The DOOR Manual for Plant Nurseries

Reprint – information current in 1996

REPRINT INFORMATION – PLEASE READ!

For updated information please call 13 25 23 or visit the website www.deedi.qld.gov.au

This publication has been reprinted as a digital book without any changes to the content published in 1996. We advise readers to take particular note of the areas most likely to be out-of-date and so requiring further research:

- Contacts—many of the contact details may have changed and there could be several new contacts available. The industry organisation may be able to assist you to find the information or services you require.
- Organisation names—most government agencies referred to in this publication have had name changes. Contact the Business Information Centre on 13 25 23 or the industry organisation to find out the current name and contact details for these agencies.
- Additional information—many other sources of information are now available. Contact an agronomist, Business Information Centre on 13 25 23 or the industry organisation for other suggested reading.

Even with these limitations we believe this information kit provides important and valuable information for intending and existing growers.

This publication was last revised in 1996. The information is not current and the accuracy of the information cannot be guaranteed by the State of Queensland.

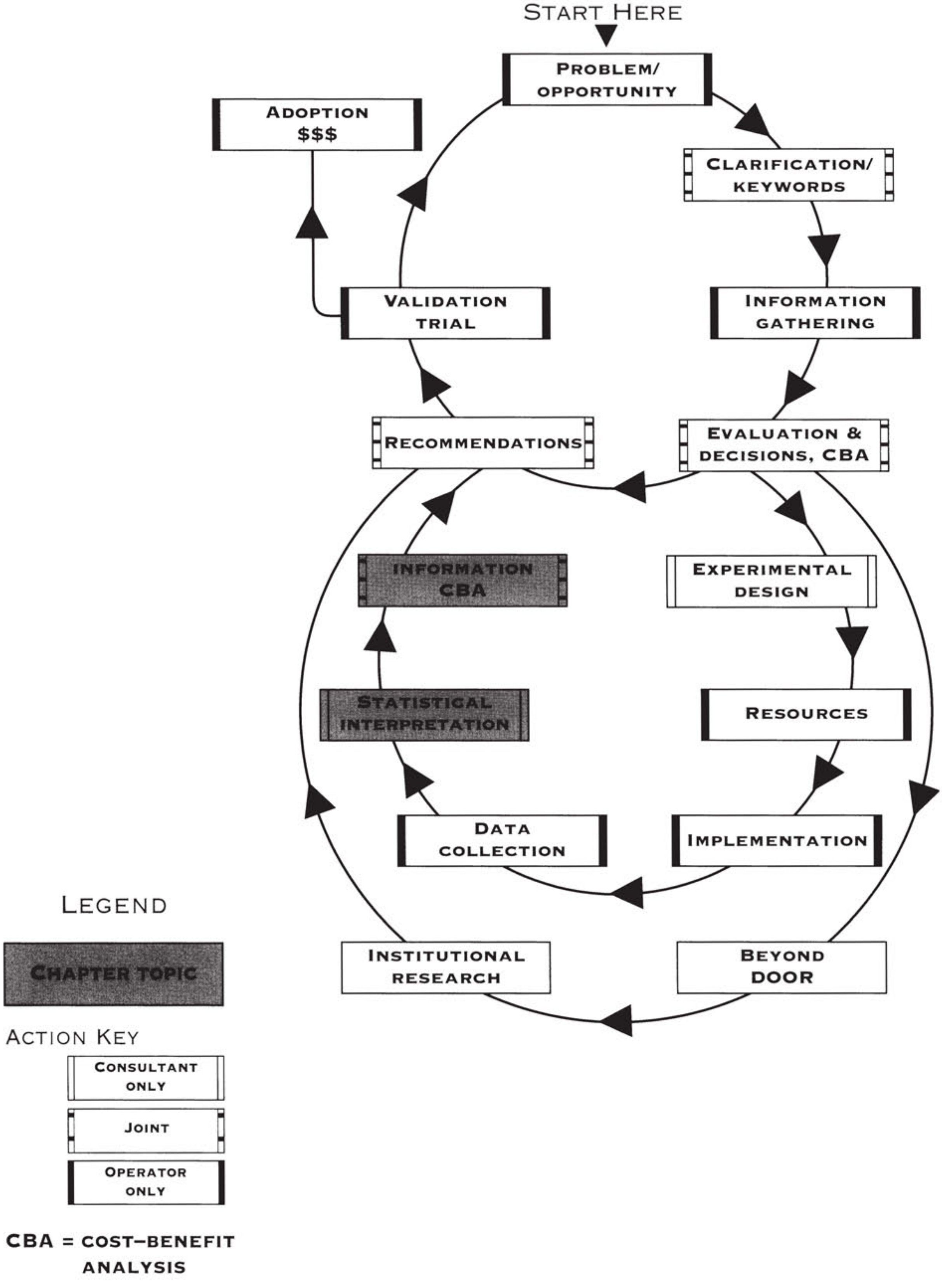
Queensland Government

Data Interpretation

J. Giles

7

DOOR
Implementation Cycle



7.1 INTRODUCTION

Most experiments set out to compare the relative performance of different treatments with the aim of finding treatments that, on average, produce the tallest plants (or the heaviest, or healthiest, etc.).

This chapter discusses how to compare average values and briefly refers to experiments where the aim is to define the response of plants to varying levels of some factor, e.g. fertiliser added.

Operators need not have a full understanding of statistics to implement DOOR since this is the consultant's area of expertise. Following the simple statistical rules given below while conducting research will suffice.

- All treatments being tested must be randomly distributed in a specified area, ensuring that none receives any preferential management or particular condition other than from the treatment itself.
- Each treatment has to be replicated enough times to meet the required level of precision.
- Treatments need to be grouped together in blocks where environmental gradients exist (slope, temperature, light, wind).
- Base data on objective assessments (e.g. height, width, weight) rather that on subjective judgements (e.g. plant vigour, appeal, colour).

7.1 INTRODUCTION

- Most experiments set out to determine the best treatment under certain conditions.
- An understanding of statistics is helpful for DOOR operators but not essential. Following the basic statistical rules is good enough.

7.2 STATISTICAL COMPARISONS

At this stage of the DOOR cycle, statistics has two important roles.

The first is to summarise the measurements made. This is usually in the form of an average value (mean) and a measure of how variable the data are.

The second important role is to carry out tests of significance to answer questions like "Does one treatment produce taller plants than another treatment?".

7.2 STATISTICAL COMPARISONS

- The means of the treatments are compared using tests of significance.

7.2.1 COMPARISON OF TREATMENT MEANS

Generally, the basis of a test of significance is the comparison of the difference between average values obtained for treatments with the experiment's variability. Big differences between treatment averages (means) and relatively low variability indicate that there are real treatment differences.

Consider two scenarios. In the first we have two treatments with average heights of 20 cm (range of 10 values between 17 and 23 cm) and 30 cm (range of 10 values between 27 and 33 cm). In the second we have two treatments with average heights of 20 cm (range of 10 values between 17 and 23 cm) and 22 cm (range of 10 values between 19 and 25 cm). We are more easily convinced that one treatment produces taller plants than the other in the first case. It is the bigger difference between means with the same amount of variability (as measured by the range of values in the experiment) that is convincing.

7.2.1 COMPARISON OF TREATMENT MEANS

- Tests of significance compare the size of the treatment differences with the experiment's variability.
- Tests of significance are usually carried out at probability levels of 0.05 or 0.01.
- When testing at $P = 0.05$, the probability of the significant difference between treatments happening by chance is less than five in 100.

Then consider another two situations. In the first we have two treatments with average heights of 20 cm (range of 10 values between 17 and 23 cm) and 30 cm (range of 10 values between 27 and 33 cm). In the second we have two treatments with average heights of 20 cm(range of 10 values between 10 cm and 30 cm) and 30 cm (range of 10 values between 20 cm and 40 cm). We are more convinced that one treatment produces taller plants than the other treatment in the first case. The difference between means is 10 cm in both cases. However, in the first case, the two treatments are separated into two distinct groups (treatment A ranges from 17 to 23 cm and treatment B varies between 27 and 33 cm) whereas in the second case there is a lot of crossover, with both treatments having plants in the range 20 to 30 cm. The variability in the second case is much higher.

Thus the test of significance considers both the differences between treatment means found in the experiment, and the variability of the plants with the same treatment. Instead of using range as a measure of variability we use something that is less influenced by extreme values. But the principle remains the same.

Using this test, you can calculate the probability of experimental differences occurring just by chance if the treatments do not, on average, have different heights.This sort of thing happens when one treatment, by chance, was allocated to all the plants with the most potential and was positioned in the most favourable locations, etc.

Thus, a test of significance can help you be at least 95 per cent sure that treatment A produces taller plants than treatment B. This is usually stated as,"treatment A produces taller plants than treatment B ($P<0.05$)" in scientific papers. The bracketed probability just tells us that the probability of the statement being incorrect is less than five times in 100 experiments or 0.05. It is conventional to select this probability or 0.01 (once in 100 experiments).

7.2.2

Comparison of two treatment means

- The simplest type of experiment compares two treatment means with no blocking.
- Consider practical significance as well as statistical significance.

7.2.2

Comparison of two treatment means

The simplest type of experiment comparing means (average values) is the experiment that tests two treatments with no blocking.

For example, an experiment is set up to test whether increasing the air-filled porosity of the growing medium by adding 20 per cent coco peat gives better or worse growth of calatheas. The two treatments are standard medium and standard medium with 20 per cent coco peat added; 48 pots are allocated to each treatment. For this example, the measurement is the number of shoots per pot after 2 months. For each treatment the number of pots with zero to five shoots is given in the table below.

Table 7.1 Number of shoots of *Calatheas* produced in standard potting medium and in standard medium to which coco peat had been added (20 per cent)

Shoot number	Standard media	Plus 20% coco peat
0	2	3
1	14	7
2	13	9
3	8	18
4	11	8
5	0	3
Mean	2.25	2.63

A comparison of the means shows the coco peat mix produces, on average, slightly more shoots than the standard medium (2.63 compared with 2.25). However, there is a lot of variability with the standard mix producing anything from 0 to 4 shoots and the coco peat mix varying from 0 to 5 shoots.

There is another aspect to statistical testing: practical significance. A statistically significant difference is not relevant if the estimated difference is too small to be practically important. Conversely, an estimated difference that is not statistically significant, but is practically worthwhile, deserves consideration. Lack of statistical significance can result from little real treatment effect or from the experiment's lack of refinement. Thus further experimentation may be justified. An improved experiment, with more replication and greater control of variable factors, might confirm a real treatment effect.

In our example, if the increase of 0.38 shoots, on average, is an important one, we could repeat the experiment with greater replication and more control of variability in an attempt to prove a significant difference between treatments.

The statistical test used here is the 't' test for independent samples. Applied to the example above, the test obtains a 't' statistic value of 1.46 by dividing the difference between treatment means (2.63 - 2.25 = 0.38) by a measure of variability (the standard error of difference). Using statistical tables, you would find that, for significance at P = 0.05, a 't' value of 1.99 or higher is needed. Thus there is not a significant difference in the number of shoots produced by the two treatments.

7.3 Analysis of variance

When more than two treatment means are being compared, analysis of variance (ANOVA) tests whether there are significant differences between treatments. Testing is a two-stage procedure. Analysis of variance tests whether there are significant differences within the group of treatments. Multiple comparison procedures are then used to find which pairs of treatment means are significantly different.

7.3 Analysis of variance

- Use analysis of variance (ANOVA) when more that two treatments are being compared.
- ANOVA tests whether there are significant differences within the group of treatments.
- Using multiple comparison procedures, pairs of treatment means are tested for significant differences.

7.3.1
Completely randomised design

- ANOVA in a completely randomised design separates the variability between treatments from natural variability or error.
- Error is the base value against which treatment effects are compared.
- If the variability between the treatments is significantly larger than the error term, there are differences between the treatment means.
- To determine which treatments are significantly different from one another, use the least significant difference test (LSD).

7.3.2
Randomised block design

- ANOVA for randomised block design is similar to that for the completely randomised design except that the variability in the experiment is split into three sources instead of two: treatments, blocks and error.
- After doing an ANOVA, do an LSD test.
- LSD is calculated using the error mean square and the number of values used to calculate each mean.
- The LSD is the smallest difference between treatment means which will give a significant difference at the probability level chosen.
- Any two means that differ by more than the LSD value are significantly different.
- Assess the block means to look at the effectiveness of blocking.

7.3.1
Completely randomised design

The analysis of variance partitions the variability in the experiment into its various causes. In the case where there are different treatments but no blocking (known as a completely randomised design), the variability is divided into that caused by treatment differences and what is left over. The part left over, or unexplained variability, also known as error, gives us a base value against which treatment effects are compared. It is a measure of the natural variability between plants and, though commonly called the error term, has nothing to do with mistakes.

If the variability between treatments is much larger than the error term (variability within treatments in this instance) then we conclude that there are differences between treatments. We can attach a probability of error in making this statement, just as we did for the 't' test.

Having established that there are significant differences between treatments, the next step is to define where the differences are. We do this using multiple comparison procedures. There are a number of these to choose from. We will only consider the least significant difference (LSD) procedure.

7.3.2
Randomised block design

The analysis of variance for a randomised block design is similar to that for a completely randomised design except that the experiment has three sources of variation instead of two: treatment differences, block (or replicate) differences, and the residual, (again known as error).

For example, an experiment compares five different pot insulation treatments on the growth of murrayas:

1. insulate continuously
2. insulate from February to April, then remove
3. insulate from May only
4. no insulation at all
5. improved insulation

The 40 pots were allocated to eight blocks, with block 1 on the western edge of the experiment through to block 8 on the eastern edge. Table 7.3 shows the collected growth data (as estimated by height multiplied by width).

Table 7.2 Data for the randomised block design experiment investigating the effect of insulation on plant growth. Collected at the end of June.

		TREATMENTS				
		1	2	3	4	5
	Height (cm)	62	73	68	70	70
BLOCK 1	Width (cm)	44	46	67	51	53
	Size (cm^2)	2728	3358	4556	3570	3710
	Height (cm)	70	70	71	80	60
BLOCK 2	Width (cm)	46	50	54	58	44
	Size (cm^2)	3220	3500	3834	4640	2640
	Height (cm)	61	65	71	63	64
BLOCK 3	Width (cm)	45	45	57	45	52
	Size (cm^2)	2745	2925	4047	2835	3328
	Height (cm)	61	64	59	63	51
BLOCK 4	Width (cm)	46	60	51	50	47
	Size (cm^2)	2806	3840	3009	3150	2397
	Height (cm)	65	58	59	75	58
BLOCK 5	Width (cm)	54	50	50	60	43
	Size (cm^2)	3510	2900	2950	4500	2494
	Height (cm)	56	62	59	63	60
BLOCK 6	Width (cm)	42	42	54	54	40
	Size (cm^2)	2352	2604	3186	3402	2400
	Height (cm)	64	67	65	55	62
BLOCK 7	Width (cm)	45	45	58	51	50
	Size (cm^2)	2880	3015	3770	2805	3100
	Height (cm)	57	69	69	60	49
BLOCK 8	Width (cm)	45	50	50	50	41
	Size (cm^2)	2565	3450	3450	3000	2009

The analysis of variance table for size is as follows:

Table 7.3 ANOVA table for data shown in table 7.2

Source of variation	Degrees of freedom	Sum of squares	Mean square	F	Probability
Treatments	4	4 453 560	1 113 390	4.33	0.0075**
Blocks	7	2 898 165	414 024	1.61	0.1736
Error	28	7 199 402	257 122		
Total	39	14 551 127			

If you are not familiar with this method, don't worry too much about all the figures in the table. Concentrate on the column labelled probability.

Table 7.3 shows significant differences between treatments. The probability 0.0075 (0.75 per cent) gives us the likelihood of this statement being incorrect. Two asterisks (**) in the probability column shows significance when testing at P = 0.01.

Although there is no significant difference between blocks when testing at P = 0.05, they would differ if tested at any level above 0.1736 (17.4 per cent). Accept this as an indication that blocking might be effective in this case.

The next step in the analysis is to establish which pairs of treatment means are significantly different. First, calculate the treatment means and then use the LSD test.

These are the treatment means (in descending order):

3.	insulate from May only	3 600 a
4.	no insulation at all	3 488 a
2.	insulate from February to April, then remove	3 199 ab
1.	insulate continuously	2 851 b
5.	improved insulation	2 760 b

From the means, we can see that the biggest plants were produced by the two treatments that had no insulation between February and May and the smallest plants came from the treatments with improved insulation.

Calculate the LSD by using the experiment's variability (the error mean square, in this example 257 122) and the number of values used to calculate each mean (the number of blocks). It is the smallest difference between treatment means which will give a significant difference at the probability level chosen.

If we choose to test at a probability level of 0.05, then the LSD for this example is 519. The actual calculation, in which the t value at 0.05 is 2.047, is given below:

$$2.047 \quad x \sqrt{\frac{2 \text{ x } 257122}{8}} = 519$$

Any two means which differ by more than 519 are significantly different (P<0. 05). We find that means 3 and 4 are significantly larger than means 1 and 5.That is, treatments receiving no insulation from February to May produced bigger plants than treatments insulated continuously (either with standard or improved insulation).

These significant differences are often represented using a lettering system next to the means as shown above. Such a table would be accompanied by a statement like,"Means not followed by a common letter are significantly different (P<0.05)". Thus treatments 3 and 4 (followed by an "a") are significantly different from 1 and 5 (followed only by "a""b") but they are not significantly different from treatment 2 because all of treatments 3, 4 and 2 are followed by the common letter a.You will find this sort of lettering in many tables that appear in scientific papers.

To plan future experiments, assess the effectiveness of the blocking method by examining the block means. The block means for this experiment are given below:

Block 1	3 584	(western edge)
Block 2	3 567	
Block 3	3 176	
Block 4	3 040	
Block 5	3 271	
Block 6	2 789	
Block 7	3 114	
Block 8	2 895	(eastern edge)

This shows that blocks 1 and 2 on the western edge of the experiment produced the biggest plants.This suggests that there is a significant enviromental gradient that is accounted for by blocking.

7.3.3
FACTORIAL EXPERIMENT

Now consider a factorial experiment in a randomised block design.

This experiment examines the effect of varying proportions of peat, sand and pine bark fines in potting mix, with or without nitrogen applied in the irrigation water, on the growth of marigold seedlings.

There were 16 treatments, consisting of three factors. Two factors related to the composition of the potting mix: proportion of peat (levels of 0, 10, 20 and 30 per cent), and proportion of sand (levels of 10 and 30 per cent). The remainder is pine bark fines. The final factor is presence or absence of nitrogen in the irrigation water.

Three replicates of each treatment used 48 pots in all. The replicates corresponded to the blocks.

The growth and health of the marigolds were assessed in a number of ways. Here we will only consider one variable: fresh weight of shoots after 36 days. The values for fresh weight of shoots are listed below in table 7.4.

7.3.3
FACTORIAL EXPERIMENT

- Factorial experiments test two or more treatment types or factors in the one experiment.
- The ANOVA partitions the variability due to the treatments into that caused by each of the factors as well as the interactions between each pair of factors.
- Look for significant interactions.
- Graph the means of the interactions to determine the trends.
- When a treatment factor consists of various levels of application, examine the response to increasing levels rather than just look at tables of means.
- Plot the points to determine the curve and the type of fit.

Table 7.4 Fresh weight of marigolds grown in different media compositions

		Fresh weight of shoot (g)					
Blocks		1		2		3	
Nitrogen		No	Yes	No	Yes	No	Yes
Treatments (% composition)							
Peat 0	Sand 10	51.0	58.2	40.1	33.1	50.8	89.9
Peat 0	Sand 30	38.9	36.6	46.6	45.2	46.1	60.4
Peat 10	Sand 10	58.9	55.6	49.4	58.0	55.3	69.1
Peat 10	Sand 30	37.1	33.0	52.5	45.4	72.3	61.6
Peat 20	Sand 10	62.5	69.9	58.5	62.4	76.7	81.6
Peat 20	Sand 30	50.6	40.5	47.1	59.0	67.6	59.4
Peat 30	Sand 10	64.8	59.2	57.0	79.4	77.4	97.2
Peat 30	Sand 30	55.9	53.4	66.0	59.8	75.2	60.6

The analysis of variance partitions the variability in the experiment into that caused by treatments, that caused by blocks and the residual (known as error). The treatment variability is further split into the effect of peat composition, the effect of sand composition, the effect of nitrogen addition and the various interactions between these factors. Remember the interaction between two factors is a measure of the extent of difference in response to one factor at varying values of the other factor. For example, the difference between how fresh weight of marigolds changes as the proportion of peat in the mix varies in the presence or absence of nitrogen. This is the peat by nitrogen interaction. The analysis of variance table is given below.

Table 7.5 ANOVA for data shown in table 7.4

Source of variation	[1] Degrees of freedom	Sums of squares	Mean square	F
Blocks	2	2 816.97	1 408.48	17.69**
Peat	3	2 147.15	715.72	8.99**
Sand	1	1 252.56	1 252.56	15.73**
Peat*Sand	3	93.22	31.07	0.39
Nitrogen	1	102.67	102.67	1.29
Peat*Nitrogen	3	138.23	42.74	0.54
Sand*Nitrogen	1	482.60	482.60	6.06*
Peat*Sand*Nitrogen	3	67.69	22.56	0.28
Error	30	2 389.01	79.63	
Total	47	9 480.10		

[1] Degrees of freedom (df) is the number of possible comparisons that can be made between a treatment (or blocks) and all others. For example, one level of peat can be compared with the other three levels (four levels in total) and thus has a df of 3.

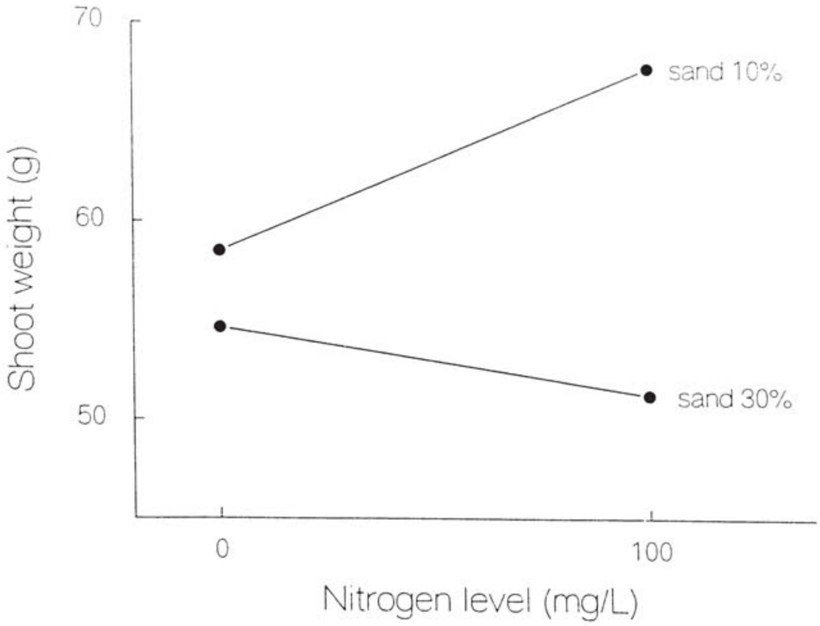
For a factorial experiment, the first step in interpreting the analysis of variance table is to look for significant interactions. In this example the sand by nitrogen interaction is significant. This tells us that the response to nitrogen depends on which level of sand is present. Thus it is not correct to look at the average effect of nitrogen or the average effect of sand. Each depends on the level of the other. Even though the sand effect is significant, we will not consider the means for the two sand levels. Instead we will examine the four means consisting of two sand proportions, each with and without nitrogen. Other effects that are significant are the effect of peat and of blocks. So the analysis of variance table for a factorial experiment tells us which table of means to study.

First we will look at the sand by nitrogen interaction. The table of means follows in table 7.6.

Table 7.6

Treatments	Nitrogen	
	No	Yes
Sand 10%	58.53	67.80
Sand 30%	54.66	51.24

The trends here are more obvious if the means are graphed as below.



The LSD for comparing these four means is 7.44. Sand 10 per cent, with nitrogen, produces larger shoot fresh weight than the other three treatments. Another way of expressing this is that, with 10 per cent sand, there is a response to nitrogen, but, with 30 per cent sand, there is not.

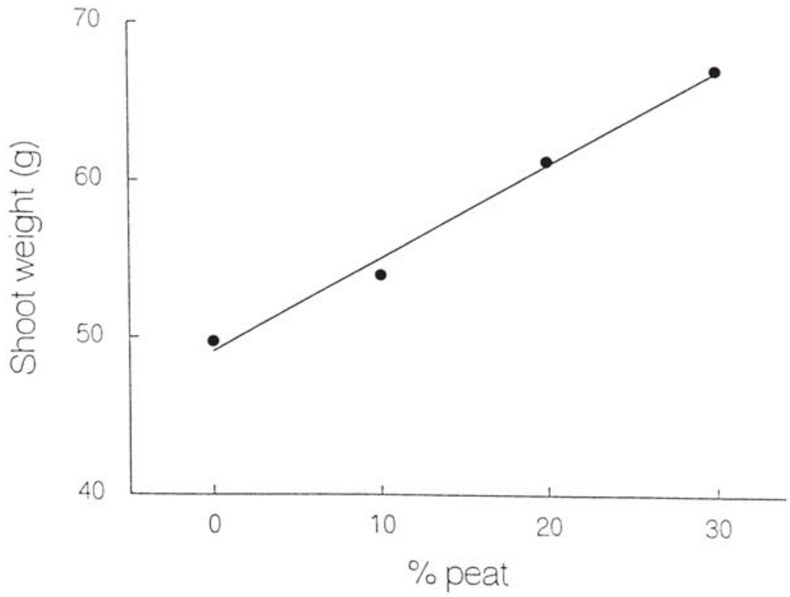
The next table of means is the effect of different peat levels. The means are listed next.

Table 7.7

Peat	0%	49.74
Peat	10%	54.02
Peat	20%	61.32
Peat	30%	67.16

The LSD for comparing the means when using a probability level of P = 0.05 is 7.17. This shows that shoot weight for 20 per cent and 30 per cent peat is significantly higher than for 0 per cent or 10 per cent peat. However, this is not a logical way to look at the response to peat.

The appropriate approach is to plot the response and fit a curve to it so that fresh weight can be predicted for any level of peat between 0-30 per cent. In this case a straight line provides a very good fit, as illustrated below.



The final table of means is for blocks.

Table 7.8

Block 1	51.63
Block 2	53.72
Block 3	68.83

The LSD for comparing blocks (P=0.05) is 6.21. Block 3 produced larger shoot weights than the other two blocks. This may have been due to the effects of enviromental factors such as shade.

7.4 RESPONSE CURVES

In some experiments the main aim is to examine the response of one variable to differing levels of another variable, for example, the response of yield to increasing levels of fertiliser. In this case, design the experiment with many levels of fertiliser (usually equally spaced). If the response is expected to increase with higher levels of fertiliser up to a certain point and then decrease with larger applications, try to plan the fertiliser levels in the experiment so that the amount required to produce this optimum is exceeded. A curve could then be fitted which will estimate this optimum. It would also be useful to estimate the fertiliser level which results in maximum profits.

7.5 INTERPRETATION OF RESULTS

Assess results from all DOOR experiments in the context of information gained in the earlier information-gathering phase. However, ensure that this information is up-to-date by having a keyword search carried out for the period since your initial search.

7.5.1 SUPPORTIVE

Results may indicate that the work is on the right track but arriving at exactly the same result as reported elsewhere would be unlikely, simply because the detail of the experiment itself would probably have varied slightly from the previously reported work.

7.5.2 NOT SUPPORTIVE

When the new results do not support information generated elsewhere, explore why this has happened. In most cases there will be some factor sufficiently different to account for the unexpected result, or it could be due a chance occurrence and should not be accepted at face value. Confirm the result in a follow-up experiment.

7.6 COST-BENEFIT ANALYSIS

Do a cost-benefit analysis as shown in chapter 4 to compare the economics of the old practice with the new. Chapter 8 discusses how to analyse all the implications of a potential new practice.

Do not overlook the value of new knowledge gained in the conduct of the research, irrespective of whether or not the new treatment was worthwhile. Such knowledge should enhance your ability to make decisions.

7.5 INTERPRETATION OF RESULTS

- Interpret the results of the data analysis in the context of existing information.
- If the results do not support the existing literature, explain why not.

7.6 COST-BENEFIT ANALYSIS

- Compare the economies of the old way with those of the new.