EXPERIMENTAL DESIGN IN CATTLE RESEARCH WHEN RESOURCES ARE LIMITING

INTRODUCTION

G.W. BLIGHT*

In conventional agronomic experiments with pasture or crops grown in plots the cost per experimental unit is low compared with, say, animal experiments. The capital cost involved in establishing a well designed large-scale field trial with cattle can easily exceed available resources. Consequently the replication rates of paddocks and animals per treatment may not be as high as we would like, and the treatment responses we can expect to detect as statistically significant in the experiment may be larger than desired.

Adequate replication rates (number of animals per treatment) are likely to be attainable in grazing trials where, because of the nature of the treatments, all of the experimental animals graze a common pasture. However, in multi-paddock grazing trials such as stocking rate or pasture comparisons, where it is the paddock which is the experimental unit, cost often prohibits setting up the necessary numbers of paddocks of sufficient size. In this contract the papers by Haydock, Duncalfe, Pepper and Mayer are concerned with the design of grazing experiments where constraints on resources are limiting opportunities to conduct research; Seebeck's paper reviews the additional design requirements when body composition information, measured at slaughter, is required from grazing experiments. In the last paper of this contract O'Rourke develops guidelines for the design of survival feeding experiments in pens.

CONSEQUENCES ARISING WHEN THE DESIGN OF A GRAZING TRIAL IS NOT IDEAL

K.P. HAYDOCK**

Grazing trials can be very expensive requiring finance to (i) acquire animals (ii) purchase and install fencing and watering points (iii) supply seed, plant and establish pastures (iv) control livestock diseases and parasites (v) build mustering yards and (vi) buy scales to weigh animals. Large resources of land are required for even a simple trial One often finds that in order to fit in with budgetary constraints the size of a trial needs to be curtailed from that which was considered ideal.

In this paper, methods of size reduction are discussed together with some of the consequences. Some suggestions are made to obtain an estimate of error and to aid the extrapolation of the results.

METHODS OF SIZE REDUCTION

(i) Herd size One compromise is to reduce the herd size. A minimum herd size generally is considered to be three; any aberrant behaviour of one or two animals may be detected by comparing the behaviour of the three. Single animal herds have been used successfully (Yates et al. 1964 ; Bryan 1968; Jones 1982). It is considered essential that the paddock sizes be small so that several animals are in view of each other in order for behaviour patterns to be normal (Mannetje et al. 1976). This approach is more relevant where the primary objective is to assess the effect of grazing on pastures. With single animals, any sickness can

* QDPI Biometry Branch, Rockhampton, Qld 4700.
** CSIRO Division of Mathematics and Statistics, St Lucia, Qld 4067.

have a serious effect on animal production.

The reduction in paddock size will decrease the precision of the trial for two reasons. Firstly, the variation between paddocks increases as the size of the paddock decreases (Smith 1938; Hatheway and Williams 1958). Secondly, the variance of mean liveweight change/paddock increases as the number of animals/paddock is reduced.

(ii) Reduced replication Another compromise is to reduce the number of replicates. This also reduces the precision of a treatment mean. In a region which is variable with respect to factors affecting pasture growth, using a small number of replicates would give wide confidence limits for a treatment effect. Restricting the region across which a trial is established would improve precision but would limit the applicability of the results. A preferable procedure would be to develop a regression model relating the variate of interest, e.g. liveweight gain or milk yield/ha, to the factors of the trial and to growth factors associated with each paddock. Factors affecting growth which. easily come to mind are rainfall (amount and pattern), temperature (maximum and minimum), soil moisture, botanical and chemical composition of the pasture, aspect of the paddock (northerly or southerly)) soil type and depth of soil.

Such a model is of inherent biological interest and would have general applicability. In its simplest form it amounts to a covariance adjustment for diverse paddock effects. For this purpose we should note that pasture production is affected by rainfall, temperature, paddock aspect, soil moisture, soil type and depth. The dry weight of green material has been shown to affect liveweight change (Willoughby 1959; Roe et al. 1959; Yates et al. 1964; Mannetje 1974). Similarly, chemical composition of a pasture is influenced by temperature, soil moisture, botanical composition and soil fertility (Andrew and Fergus 1976; Ludlow 1976). Howard et al. (1962) showed that mineral composition of a pasture was correlated with protein content. They used chemical composition to explain the variation in animal production over seasons of the year. Thus, a simple form of the model suggested would involve only animal production, dry weight of green material and protein content of the pasture.

A reduction in the size of an experiment also may be achieved by reducing the replications of some treatments. The greatest replication would be maintained for the treatment of most interest. For example, if the purpose of a trial was to determine the optimum stocking rate for a number of pastures, the stocking rates initially judged to be near the optimum could have the most replication.

(iii) No replication In any grazing trial there is a need for an unbiased estimate of random error. A trial without replication usually would preclude the estimation of an error term. One exception is where a treatment (say A) has quantitative levels. Provided three or more levels are used, one can fit a suitable response model to the levels of A e.g. linear. Orthogonal polynomials should be used so that independent sums of squares for the parameters of the model can be calculated. The deviation from the response model (DEV) may be used as an estimate of error. If other factors are also being tested in the trial, the error term could be composed of DEV and interactions of the form DEV X factor, after testing for homogeneity.

Error degrees of freedom usually are few and they may be increased by transforming the levels of the factor to linearize the response. In this way only one degree of freedom is assigned to the factor A effect and any interaction effects involving A will be adequately tested. If a polynomial is not a biologically meaningful model, one may fit an appropriate model to the response

after determining the reality of any interactions,

For unreplicated grazing trials in which a paddock is the experimental unit, an analysis of variance is possible using the variation between animals within paddocks as an error term. However, the effects of treatments and of paddocks are confounded; also it is difficult to extrapolate the results to some population as most research workers would desire to do.

Unreplicated experiments usually are used in the early stages of an investigation. Exceptions to this were the work of Shaw (1961) and Shaw and Mannetje (1970). These workers compared animal production from native pasture with that from native pasture and Townsville lucerne with and without fertilizer. Five fold increases in production/ha were obtained and the results were enhanced by running the trial for seven years using replacement animals each year. There is little doubt that the results were accepted as real because the change in production was large, biologically meaningful and consistent over seven years with different samples of animals. We might say that the magnitude of the difference was greater than any which could arise by chance due to any conceivable random effects of paddocks and animals. Where the differences are much smaller, the problem reduces to one of eliminating the confounding effects of paddocks and of having an estimate of the underlying random error.

If the regression model relating animal production to environmental and growth factors has been developed, it could be used to adjust animal production for the confounded paddock effects. An estimate of error must be composed of two components, one due to paddocks and one due to animals within paddocks. A component due to paddocks might be estimated by collecting information on the variance component for paddocks from experiments reported in the literature and from unpublished results. If these components have a narrow distribution one could use with confidence their weighted mean as an estimate of paddock variance; if they exhibit a wide range, their values might be associated with paddock size, as found by Smith (1938) for dry matter production of different crops. As well, the paddock variance components might be associated with soil type, region etc. If this is the case, an estimate of the variance component for paddocks, for the appropriate conditions of a particular trial, could be made. A component of variance for animals also should be estimated. In this regard, the standard deviations/animal for cumulative liveweight change, in lb/ac, for one year old animals after periods of four weeks, produced by Haydock (1964) from data gathered by the late M.C. Franklin, should prove useful.

Even if these efforts are successful in constructing an error variance, there would be many research workers who would prefer to compare treatments using information based on genuine replication.

PADDOCK ROTATION AND REPLICATION RATES FOR DAIRY PRODUCTION EXPERIMENTS

F. DUNCALFE*

BACKGROUND

This paper examines the design of long-term dairy production experiments in which paddocks, each of which is subjected to a specific pasture treatment, are grazed for a whole lactation by dairy cows. We show that the resource requirements of such experiments are very large when the experimental unit is taken to be the herd of animals and its associated paddock. In many cases the

* QDPI Biometry Branch, Brisbane, Qld 4001.

requirements would limit research opportunities.

In a typical experimental set-up, treatments are assigned to paddocks and a herd of lactating cows is allocated to each paddock. During the course of the experiment each herd of cows has to be taken to the dairy, milked and then returned to the same paddock. Measurements related to dairy production, such as milk production, butter fat yield and solids-not-fat, are made for each animal in each herd. Herds of five cows are commonly used by QDPI officers to simulate a commercial herd, and to minimize possible within herd correlation. An experiment designed to compare eight treatments, with a minimal replication rate of two paddocks per treatment, will have sixteen herds of animals to be handled twice daily seven days a week for the duration of the experiment, Thus even minimal replication and a herd size of five produces an experiment with large labour requirements.

A normal management practice used by dairy farmers and also in some experiments on irrigated improved pastures, involves the rotation of a herd of animals sequentially through a series of paddocks. For example, with a rotation sequence of one week in and three weeks out for each paddock in turn, a herd of animals becomes associated with a set of four "sub-paddocks", each being one quarter of the size of the original paddocks. While paddock rotation may involve a lot of extra fencing and extra work in the planning stages, it offers a reduction in size of the error mean square, and thus improved power in significance tests, through extra blocking and randomization. For eight treatments and four rotations, there are 32 sub-paddocks for each replicate. For two replicates, rotational grazing involves blocking the 64 sub-paddocks into eight homogeneous "sub-blocks", each of eight sub-paddocks. Since the sub-block is one quarter of the size of blocks in an experiment without rotation, the experimenter may be able to take advantage of the smaller size to create more uniform sub-blocks. In order to be able to block effectively, some knowledge of the expected productivity of different parts of the available land is required. Even when such information is not available, the extra randomization in allocating treatments to sub-paddocks will tend to remove the confounding between treatment and paddock effects. If the experimental area is completely uniform, no gain in power of significance tests would be achieved by rotation.

METHODS AND RESULTS

We can use the estimates of paddock error variation from past experiments to calculate the probability (power) of detecting significant differences between treatments for a specified replication rate and true difference between treatments. Three rotationally grazed randomized block experiments conducted at Ayr in North Queensland, having 3 or 4 treatments, 2 or 3 paddock replicates and 2 or 3 rotations have been studied in an attempt to quantify the power of such experiments. The approximate power when a 5% level significance test is applied were calculated for a range of replication rates and true differences, for an experiment with eight treatments and a herd size of five. The calculations were based on Tang's tables (Kempthorne 1952). Coefficients of variation were calculated from paddock mean squares for 12 week milk production data from these experiments. The paddock mean squares are associated with only 3 or 4 degrees of freedom, and so power is poorly determined. Coefficients of variation ranged from 12% (12 week summer production) to 70% (12 week winter production) for the three experiments, and these two values were used in the calculations of power.

Our study suggests that unless the researcher is looking for gross differences, he is unlikely to be able to obtain reasonable power in his significance tests. To achieve an 80% chance of detecting commercially important differences in milk production (say 0.5 kg/cow/day for inexpensive treatments to

1.5 kg/cow/day for expensive treatments), the necessary replication rates are impractical (more than five replicates). The conclusion that can be drawn is that the cost of establishing experiments of the type considered in this paper is extremely large; equally, the ongoing labour requirements are likely to be beyond available resources. This is especially so for winter milk production, which is an important subject of investigation. The lack of dairy grazing experiments having sufficient replication to assure reasonable power is surely a consequence of the factors discussed above.

The researcher often has to be content with an experiment without paddock replication, particularly where a number of treatments are proposed e.g. factorials. The technique of rotation as described above removes the confounding of treatments and paddocks, but does not provide a valid error term for use in treatment comparisons since the experimental unit is the herd and its associated paddock. Investigation of data from available replicated grazing experiments indicates that the animal error term provides a satisfactory approximation, although the results are not conclusive. I consider that the design of dairy grazing experiments needs to be the subject of further research.

REPLICATION RATES FOR GRAZING EXPERIMENTS WITH BEEF CATTLE ON NATIVE PASTURES IN CENTRAL AND NORTH QUEENSLAND

P.M. PEPPER* AND R.J. MAYER*

Cattle growth rate data from two environmentally diverse sites in north and central Queensland (NQ and CQ, respectively) were analysed to determine suitable replication rates for grazing experiments. For this purpose, probability tables were derived which show the chance of detecting true differences between treatment means; the probability values have been calculated for treatment differences within a size range considered to be of practical importance.

MATERIALS AND METHODS

At the NQ site near Millaroo, there were three series of experiments which included two series of supplementary feeding experiments of randomised block design (5 or 6 treatments x 2 replicates) over 7 consecutive years, followed by a uniformity trial on the same paddocks for 2 years. Each year a fresh draft of Brahman cross steers grazed speargrass pasture growing on solodic soils and uniform sands at .47, .43 or .33 beasts/ha, respectively, for the three series of experiments. In the second year of the uniformity trial 10 of the paddocks, being long and rectangular, were fenced in half and stocked at the same rate (.33 beast/ha) as in the previous year. For all experiments each year was divided into dry and wet seasons; average daily gain was calculated for these two periods and the overall period. The dry period corresponds to the supplementation period.

The CQ site is near Gayndah in the Burnett valley, on an open forest site dominated by speargrass and forest bluegrass. A stocking rate experiment of 4 replicates x 3 stocking rates in a randomized block design was conducted over 9 consecutive years at the same site. Each year a new draft of Hereford weaners grazed at the 3 rates (0.74, 1.24, 2.47 beasts/ha) during the active pasture growing periods (December-May).

In all experiments animals were allocated to paddock groups by stratified randomisation on the basis of fasted liveweight. The effectiveness of this strategy for controlling variation in growth rate was tested by analysis of

* QDPI Biometry Branch, Brisbane 4001 and Toowoomba 4350, Qld, respectively.

variance. . Similar tests determined if grouping the paddocks into blocks could increase precision by isolating variation from the experimental error. For each site, Bartlett's test for homogeneity of variances was used to determine if the experimental error could be pooled over years. From the expectation of the mean squares, the components of variance were calculated and Tang's tables were used to determine the approximate probabilities (power) of detecting true treatment differences of various magnitude when using a 5% significance level test (Kempthorne 1952).

RESULTS AND DISCUSSION

NQ site Analyses of the first year's data of the uniformity trial, which used the larger paddocks (30 ha) showed that contiguous paddocks grouped into blocks only succeeded in isolating variation for the dry season. For the supplementary feeding trials variation due to blocks on these same 30 ha paddocks, was significant in only 3 out of 21 analyses. Hence for experiments on these 30 ha paddocks where the paddock is the experimental unit, a completely randomised design would be preferable; since the number of degrees of freedom for error would be small, losing precious degrees of freedom with no compensating reduction in error mean square would decrease the power of significance tests. In the second year of the uniformity trial with the smaller 15 ha paddocks, blocks did effectively isolate variation from the paddocks error variation because there was a significant difference between the back and front halves of the 30 ha paddocks.

Because poor and good performing paddocks had been noticed in the supplementation trials, the effect of choosing replicates on previous performance rather than position was investigated. The performance from the first year of the uniformity trial was used to group the 30 ha paddocks into replicates. From the analyses of the second years data, this did not prove to be a successful strategy, as the sets of paddocks did not maintain their relativity across years.

Blocking animals on initial fasted weight was effective in reducing animal variation in only three of the 27 analyses at the NQ site, which indicates that animals could have been allocated to paddocks completely at random.

The paddocks variance components for the 30 ha paddocks from the first year of the uniformity trial were similar to those for the 15 ha paddocks in the second year. The "all trials" estimates of paddock and animal components of variance (Table 1) were calculated from mean squares pooled over years of both the supplementation and uniformity trials. The "all trials" variance components were used to calculate power values given in later tables.

TABLE 1 Components of variance

	Uniformity trial		All trials	
	paddocks (yr 1)	paddocks (yr 2)	paddocks	animals
dry season	3086	2366	1959	3597
wet season	2394	2148	1315	7614
overall	1015	1007	740	2619

CQ site Variation due to the replicates and animal blocks terms were not significant in any of the nine analyses. Bartlett's test showed that the paddock mean squares could be pooled over years; similarly the animal mean squares were pooled over years. Components of variance of 3453 for paddocks, and 5955 for animals, were calculated from mean squares pooled over the 9 analyses. The paddock component is larger than the corresponding figure from NQ (1315) and may be due to the much smaller paddock sizes. These were: 1.21, 2.43 or 4.05 ha,

depending on the stocking rate treatment.

Both sites For experiments where the paddock variation is the appropriate error, such as stocking rate trials, approximate probabilities of detecting true treatment differences of various magnitudes in average daily gain are given in Table 2. The probabilities are for an experiment of a randomised block design with 6 treatments.

TABLE 2 Approximate probabilities (%) of detecting true differences between treatment means as significant (P=0.05)

No. of animals/	No. of paddock	True difference (average daily gain, kg/hd/day) NQ (wet season)			CQ		
paddock	replicates	.1	.15	.2	.1	.15	.2
5	2	33	61	83	22	42	64
	3	54	87	98	38	68	90
	4	69	95	99	49	82	97
10	2	42	73	92	25	47	69
	3	67	94	99	42	74	92
	4	81	98	99.9	54	87	98

Increasing the number of animals does not have the same dramatic effect as increasing the number of paddock replicates. Hence, smaller paddocks with 5 animals each are to be recommended in preference to larger paddocks with 10 animals each. The number of replicates needed is determined by the size of true treatment differences which the experimenter considers meaningful and would like to be able to detect as significant. Although stratifying animals on initial fasted liveweight was not effective in the experiments considered here, it is a recommended precaution especially if there are only 5 animals per paddock and an animal needs to be removed for any reason.

Probability estimates are given in Table 3 for experiments where the animal carries the treatment with it and the animal variation is the appropriate experimental error. Probabilities are calculated for two treatments in a completely randomized design where all experimental animals graze together.

TABLE 3 Approximate probabilities (%) of detecting true differences between treatment means as significant (P=0.05)

No. of animals/	True difference (average daily gain, kg/hd/day) NQ dry		NQ wet		NQ overall		CQ	
treatment	.05	.1	.05	.1	.05	.1	.05	.1
10	43	93	-	67	54	98	30	81
20	72	99.5	43	94	84	99.9	52	98
30	87	99.9	58	99	95	99.9	68	99.3

The variation for animals was higher in the NQ wet season than in the dry season and this is reflected in the lower chance of detecting a true difference. In experiments where the animal variation is the experimental error there is usually little problem in obtaining enough animals to give an acceptable chance of detecting treatment differences of a magnitude regarded as economically important.

GUIDELINES FOR THE DESIGN OF EXPERIMENTS INVOLVING MEASUREMENT OF BODY COMPOSITION

R.M. SEEBECK*

BACKGROUND

The design of body composition experiments, incorporating the latest biometrical methods for expressing changes in body composition, has been recently discussed by Seebeck (1983). The present paper is intended to provide guidelines for the design of experiments involving effects of some treatment on animal production, where the primary aim is to give information on body weights and/or economic returns, but where some subsidiary information is desired on body composition measured at slaughter. Aspects of body composition that may be considered in this type of experiment include dressing percentage, yield of saleable meat, dissected or chemical composition of a sample joint, complete dissection of a side, or even the proportions of the different grades of carcass.

Collecting meaningful information at slaughter from experiments involving growing stock requires much more rigorous experimental designs than those where information on live weights is the only requirement. The main reason for this is that the composition of an animal changes with increase in size. Therefore any difference in composition that results from a treatment should be partitioned between what was caused by a simple weight difference and what was not. If this is not done, nonsensical statements are often made, such as (when referring to an experiment where all animals are slaughtered at the same time), "not only did treatment A (e.g. the highest stocking rate) decrease the live weight of the steers at the end of the period, but there were less first grade carcasses (or they were less fat or whatever)". The second part of the statement may have been just because they were smaller, or there may have been another effect operating separately from the effect of size. A more appropriate experiment and analysis would have been able to clarify this point.

Although size is the major determinant of body composition if one considers animals over a reasonable part or all of their growing period, body composition is also affected by growth rate (Seebeck 1983), presumably mainly by variation in nutrient intake. In some cases, e.g. if the treatment is a hormonal one, one should try to isolate the growth rate effect from the treatment effect. In other cases, the inclusion of a variate for growth rate may eliminate the significance of the treatment effect that appeared in an analysis without that inclusion. This is, in itself, a worthwhile conclusion.

DESIGNS AND ANALYSES

General If probability values are required for treatment differences in body composition, then both the design and statistical analysis have to be compatible with this desire. This requirement means that the design includes variation in size and possibly also growth rate and is such that it can be analysed by regression techniques. Other statistical techniques, such as factor analysis, while they may better represent patterns of development, have limited capabilities for comparing groups. To achieve a range in size unconfounded with growth rate, serial slaughter designs have to be used. These designs can be set up by slaughtering groups of animals at a series of either weights or ages, the latter necessitating that a rate of growth effect be removed statistically to enable estimation of development patterns in an unbiased form (Seebeck 1983). Removal of the rate of growth effect is also desirable when slaughtering at a

* CSIRO Tropical Cattle Research Centre, Rockhampton, Qld 4701.

series of weights, because it is difficult to kill animals at exactly the same size, due to variation in gut fill, etc. (Seebeck 1983). The allocation of animals to the slaughter groups must be done on a random basis (preferably stratified on the basis of initial live weights) before the start of the experiment.

It is suggested that a treatment group should consist of at least 10 animals, spread over at least 3 and preferably 5 slaughter groups. This will give results that will be able to be partitioned into treatment, size and rate of growth effects. If the treatment is imposed at or near the start of the slaughter programme, the treatment will have a progressive effect and the regression of composition on size (and/or growth rate) will be expected to have different slopes. The different slopes can be dealt with in the analysis by comparison of regression line techniques, or in the general least squares sense by including treatment x covariate interactions.

With experiments involving an imposed treatment(s), it is useful to know what the composition of each animal is at the start, so that the change in composition, adjusted for size and rate of growth for each animal, can be determined. However, using such an initial slaughter group to calculate the composition of each other animal is not very satisfactory, because the calculation is only based on the average animal. The animals that would have comprised the initial slaughter group are best slaughtered at a subsequent time spread over the treatments and used to better define the treatment effects.

A more useful strategy is to perform one or more in vivo estimates of composition on all animals at the start of the experiment (such as TOH injection), slaughtering some of the animals then as an initial slaughter group, and repeating the in vivo measurements on all other animals at least at the time the individual is slaughtered. This type of information can greatly improve the precision of estimation of changes within an animal.

Using serial slaughter, rather than killing all animals at the same weight or time, naturally decreases the information available from the experiment directly concerning that particular weight or time. However, with a good serial slaughter design, the loss of accuracy is not very great and is compensated for by more information at other weights or times. If a specific weight or time is of particular interest, there is nothing to stop the experimenter slaughtering more animals at that point than at other points.

Pen experiments With pen experiments, the use of serial slaughter reduces the information on feed efficiency comparisons, because of the different times that the animals are on feed. Consequently, it is probably better to space the slaughter programme unequally, with an initial slaughter group and, say, as a compromise, slaughter at weights or times corresponding to 50, 75 and 100% of the total experimental period envisaged. Initial in vivo composition estimates are almost a necessity with this type of experiment. For experiments that are reasonably short, say of six months or less duration, the slaughter of the animals may be better done on a weight gain or time basis rather than on a weight basis. This is because the range of initial weights may be large compared to the weight gain over the experiment. However, the range of initial weights should probably not be artificially restricted because the treatment may have differential effects on animals of different initial weights.

Grazing experiments The major constraint on body composition information from an experiment with grazing animals is that slaughter of some animals during the experiment, as a consequence of serial slaughter, will change the grazing pressure. This is of particular importance if the experiment involves treatments

in different paddocks, such as a stocking rate experiment. The only practical solution to this is to have a matching group of animals, members of which can be brought in to maintain the grazing pressure. Such a scheme presupposes that there are sufficient animals in a paddock that can be slaughtered at three (or certainly a minimum to two) different times or weights. To have any variation in rate of growth at least two animals have to be killed at each point, so that the minimum number of animals per paddock is four. If other constraints put the number at less than that, it would be better to forget any thought of doing body composition estimates.

Where paddocks are the experimental unit, as in stocking rate or pasture evaluation trials, then the variation between paddocks treated alike is used as the error for between treatment comparisons. Regression slopes, between composition and size and between composition and rate of growth (if fitted) will almost certainly have to be estimated by pooling across replicates because of limitations on degrees of freedom. However, before pooling across treatments, treatment differences in slope should be tested.

REPLICATION RATES FOR SURVIVAL FEEDING EXPERIMENTS IN PENS WITH WEANER BEEF CATTLE

P.K. O'ROURKE*

The contributions to total variance of liveweight change from replicates, pens, liveweight strata and animals have been collated from 12 survival feeding experiments carried out in the pen complex at Swan's Lagoon, north Queensland. Recommendations have been derived for the number of pens per treatment and animals per pen which give a high probability of significance for detecting a specified difference between treatments. Guidelines for the design of survival feeding experiments are developed.

MATERIALS AND METHODS

Brahman cross weaner steers aged 9-12 months were used in all 12 experiments. The number of animals, their mean initial liveweight and standard deviation are summarised in Table 1. A low quality diet of either rice straw or native pasture hay was fed daily and ad libitum on a pen basis. Experiments compared a range of nitrogen and energy supplements with an unsupplemented control group. The length of the feeding period and rate of change in liveweight for the control group are given in Table 1.

The pen complex at Swan's Lagoon consists of 24 identical pens, each 6 m x 9 m. Experiments generally used a randomized block design with two replicates, one on either side of a central laneway. The exceptions were Experiments 6 and 12 with two and three replicates, respectively, in completely randomized designs. For Experiments 6 and 7 animals were allocated to pens completely at random. Stratified randomization on initial liveweight was used for all others. In Experiments 3, 4 and 7 the major stratification into heavy and light animals corresponded with pen replication; otherwise stratification was across all pens. Four animals per pen were used except for Experiment 3 which used five. The number of treatments ranged from 3 to 12 and may be inferred from the residual pen degrees of freedom in Table 1.

Daily change in liveweight, calculated from initial and final unfasted liveweights, was analysed for each experiment to separate variation due to treatments, replicates, pens, liveweight strata and animals. Degrees of freedom

* QDPI Biometry Branch, Brisbane, Qld 4001.

TABLE 1. Initial liveweight (kg), control group average, residual pen and animal mean squares for average daily gain (g)

Expt. no.	Initial liveweight (kg) n	Mean	SD	Control group ADG (g)	No. of days feeding	Residual Pen DF	Mean Square	Residual Animal DF	Mean Square
1	40	193	16.5	-127	126	4	4420	27	5020
2	24	206	19.6	77	63	2	2374	15	5929
3	100	190	20.8	-336	112	9	4839	71	5778
4	96	194	18.9	-283	112	11	8592	65	4586
5	80	190	16.0	-227	84	9	17385	57	8162
6	32	198	19.7	-434	120	4	11853	22	8076
7	64	163	21.8	-319	60	7	30597	46	14177
8	72	196	10.7	-407	71	8	9244	47	31572
9	80	208	16.3	-96	64	9	23301	57	29137
10	64	174	11.5	-206	57	7	47718	45	12038
11	72	215	16.6	-421	65	8	41385	50	35803
12	72	148	11.2	-273	65	12	27462	51	23625
Pooled	796	189	16.9	-261	83	90	20632	553	16111
Median	72	193	16.5	-280	68	-	14619	-	10100
Minimum	24	148	10.7	-434	57	-	2374	-	4586
Maximum	100	215	21.8	77	126	-	47718	-	35803

and mean squares for residual pen and animal terms are given in Table 1. The descriptive statistics pooled mean, median, minimum and maximum summarise the data from individual experiments.

RESULTS AND DISCUSSION

The randomized block design did not increase precision in that the replicate effect was not significant ($P>0.05$) in any of the experiments. Hence, the completely randomized design, which would also contribute extra error degrees of freedom, should be preferred. A significant proportion ($P<0.05$) of the animal variation in experiments 3, 4 and 10 was accounted for by liveweight strata. Thus stratification is a worthwhile precaution which involves little extra effort. The residual pen term was significantly greater ($P<0.05$) than the residual animal term for Experiments 5 and 10.

Bartlett's test for homogeneity of variances was applied across the 12 experiments. Both the residual pen mean squares and the residual animal mean squares showed heterogeneity with chi-squared values of 23.2 and 147.6, respectively ($P<0.05$); however, the pooled mean squares provide the most appropriate basis for general recommendations on replication rates for future survival feeding experiments in pens. Variance components of 16111 for animals and 1103 for pens have been estimated from the pooled mean squares for use in calculation of appropriate replication rates.

The procedure outlined by Cochran and Cox (1957) has been used to calculate the approximate probabilities (power) of detecting a significant difference ($P<0.05$) between treatment means for a specified true difference, number of pens per treatment and animals per pen, and error degrees of freedom. Ten has been used as a conservative and recommended minimum value for the error (residual pen) degrees of freedom. Probabilities have been set out in Table 2 for a range of the other three parameters. All probabilities are much higher when the true

TABLE 2 Approximate probabilities (%) of detecting a significant difference in treatment means for a range of true differences, numbers of pen replicates and numbers of animals per pen and using 10 for the residual pen (error) degrees of freedom

True difference (g/day)	No. of pen replicates	No. of animals					
		2	3	4	5	10	20
100	2	13	17	21	25	38	52
	3	18	25	31	36	55	71
	4	24	32	40	47	68	83
200	2	45	60	71	78	93	98
	3	63	78	87	92	98	100
	4	76	89	94	97	100	100

difference is increased from 100 to 200 g/day. Since the variance component for animals is so much larger than that for pens, the major determinant of power is the total number of animals per treatment; for a fixed total there is little benefit in increasing the number of pens and reducing the number of animals per pen. On the other hand, increasing the number of animals per pen and minimising the number of pens is likely to be more cost efficient. However, this must be balanced against the need to retain sufficient pen replication to generate a minimum of ten degrees of freedom for the residual pen term.

The 12 experiments used to generate the probabilities in Table 2 had basically two replicate pens, each of four animals per treatment. Hence, they have probabilities of 21% and 71% of picking up differences between treatments of 100 and 200 g/day, respectively. The corresponding probabilities for four pens with two animals in each, are 24% and 76% ; this illustrates the marginal advantage to be gained by maximising the number of pens for a fixed total number of animals.

Experiments 1-6 had feeding periods which were, on average, 39 days longer than those used in the more recent experiments. In spite of similar handling, type of animal, mean and standard deviation of initial liveweights, the mean squares for rate of liveweight change during the feeding period were much larger for experiments 7-12. The variance components of 759 for pens and of 6131 for animals, estimated from the pooled mean squares for experiments 1-6, were lower than the corresponding variance components of 1122 and 24776 for experiments 7-12. Hence, one cost of a shorter feeding period could be markedly higher variability, particularly for the animal component. Calculations similar to those presented in Table 2 could be repeated for variance components from any individual experiment. The lower probabilities of detecting specified differences for experiments with shorter feeding periods need to be balanced against the reduced cost, time and physical effort for shorter experiments.

In survival feeding experiments the residual pen term is the appropriate estimate of experimental error. The potential exists to use animal variation as the error term, thereby taking advantage of higher error degrees of freedom. The opportunity to reduce costs by having only a single pen replicate of each treatment is also available. Neither of these approaches is recommended. Since the pen component of variance is positive, to ignore it could lead to overstating levels of significance. As well as this statistical consideration, the pen is the logical experimental unit in this type of experiment for other data, such as feed intake, which is typically collected on a pen basis.

CONCLUSIONS

G.W. BLIGHT

If constraints on resources fix the number of replicates at what are thought to be inadequate levels, there are two artificial ways to increase the power of tests of significance of treatment differences. Firstly, change the test e.g. carry out 10% level tests instead of the traditional 5% (or 1%) level test of significance. Secondly, increase the size of the true treatment difference that we wish to be able to detect. Otherwise accept a reduced power, say 50%, rather than a power figure of around 80% - the usually accepted figure.

Investigations of replication rates for grazing experiments with beef cattle suggest that power is maximized for a given area of land and total number of animals if paddock group sizes are kept to a minimum (3-5 animals) and the number of paddocks per treatment increased. For survival feeding experiments in pens it is the total number of animals per treatment which affects power and it is cost efficient to increase the animal numbers per treatment by having larger pens; two pens per treatment was recommended provided this was sufficient to give ten degrees of freedom for the residual pen error term. In pen experiments and multi-paddock grazing experiments in large paddocks (30 ha) a completely randomized layout of pens or paddocks is preferred to a randomized blocks layout; having replicates in blocks was found to be inefficient and wasteful of precious error degrees of freedom. Blocking animals on initial weight classes was recommended for both pen and grazing experiments.

Where body composition information measured at slaughter is required, the ideal experiment will probably never be done. In the face of competing experimental requirements, there are some good design principles that can be followed to achieve the best compromise. In particular, serial slaughter designs should be used if a range in size unconfounded with growth rate is to be achieved.

With grazing experiments, a lack of uniformity amongst paddocks often is a problem; using animal uniformity data at the design stage as a criterion for assigning paddocks to blocks was shown to be ineffective in north Queensland. There may be an opportunity to use simple regression models, relating animal performance data to herbage quality and quantity covariates, to disentangle confounded paddock and treatment effects in unreplicated (or inadequately replicated) grazing experiments. There would still remain the difficult problem of constructing an error variance. The development of (i) regression models to adjust animal production for the confounded paddock effects and (ii) methods for constructing an error variance, may be worthy of further research in the case of dairy production experiments, where resource requirements rarely permit adequate paddock replication rates.

In concluding his address given at the last biennial conference, the President of this Society affirmed that the competitive advantage in overseas markets enjoyed by Australia's pastoral industry depended on low cost production (from extensively grazed systems). Queensland's beef industry is based on lightly stocked native pastures in central and northern environments - typical low cost production systems. Good quality research aimed at improving the productivity and management of these native pastures is essential. Research in which comparisons are made between stocking rates or fertilizer levels or pasture additives, requires adequately replicated experiments in multi-paddock layouts if the results are to be statistically and scientifically acceptable.

ACKNOWLEDGEMENTS

The authors wish to thank the following officers of the QDPI who made data available: Messrs G. Chopping, L. Winks, S. McLennan, W. Scattini, A. Ernst and J. Lindsay.

REFERENCES

ANDREW, C.S. and FERGUS, I.F. (1976). In "Tropical Pasture Research, Principles and Methods" Bul. 51, p.101, editors N.H. Shaw and W.W. Bryan. (Commonwealth Agricultural Bureaux: Farnham Royal).

BRYAN, W.W. (1968). Aust. J. Exp. Agric. Anim. Husb. 8:683.

COCHRAN, W.G. and COX, G.M. (1957). "Experimental Designs". 2nd Ed. (Wiley:New York).

HATHEWAY, W.H. and WILLIAMS, E.J. (1958). Biometrics 14: 207.

HAYDOCK, K.P. (1964). In "Some Concepts and Methods in Sub-tropical Pasture Research" Bul. 47, p.159, editor W.W. Bryan et al. (Commonwealth Agricultural Bureaux: Farnham Royal).

HOWARD, D.A., BURDIN, M.L. and LAMPKIN, G.H. (1962). J. Agric. Sci. 59: 251.

JONES, R.M. (1982). Trop. Grassld. 16: 118.

KEMPTHORNE, 0. (1952). "The Design and Analysis of Experiments". (Wiley: New York).

LUDLOW, M.M. (1976). In "Tropical Pasture Research, Principles and Methods" Bul. 51, p.251, editors N.H. Shaw and W.W. Bryan. (Commonwealth Agricultural Bureaux: Farnham Royal).

MANNETJE, L. 't (1974). In "Proceedings 12th International Grassland Congress, Moscow" Vol. 3, p.299, editors, V.G. Iglovikov and A.P. Mousissyants. (Izdatel'stus Mir; Moscow).

MANNETJE, L. 't, JONES, R.J. and STOBBS, H. (1976). In "Tropical Pasture Research, Principles and Methods" Bul. 51, p.194, editors N.H. Shaw and W.W. Bryan. (Commonwealth Agricultural Bureaux: Farnham Royal).

ROE, R., SOUTHCOTT, W.H. and TURNER, H.N. (1959). Aust. J. Agric. Res. 10: 530.

SEEBECK, R.M. (1983). Anim. Prod. 37: Part 1 (In Press).

SHAW, N.H. (1961). Aust. J. Exp. Agric. Anim. Husb. 1: 73.

SHAW, N.H. and MANNETJE, L. 't (1970). Trop. Grassld. 4: 43.

SMITH, H.F. (1938). J. Agric. Sci. 28: 1.

WILLOUGHBY, M.W. (1959). Aust. J. Agric. Res. 10: 248.

YATES, J.J., EDYE, L.A., DAVIES, J.G. and HAYDOCK, K.P. (1964). Aust. J. Exp. Agric. Anim. Husb. 4: 326.