational' problems (recommended reading for everyone interested in doing trials) then we'll look at some of the more technical aspects of trial design.

**Using the 'big three' of the four Rs of trial design from Chapter 4 – Replication, Randomisation and aRrangement – will get you sailing over the most common trial design hurdles, no worries.**

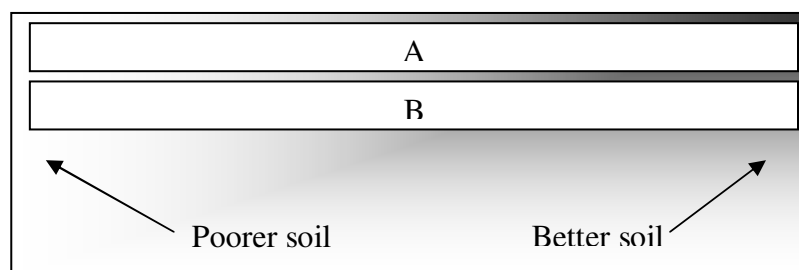
#### 11.1.1 Unreplicated side-by-side comparison of two treatments

This is the most common trial design seen on farms (Fig. 9 below). There is no replication, which means that the effects of site simply can't be separated from the effects of treatment. Consequently, any differences between the plots are just as likely to be caused by variation in the paddock as differences between treatments.

The obvious attraction of these sorts of trials is that they're easy to establish, monitor and 'analyse'. Essentially, they minimise effort.

Using the same rationale, why not save all the effort by calling one treatment 'heads' and the other 'tails' and then using a coin toss to decide which treatment is best?

This method is almost as reliable (and a lot less work) than an unreplicated side-by-side trial!



*Fig. 9 . An example of the most common on-farm trial design – the side-by-side strip trial. This incorporates most of the worst elements of trial design. The 'big three' of the four Rs (replication, randomisation and arrangement) are missing and, with them, the chance of getting reliable results.*

The basic problems have been described in the preceding sections but, to summarise: the 'big three' of the four Rs of trial design are missing which means that (1) there's only one contrast between the treatments which makes any comparison very risky; (2) the variability within plots is large because the plots are very long; (3) the variability between plots is large because one has more 'good' soil than the other and one has been placed next to a fence line and the other hasn't; (4) each plot contains atypical 'headland' areas.

The example above obviously has lots of problems, but it wouldn't take much effort to fix them. Let's use the 'big three' Rs approach.

The replication problem could be fixed either by

- (a) repeating the trial in three more paddocks,
- (b) repeating the trial over a number of years, or
- (c) a combination of these that gives four replicates. Easy!

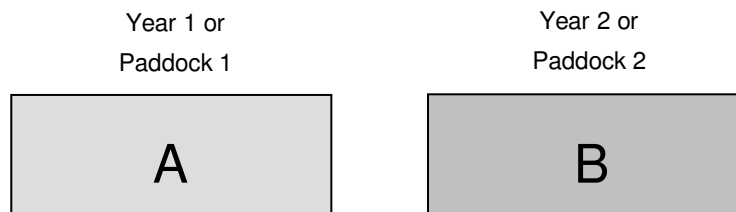
Once there's some replication, it's a simple matter of randomising the treatments within each 'block' (paddock or time) to fix the second R (randomisation).

The arrangement problem could be fixed by moving both plots towards the bottom of the paddock where the soil is more uniform. This would get away from the fence line, too. Making the plots a little shorter, or measuring from only the central areas would remove the 'atypical' influence of the headlands.

The simple use of the 'big three' Rs has made a silk purse from a sow's ear! Let's look at a few more examples to show how the 'big three' Rs can be used to solve some common 'real paddock' trial design problems.

## 11.2 Applying a single treatment to a whole paddock...

...and comparing the results with other treatments planted in different years or different paddocks is a common 'trial design' used on-farm (Fig. 10). The main problem here is that crop performance and characteristics vary widely from year to year and from paddock to paddock even when there's no change in treatment. Consequently, it's simply not possible to ascribe any observed differences to the new treatment that's been applied.



*Fig. 10 . An example of a common on-farm trial design – one treatment applied per paddock or per year. A favourite of snake-oil merchants, because the trial can only prove what you want it to prove.*

**This sort of 'trial' is a real waste of time and effort.**

Despite this, it's often used by those who want to reinforce a prejudice or sell some snake-oil, because there's a 50:50 chance that the year or paddock chosen for the new treatment will be 'better' than that used for the old treatment. When results are favourable it is used to 'prove' that the new treatment is best. When they're not, it was because this year or this paddock 'wasn't much good'. Either way, you lose!

***Apply some healthy scepticism to trials that don't incorporate the big three Rs.***

**11.3 Fertiliser trials without accompanying soil tests...**

...are quite common, despite the fact that they are usually a complete waste of time and money. It's well known that you will get a response to an added nutrient if it's deficient and that you won't get a (positive) response if it's not. You can most easily find out if a nutrient is deficient by doing a soil test undertaken using approved procedures.

Without a properly conducted soil test, the results from a fertiliser trial cannot be applied with any confidence to different paddocks or different seasons. The results are applicable only to the crop on which the trial was undertaken, so it would be unwise to use them to formulate a broader-scale fertiliser plan.

**11.4 'Head-to-head' trials that aren't really head-to-head trials...**

...are rather common and occur in several forms. The most common typically involve an intention to undertake a head-to-head comparison of fertiliser treatments. This is often done when the commercial planting has been finished, and left-over bags of seed are used for a trial. One fertiliser treatment ends up with one variety and the other treatment gets a different variety or a mix of two (because the seed ran out half way). Clearly, because the varieties differ to an unknown extent, it's not possible to determine whether any 'treatment' differences were caused by fertiliser.

**Unless you're doing a variety trial, use only one crop variety in your trial.**

Another common problem is the comparison of fertiliser types that also include changes in rate, and vice versa. A typical example would be a trial designed to see whether P was required as a starter fertiliser. The 'plus P' treatment would often be applied as MAP or 12:10:10 and the 'minus P' treatment would receive no starter fertiliser. This (common) example isn't really a comparison of plus and minus P because the amounts of nitrogen and potassium also differ between the treatments. In this sort of trial, you can only conclude that the starter fertiliser (whatever it was) was or wasn't different from no starter fertiliser. You can't tell which nutrient did or didn't do the work and, consequently, can't really talk about the effect of P, N or K.

**If you're doing a comparison of fertiliser types, be certain that the rates of all of the nutrients are the same.**

In addition, trials like this are often worthless because visual early season responses to a treatment (such as starter fertiliser) are often assumed, without justification, to have translated to a yield response 4 or 5 months later. Lots can happen to a crop during the season, so you can't assume anything about yield from the appearance of the crop early on. If you want to know about yield, measure it!

**If you're doing a trial to look at yield responses, you must measure yield at the business end of the year. Mid-season results aren't where the money's at.**

### **11.5 Treatment comparisons where one treatment isn't managed properly.**

This is a common and understandable problem in trials that are comparing a 'new' with an 'old' technique. It's also a difficult problem to solve, because the trial of the new treatment usually involves two distinct functions. First, you want to see if it's better than the old technique but, second, you need to learn how to use the new method effectively. In many ways, the old treatment has an unfair advantage, because you already know how to get the best from it. This problem often leads to good ideas being left on the shelf – a waste of profit potential.

There's no simple solution to this problem, but being aware of it is a good start! If you're trying something new, speak with people who've already had success with it, as well as those who've had failures. It'll help you to piece together a plan that best matches your situation. In addition, it may be worth persisting for more than the usual season or two. This will give you time to get better with the new method and will increase the chances of it getting a fair run against your 'tried and true' technology.

New treatments that involve changes in more than one aspect of crop management often suffer from this problem. Reduced tillage, weed control using stale seedbeds, and integrated pest and disease management are techniques that have suffered from 'unfair' comparison with existing practices.

### **11.6 Common design problems**

Hopefully the previous sections will have helped to iron out some kinks in the application of treatments to trials. The following section is intended to highlight common problems arising from incorrect arrangement of treatments on the trial site. In each case, simple solutions are presented.

The blocking used in Fig. 11 is incorrect because it makes a path that goes across the variation of the site. Because it doesn't travel along the trend in the paddock (from poor to better soil) each block contains a mix of soil types, from the very worst to the very best. This almost completely defeats the purpose of blocking, because

each head-to-head treatment comparison (of A vs. B) within these blocks will occur under very different conditions.

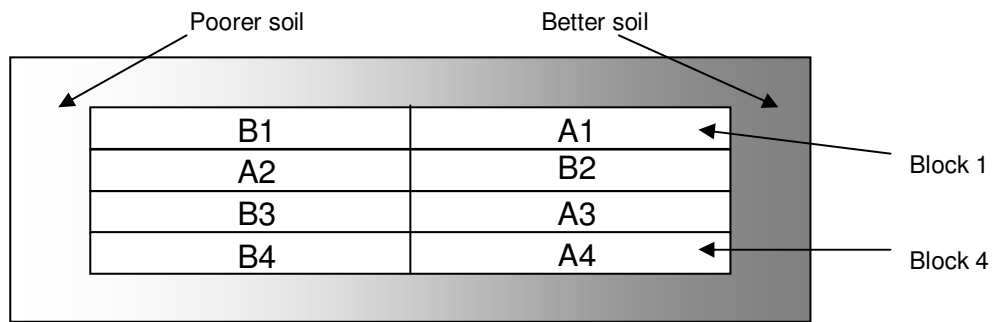


Fig. 11. An example of really poor blocking – the blocks are arranged to make a path that leads across the site variation, so each block contains a wide mixture of soil types

The blocking used in Fig. 12 is incorrect because it also makes a path that goes across the variation of the site. Because it doesn't travel along the trend in the paddock (from most to least sheltered) each block contains a mix of environments.

By the same token, if the shelter belt were instead at the top of a hill that sloped down to the bottom of the diagram, the blocking would be incorrect for much the same reason.

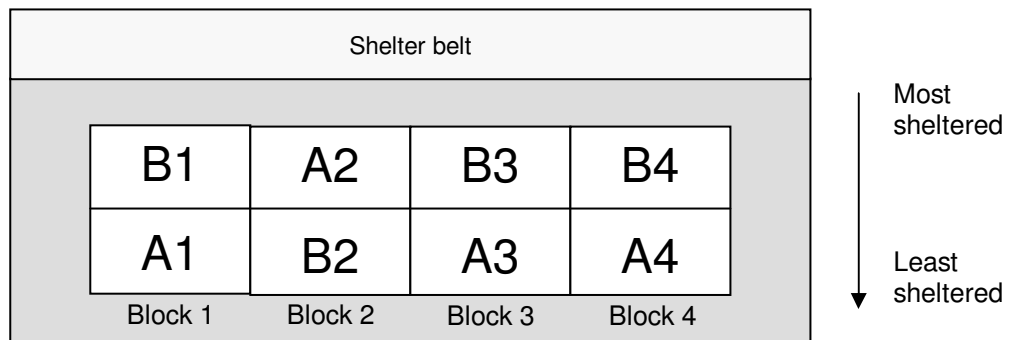
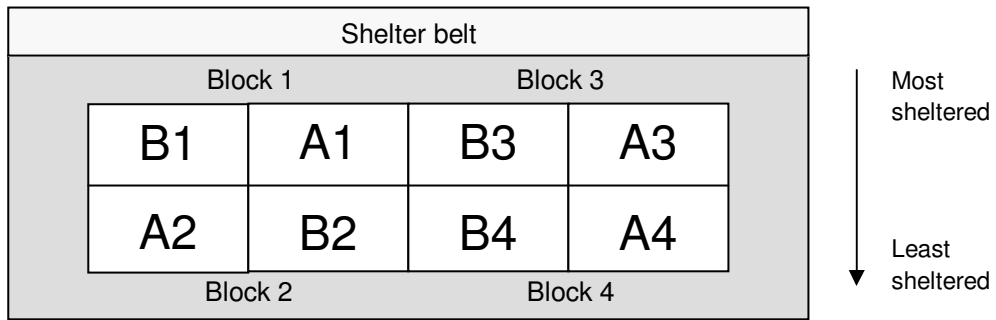


Fig. 12. An example of poor blocking – the blocks are arranged to make a path that leads across the site variation, so each block contains a mixture of environments

The error made in the design above has been fixed without the need to fiddle around with plot dimensions or general layout (Fig. 13). We simply took the existing plot layout (4 plots across and 2 down) and reassigned the blocks so that each block runs horizontally (rather than vertically). We then re-assigned treatments within each

block, using the same randomisation as before (B-A in block 1, A-B in block 2, B-A in block 3 and B-A in block 4).



*Fig. 13. An example of good blocking – the blocks are arranged to make a path that leads along the site variation, so that in each block two plots are compared in the same environment*

No doubt you'll be aware of other common pitfalls in trial design. Keep an eye out for them, and consider offering gratuitous advice to help 'offenders' mend their ways.

#### **A final word**

*We hope that you've found the Guide useful and interesting. Most of all, we hope that it helps you to improve your farm business.*

*If you think that there's something missing, or that there are issues that need clarification, please do not hesitate to contact the authors. As trial professionals, they are as interested in helping you to run top-notch trials of your own design on your own farm.*

**Happy trialling!**