Fundamental Considerations of Measurement, Analysis and Interpretation in Designing Performance Recording Systems in Animals

J.W. James

Dept. Wool & Animal Science, NSW University, Australia

1.0 Introduction

In traditional husbandry the measurement of production has played little part, with assessment of animals usually being made by subjective evaluation of performance. Such evaluation may be fairly accurate when flocks or herds are small and closely supervised because under such conditions individual animals are very well known to the people concerned. However, when groups of animals are too large or husbandry systems are not suitable for detailed individual knowledge of animal performance, measurement provides the only possibility for comparison among different animals or groups. Measurement is particularly important when animals are to be compared and the animals involved are not all known by those who must make decisions on choices between them. Even when there is good personal knowledge of animals in a particular population, that knowledge may not be generally available and may not be in a form convenient for analysis of differences in profitability of various management procedures or for planning breeding programs. Therefore even in low or medium input farming systems it may be desirable to institute performance recording in order to improve general management or genetic improvement.

The rational management of an animal population of any type may be regarded as involving the attainment of the goals of the owners of the population in the most efficient manner. One way of approaching this ideal is to identify the desired goals in terms of measured performance. Goals of the managers may be considered to consist of inputs and outputs, with the aim being to maximise the difference between the value of inputs and the value of outputs. Unless inputs and outputs are quantified their comparison is necessarily somewhat arbitrary. While it is possible to manage a population to obtain a satisfactory outcome without measurements, it is likely that some use of measurements will improve flock or herd productivity. When national or regional goals are concerned, it seems that only a degree of performance recording would allow their pursuit. With the growing realisation of the importance of sustainable agricultural systems there will be increasing pressure to quantify the effects of different husbandry practices.

Both for normal management and for genetic improvement programs there are similar considerations to be taken into account in formulating a performance recording scheme. The measurements to be taken and the recording and analysis systems implemented must be relevant to the goals of the overall program, they must be practical and cost effective, and they must be acceptable to farmers. Normal management may be seen as the taking of decisions in order to maximise the profitability of existing resources, while genetic improvement may be seen as providing improved animal resources for the future. Strictly speaking this is an arbitrary distinction, since some management decisions, such as a decision to reduce stocking rates to avoid overgrazing, will have impacts on future productivity, and some management procedures will affect genetic improvement. However, in a broad sense this distinction between optimising use of existing resources and providing better animals for the future can be seen as the major difference. Given that there is overlap between the two types of performance recording schemes, and that a specific scheme may serve both purposes, this discussion will not treat the two types separately. As the requirements for a breeding program may often be more stringent, the major emphasis will be on planning of systems for this purpose, but except for some special requirements for genetic improvement which will usually be obvious the principles apply equally to both management and breeding recording systems. For example, records may be used for making management decisions, or for providing information required by government bodies but not used in management. This would in many ways be analogous to the distinction made below about reasons for performance recording in breeding programs. In fact, there is in general a close analogy between the information used to make decisions for selection and to make management decisions. Farmers and their advisors concerned to identify problems leading to sub-optimal rates of growth or reproduction will need the same kinds of data on growth rate or breeding success as will breeders concerned to increase rates of genetic improvement.

Broadly speaking there are two main reasons for measuring performance in breeding programs: the provision of information on which selection decisions are to be made, and provision of information which may be of value but which is not intended for use in selection decisions. There are good reasons for making a distinction between these purposes, as will be discussed in due course. The more important purpose is to aid in making selection decisions, and this will be the main focus of the present document.

2.0 Choice of traits

There are in principle very many measurements which could be made as part of a performance recording system. Since there are costs associated with every measurement, it is only those traits whose measurement makes a cost-effective contribution to genetic progress which can justifiably be included in such a system. How can such traits be chosen?

2.1 Importance for breeding goal

In a strict sense, choice of traits can be made only after extensive investigations have been made. Before it is possible to assess a contribution to genetic progress it is necessary that genetic progress be defined, and this can logically be done only after a breeding goal has been defined (James 1982). The definition of a breeding goal is a decision to be made by the owner of the breeding population in question, and is in some sense arbitrary, but in recent times there has been extensive discussion of how this might be done in a structured way (e.g. Amer and Fox, 1992; Smith, James and Brascamp, 1986; Brascamp, Smith and Guy, 1985; Ponzoni, 1986; Goddard, 1997). Essentially it appears to be an emerging consensus that the approach recommended by Amer and Fox (1992) can be used as a basis for the definition of goals, though in practice there may be difficulties in a rigorous application of the method because of market distortions for political or other reasons. It is also recognised that there will always be considerable uncertainties in such definitions because the goals will not be achieved for many years, by which time conditions may have changed, so that it may be that breeders who best guess the future may be more successful than those who derive their goals in the most logical manner but guess the future badly. Nevertheless it remains true that only in relation to a defined goal is it possible to assess the contribution of a performance recording system. In the same way, for a management oriented system it is only when the crucial decisions have been identified that the value of performance recording in making those decisions can be assessed.

In making such an assessment it is important to keep a clear distinction between traits which are to be improved because they are part of the breeding objective, and traits which are to be used as selection criteria, which may include some or all of the traits in the objective, but may also include traits which are of no intrinsic value but because of their correlations with important traits can be used to effect genetic improvement. As an example of the distinction between these traits, we may cite feed intake of grazing animals as a component of the breeding goal in most, if not all husbandry systems, but it is one which would not be routinely measured as a selection criterion because of the impracticability of such measurement. On the other hand, a measurement of backfat thickness at a specified location is of minor importance in itself, but may be a valuable selection criterion because it has a strong correlation with the total amount of fat in the carcase after slaughter. In designing a performance recording system it is not sufficient to identify economically important traits and decide to measure them, as indicated by the example of feed intake of grazing animals. But the identification of economically important traits can lead to the search for correlated traits which may be used as selection criteria for improvement of characters which are too difficult or too expensive to measure.

Before reliable choices can be made it is also necessary that accurate estimates of genetic parameters are available, and these will often be lacking. Indeed, in some cases a good reason for embarking on a performance recording program is to obtain data for the estimation of such parameters. Accurate estimation of genetic parameters requires quite extensive data sets even with modern software such as DFREML (Meyer, 1989) or VCE (Groeneveld and Kovac, 1990) which is capable of extracting information from field data in an efficient manner. In the first stages of a recording program, then, it is likely that the necessary information will not be available. The way to proceed in this case is generally taken to be as follows. For most traits there are estimates available in some populations, even if they are not closely related to the population of interest. A literature survey will give some guidance as to the estimates to be expected in the population of interest. Of course, it is a truism that the heritability of a trait is defined for a particular population in a particular environment, but experience suggests that if no better information is available, a value taken from a literature search is better than nothing and may even be better than a very inaccurate estimate from a small amount of data taken from the relevant population. Similar remarks apply to genetic correlations, which are notoriously difficult to estimate accurately, so that in some circumstances the use of a phenotypic correlation as an estimate of a genetic correlation is justified. In making such guesses from literature reports it is wise to note that variances and covariances have often been found to differ substantially between environments with low inputs and high stress levels and environments with high inputs and low stress. Populations of animals adapted to such different environments are also likely to differ in variance structures when maintained in the same environment. If possible guesses should be based on estimates for broadly similar environments.

In the long run there is no real substitute for estimates from an appropriate data set, but decisions must often be made in the short run. For the moment we shall assume that a breeding goal has been defined and a range of potential selection criteria have been identified, following which a set of estimates of genetic parameters for the traits in the breeding goal and the selection criteria have been assembled.

The information obtained in this way can be then used to assess the contribution each trait to be used as a selection criterion can make to genetic progress. A simple way to approach this problem is to use selection index methods, as done for example by Ponzoni (1986). Once the breeding goal has been defined, the response to selection using an index which includes all potential criteria can be predicted from selection index theory. It is then simple using freely available software to check the effects of dropping certain criteria from the index on the response to selection. If a criterion can be left out of the index with a negligible loss of response we must question the value of measuring the trait, recording it and using it in the prediction of breeding values. If the omission of a trait from the index leads to substantial loss of response we would be reluctant to remove it from the criteria to be used even if it were expensive to measure. A survey of possible selection criteria by this method can be very informative and give rapid insight into the relative importance of different criteria in a breeding program.

So far we have been considering selection within a population, where continuing genetic improvement is being sought. Measurement is also important for selection between different populations, and may be especially important because the various populations may be familiar to different groups of people. However, measurements in themselves are of little value unless they are made in comparable conditions (ideally in the same environment) and are relevant to breeding goals in the same way as in selection within populations. For this purpose it may be that individual breeding values are of lesser importance than the mean breeding values of the different populations, and the considerations involved are not identical in the two cases.

2.2 Practicability

It has already been mentioned above that the practicability of measurement must be taken into account when planning a performance recording system, using the example of feed intake of a grazing animal. The taking of performance measurements is sometimes a part of normal husbandry, but in most commercial populations and many seedstock populations the extent of measurement is minimal. Any performance recording system needs to be integrated into the management of the population, and thus must fit in with other management operations. This can be done more easily in a seedstock population in the sense that genetic improvement in such a population can more readily lead to increased returns than in a commercial population where the sale of breeding animals is not a source of income.

In wool sheep, measurement of the quantity and quality of wool can be integrated into the shearing process if the fleece is normally handled for skirting by placing the fleece on a weighing table, and taking a sample for later determination of quality (e.g. yield, fineness) before it is packed with other fleeces in a bale. It will be necessary to introduce some changes into the shearing operation compared to one in which no measurement is done, where fleeces are subjectively classed, but the incorporation of the measurement process into the shearing shed procedure does not call for a major reorganisation. On the other hand, if a sheep breeder decides to breed for parasite resistance by using faecal egg count as a selection criterion, particularly if this is to be done using a challenge with a known dose of larvae, a whole new management procedure needs to be introduced. This is not necessarily an argument against such an operation, but it does illustrate the fact that some measurements can easily be added to existing management practices while others need substantial changes to husbandry if they are to be used. An important part of the design of a performance recording system is the establishment of a method of putting the system into operation, and this may need to be done in different ways in different populations if their management is sufficiently different. A system which is seen as disruptive by the manager of the population is not likely to be favourably regarded. Thus at an early stage it is important to find out what management procures are employed in order that an appropriate way of including measurements in the overall program can be devised.

2.3 Cost

Although there may be no clear distinction between practicability and cost, in the sense that at a sufficient expenditure virtually any procedure may be made practicable, there are nevertheless real differences. As an example, we may consider the case of wool measurement mentioned above. It may be quite practicable to institute a fleece measurement program, with samples of wool being kept for later laboratory measurement, but the cost of such measurements may be regarded as too great to be acceptable.

Where cost is an important consideration, as it very often is, it may be desirable to reduce expense by using a cheaper measurement procedure or by measuring only a proportion of the animals which might have been measured if costs were much smaller. The true costs of measurement and recording need to be recognised, since these costs may include the labour of gathering the animals together for measurement if the measurement cannot be done at a time when the animals would be normally collected together.

In general a cheap measurement is more attractive than an expensive one, not necessarily because the total expenditure can thus be reduced, but there is also the chance that the total expenditure may be the same, but several cheap measurements may be taken instead of a single expensive one.

The general principle which needs to be recognised is that a given total measurement budget can be spent in many different ways, and for any given budget, one should seek the most effective set of measurements which may be feasible for such a total cost. The contribution of the measurements to progress in the direction of the breeding goal is the criterion which should be used in the optimisation of the measurement procedure. Once the most efficient process has been found for a given budget, one may ask whether a larger or a smaller budget would be advantageous. But the optimum performance recording budget can be determined only when the budget is in fact spent in the optimal fashion. This involves the choice of which measurements are to be taken and on which animals these measurements are to be made.

3.0 Accuracy

It should be remembered when assessing a performance recording system that what we have are not actually performances, but records of measurements of performance. Thus there will be errors in our records, because of the precision of our measuring instruments, the accuracy of the procedure in which the instruments are employed, and the extent to which transcription and other errors arise in the recording process. A system should be designed with these potential sources of error in mind, but an obsession with unnecessary precision can be counter-productive. We thus need to consider the relative importance of such factors for the planning of a performance recording system.

3.1 Precision of instruments

Other things being equal, we will usually prefer a precise measuring instrument to an imprecise one, but other things are seldom, if ever, equal. Nobody, I believe, would want to weigh beef cattle with scales which recorded weight to the nearest gram, because such precision of measurement would be seen as unnecessary, and scales which would weigh to this precision would be much more expensive than scales weighing to (say) the nearest kilogram, and the weighing process for such precise scales would no doubt be too complex. It is easy to recognise extreme cases where precision of measurement may be much too great or much too small, but we should have a general approach which will allow a suitable precision to be chosen.

If the unit of measurement for a continuous variable is denoted as U, values of the variable are rounded up or down by an amount depending on how far they are from the nearest scale point. It has been known for a long time that the effect of this is to increase the variance of measurements by an amount $U^2/12$, known as Sheppard's correction. If the variance of the measured performance is σ^2 then a part of this variance is the rounding error. The rounding error variance as a fraction of the total is $U^2/12\sigma^2$, and a sensible choice of the unit of measurement can be made on this basis. For example, if the fraction of the variance due to rounding is regarded as acceptable if it is no greater than one percent, the implication is that $U^2 \leq 0.12\sigma^2$ or $U \leq 0.35\sigma$. Thus for a trait with a coefficient of variation of 10% the unit of measurement should be no greater than 3.5% of the mean. Or, if we consider the weighing of cattle whose weights have a standard deviation of 30 kg, an acceptable unit of measurement by this criterion would be 10 kg. Of course it is open to anyone to choose what level of rounding error variance is acceptable, but in my opinion it is hard to argue for greater precision than about one-third of a standard deviation, although more precision may be acceptable if it can be achieved without further cost.

Such a conclusion is sometimes resisted by practical breeders, on the grounds that it may lead to failure to recognise genuine differences. However, a proper consideration shows that the overall effect of such differences must be negligible, as removal of the rounding error would increase the heritability by at most one percent. An increase of this order can hardly be regarded as of major significance.

3.2 Precision of measurement methods

As a rule, the total variance of measurement errors will be appreciably greater than that due to rounding errors. Though it has been shown above that rounding errors are unlikely to be a serious problem, other types of measurement error are not so easily disposed of. For instance, in the weighing of animals it is possible that unless precautions are taken variations in factors such as gut fill will contribute substantial amounts of variation, much more than will arise from rounding. We must therefore consider the effect of measurement errors on response to selection.

This can be done by looking at the correlation between estimated and true breeding value. For individual single-trait selection this is the square root of the heritability. If the fraction of the phenotypic variance which is due to measurement error is denoted by m, then the ratio of selection accuracies with and without measurement errors is $\sqrt{(1-m)}$ so that the loss of response due to measurement errors is approximately 0.5 m. The exact percentage losses of response for different percentages of variance due to measurement errors are shown below.

	Percentages					
Measurement error	0.1	01	2	5	10	20
Loss of response	0.05	0.50	1.01	2.53	5.13	10.56

The situation is different if selection is to be on a progeny test, since in this case the measurement errors are averaged over the progeny group. The accuracy of a progeny test based on n progeny is $\sqrt{[n/(n + a)]}$ where $a=(4-h^2)/h^2$. On substituting the values of h^2 with and without measurement error we find that the ratio of accuracies is $\sqrt{[1-m/\{1+(n-1)h^2/4\}]}$. Thus the fraction of response lost due to measurement error is approximately $0.5m/\{1+(n-1)h^2/4\}$. As expected, this is less than for individual selection, and becomes less important as heritability and family size increase. If we consider an index combining individual performance with the average of n relatives we find that the loss due to measurement errors in the case when relatives are half sibs is intermediate between the losses with individual selection and with progeny testing. This is not surprising since the combined index is a weighted average of the selection criteria in the other two cases. These points are illustrated in the following table.

These cases involve selection for a single trait, but in practice selection is commonly based on several traits. The general case obviously has too many parameters for a useful summary, but to illustrate the nature of the effects of having to consider more than one trait we may treat the case of individual selection on an index of two traits each of equal economic importance, with equal heritabilities and equal fractions of measurement error variance. The analysis of this case is straightforward, and the ratio of accuracies with and without measurement errors turns out to be $\sqrt{[1-m/(1+r_P)]}$ where r_P is the phenotypic correlation between the traits. In this case the loss of response due to measurement error is approximately $0.5m/(1+r_P)$. If the traits are positively correlated the effect is less than for a single trait, while if they are negatively correlated the effect will be greater than for a single trait. In this balanced case the genetic correlation does not matter. Percentage loss of response due to measurement error for a range of values of heritability, family size, and fraction of variance due to measurement error. Progeny test values in normal and combined selection values in *italics*.

m	h^2	n				
		5	10	20	50	100
0.01	0.1	0.46 *0.48*	0.41 *0.45*	0.34 *0.40*	0.22 *0.30*	0.14 *0.23*
	0.25	0.40 *0.46*	0.32 *0.42*	0.23 *0.36*	0.12 *0.29*	0.07 *0.25*
	0.5	0.33 *0.45*	0.24 *0.42*	0.15 *0.38*	0.07 *0.35*	0.04 *0.34*
0.02	0.1	0.91 *0.96*	0.82 *0.90*	0.68 *0.80*	0.45 *0.60*	0.29 *0.46*
	0.25	0.80 *0.92*	0.64 *0.83*	0.46 *0.72*	0.25 *0.58*	0.14 *0.51*
	0.5	0.67 *0.91*	0.47 *0.84*	0.30 *0.77*	0.14 *0.71*	0.07 *0.68*
0.05	0.1	2.30 *2.42*	2.06 *2.26*	1.71 *2.00*	1.13 *1.52*	0.72 *1.16*
	0.25	2.02 *2.31*	1.61 *2.10*	1.15 *1.83*	0.62 *1.48*	0.35 *1.29*
	0.5	1.68 *2.28*	1.18 *2.12*	0.74 *1.96*	0.35 *1.81*	0.19 *1.74*
0.1	0.1	4.65 *4.89*	4.17 *4.59*	3.45 *4.06*	2.27 *3.09*	1.45 *2.37*
	0.25	4.08 *4.69*	3.25 *4.26*	2.31 *3.71*	1.24 *3.03*	0.70 *2.64*
	0.5	3.39 *4.64*	2.38 *4.32*	1.49 *4.01*	0.70 *3.71*	0.37 *3.59*
0.2	0.1	9.55 *10.06*	8.53 *9.42*	7.03 *8.33*	4.60 *6.37*	2.92 *4.93*
	0.25	8.35 *9.65*	6.62 *8.79*	4.68 *7.71*	2.49 *6.37*	1.40 *5.66*
	0.5	6.91 *9.60*	4.82 *8.99*	3.01 *8.42*	1.41 *7.90*	0.75 *7.68*

If we are concerned not with the choice of breeding animals within a single population but with the choice between two or more populations, the importance of measurement errors will be reduced because their effects will be averaged across the samples from the populations to be compared. In a sense this is rather like the case of progeny testing.

We see that measurement errors have a greater effect when selection is based on individual measurements than when selection is based on family information, and that when multitrait selection is practised the effects of measurement errors in the traits may be greater or smaller than for a single trait. On the assumption that errors of measurement may contribute appreciably greater variance than does rounding we must pay more attention to obtaining measurements in a way which will avoid unnecessary errors.

An important source of measurement error in many situations is the sampling of performance for measurement. Some examples of sampling would be:

- measurement of lactation milk yield of cows by testing daily milk yield on a number of test days;
- measurement of fibre diameter of the fleece on a wool sample taken from the fleece;
- measurement of egg weight in chickens by weighing eggs laid during a specified period;
- measurement of carcase fat content by measurement of the fat thickness at a specific body site.

In such cases it is perhaps more appropriate to take these as correlated characters rather than as inaccurate measurements of the trait of interest. From this viewpoint we are concerned with the heritability of the measured trait and its genetic and phenotypic correlations with the target trait. This is different from the pure measurement error case, where it is assumed that the genetic correlation between the measured trait and the target trait is unity, whereas it is sensible to assume an imperfect genetic correlation between total carcase fat and fat thickness at a given point. In fact this example is so clear that the fat thickness measurement would normally be regarded as a correlated trait rather than a measurement on a sample of the total phenotype. Such characters are perhaps best regarded as indicator traits, included in the performance recording program not for their own sake but to help estimation of breeding values through their correlations.

3.3 Precision of recording methods

As stated above, the data available for analysis and decision-making are records of measurements, and errors may be introduced in the recording process. Such errors are difficult to assess because they are often difficult to quantify. Recording errors may be of several kinds. For example, a record of 254 kg may be mistakenly written as 245 kg. Such transcription errors are not uncommon, and may be quite undetectable, as would probably be true in the example unless all weights were taken to the nearest 5 kg. On the other hand a similar error of transcription resulting in a record of 524 kg is almost certain to be detected as an error, and one may even guess that it should be 254 kg, but the record would usually be discarded.

In the same way it may be that a measurement is correctly recorded but is ascribed to the wrong animal, either because the animal identity is wrongly entered at the time of recording, or because the measurement was entered in the wrong place on a prepared data entry form. Experience suggests that such errors are less likely when prepared data entry forms are available, if only because there are then fewer entries to be made, so that having prepared data forms is to be recommended where it is feasible.

Technology is now available to enable automatic data capture at the time of measurement. If animals have electronic identification and appropriate sensors are used, the identity and the measurement can be entered on an electronic data base without the need for transcription by people involved in the measurement process. When suitable systems have been developed they not only reduce the frequency of recording errors because electronic devices are less prone to error than are humans, the automatic recording frees the attention of those involved in the measurement procedure from the needs of recording and is likely to lead to better concentration on other aspects of the process, thus leading to even better accuracy. The problem is that the technology is not yet developed for many measurement procedures and cost may be very high, as for instance in the measurement of individual food intake in pigs. Further development of automated measuring and recording devices could aid significantly in the improvement of performance recording.

When information is based on an individual animal's own performance alone it does not matter for selection whether its pedigree is known or not. But when information on relatives is used in making selection decisions it clearly does matter that the pedigree records should be correct. It is also necessary to know pedigrees if attempts to avoid close inbreeding are made. We know that pedigrees are often in error, and as a result records on animals are wrongly combined when information on relatives is used to estimate breeding values. Many methods of varying accuracy are available for determination of parentage, but the current range of DNA markers certainly provide the best method for identifying the pedigrees of animals in a group. Unfortunately the costs of this technology are as yet too high for its use except in special circumstances. While its cost will no doubt decrease, it seems unlikely to play a major part in animal breeding in the short to medium term.

The importance of pedigree errors can be illustrated by considering their impact in progeny testing of sires, assuming that a fraction f of progeny are falsely attributed to a sire and are randomly drawn from the progeny of other sires. The accuracy of the estimated breeding value of a sire under such conditions as a fraction of the accuracy when there are no pedigree errors can be shown to be

$$(1 - f) / \sqrt{[1 - f (2n - 1 - nf) / (n + a)]}$$

where n is the total number of progeny attributed to the sire and $a=(4-h^2)/h^2$. As the "family" size becomes large this approaches unity because the number of true progeny is very large. The following small table shows the percentage losses in accuracy for a few cases.

Percentage loss of response in progeny testing due to pedigree errors

Percentage errors in pedigrees	Heritability	Family size (including errors) 10	20	50	75	100
1	0.1	0.81	0.67	0.45	0.35	0.29
	0.25	0.62	0.45	0.24	0.17	0.14
	0.5	0.45	0.28	0.13	0.09	0.07
2	0.1	1.62	1.35	0.90	0.71	0.58
	0.25	1.25	0.90	0.49	0.35	0.28
	0.5	0.90	0.57	0.27	0.19	0.14
5	0.1	4.09	3.43	2.32	1.82	1.50
	0.25	3.19	2.31	1.26	0.92	0.72
	0.5	2.30	1.47	0.70	0.49	0.38
10	0.1	8.30	7.04	4.83	3.83	3.18
	0.25	6.57	4.83	2.69	1.96	1.55
	0.5	4.82	3.12	1.51	1.06	0.81
20	0.1	17.07	14.79	10.56	8.53	7.16
	0.25	13.93	10.56	6.12	4.53	3.60
	0.5	10.56	7.05	3.53	2.49	1.93

For small family sizes and low heritabilities the loss of response is almost equal to the rate of pedigree errors, but as family size and heritability increase the loss of response declines. It may be noted that erroneous pedigrees are worse than missing pedigrees because they contaminate the information on a sire whereas missing pedigrees simply reduce the amount of correct information. The greater the use made of pedigree information in a breeding program the greater will be the impact of missing and erroneous pedigrees.

3.4 Need for correction factors

Performances of animals are of value to the breeder as guides to their breeding values, non-genetic sources of variation essentially making recognition of genetic differences more difficult. Thus if some environmental factors which affect performance can be identified, making allowance for the influence of these factors will increase the accuracy with which breeding value can be estimated. For example, the weaning weight of a lamb will be influenced by its age at weaning, by whether it was born as a single, twin, or higher order multiple birth, whether it was reared as a single lamb, twin etc., by the age of its dam and so on. When information on these factors is available, records can be adjusted to a standard condition, for example, a single-born, single-reared lamb born to a mature ewe and weaned at 90 days of age. Provided that appropriate correction factors are applied the variation due to these influences can be removed from the phenotypic variance resulting in a higher heritability and therefore more accurate estimate of breeding value.

Before such corrections can be made the data on the factors must be available. Thus in the example the date of birth of the lamb must be recorded, as must its birth and rearing types, and the age of its dam must be known. If these pieces of information are not routinely recorded, the question arises as to the value of this information, which can be expressed as the reduction in variance brought about by the corrections. We can illustrate the nature of the considerations by taking the case of type of rearing and age at weaning in the above example.

Let us suppose that there are only two classes of animals reared, singles and twins. Let the fractions of singles and twins in the population be (1-t) and t respectively, and let the difference in weaning weight between the two groups be D. The variance due to type of rearing is $t(1-t)D^2$ and as a fraction of the total phenotypic variance it is $t(1-t)D^2/\sigma^2$. Its importance thus depends on two factors: the relative magnitude of the effect D/σ, and the incidence of twin-reared animals in the population. If they are rare, then t(1-t) will be small and even a moderately large value of D/σ will not cause a significant loss of accuracy if correction is not made. On the other hand, if the incidence is intermediate, of the order of 20% to 80%, then correction is likely to make an appreciable improvement in the accuracy of estimated breeding values. The incidence of the different groups is then crucial in assessing the need for corrections of effects of this kind. The percentage loss of response when a binomial factor is not corrected for is illustrated in the following table.

Percentage loss of response when a binomial factor with an incidence t or 1-t and an effect D/σ as a fraction of the total variance (including its contribution) is not corrected for in individual selection.

Factor effect	Incidence of the factor						
	0.01	0.05	0.01	0.02	0.03	0.04	0.05
0.01	0.00	0.02	0.04	0.08	0.11	0.12	0.13
0.02	0.02	0.10	0.18	0.32	0.42	0.48	0.50
0.05	0.12	0.60	1.13	2.02	2.66	3.05	3.18
0.75	0.28	1.34	2.56	4.61	6.09	6.99	7.30
1.00	0.50	2.40	4.61	8.35	11.12	12.82	13.40

If we consider the correction for age, we can assume that the correction is done by regression, so that if g represents the rate of gain per day at weaning, A represents age at weaning, and M the mean age at weaning, then the correction to weaning weight is g(M - A). Sometimes corrections are made on a proportional basis so that heavier lambs have larger adjustments than light lambs, but the differences due to such alternative correction methods are too small to be important here. The variance removed by such corrections is $g^2 \sigma_A^2$, where σ_A^2 is the variance in the population of age at weaning. As a fraction of the phenotypic variance this is $g^2\sigma_A^2/\sigma^2$. A very simple assumption would be that the growth was approximately linear over the pre-weaning period, in which case the mean weaning weight W would be approximately gM. Then we would have $g^2\sigma_A^2 / \sigma^2 = C_A^2/C_W^2$ where C stands for the coefficient of variation. Thus the crucial factor in deciding whether variation in age is important is the coefficient of variation of age as a fraction of the coefficient of variation of weaning weight.

In this example, it is clear that if dates of birth and therefore ages at weaning are not very different, there is little point in recording dates of birth and making corrections for weaning age. In the same way if there is a factor which has a fairly large effect on the trait in which we are interested but which is of very rare occurrence, there will be little value in trying to correct for it as so few animals will be affected. If corrections do not appreciably reduce the average squared error they will not make an important contribution to increasing genetic progress.

3.5 Group identification

One potential source of variation which should always be taken into account is the occurrence of what may be called management groups, that is, groups of animals which are run together, treated in the same way and generally differentiated from other such groups. Management groups may sometimes differ very distinctly in performance from one another, or in some cases may differ very little, but experience shows that even when on *a priori* grounds there is little expectation of large group effects they nevertheless can occur. The only safe way to deal with such group effects is to assume that they may be large and to allow for them. This means that a group identification system is required, and should be planned in the design of the performance recording system. In particular it needs to be made clear to those managing the animals that animals in different groups should be regarded as different and records be kept of group formation. This is so even when different management is applied only for a limited period, with the animals being later combined into a single group. If the importance of group identification is not stressed it is likely that in such cases the previous difference in treatment will not be remembered when performance is being measured and recorded. Failure to identify such groups properly has been a serious problem in many performance recording systems.

The number of ways in which animals may be grouped is very large. A few examples of effects which should be taken into account are:
1. animals born in different seasons;
2. animals born outside the herd;
3. animals cared for by different farm workers:
4. animals which have been given different veterinary treatments from others;
5. animals which have been given special feeding for any purpose, such as pregnant females;
6. animals which have been weighed in groups on different days;
7. animals which have been mated by artificial insemination versus those mated naturally.

All of these factors and many more can in some circumstances greatly affect performance, while in some cases the effects may be negligible. It is usually impossible to be sure that the effect would be negligible, and so the animals ought to be identified as forming different groups.

4.0 Choice of animals

When performance is being measured and recorded there is a wide range in the extent to which it may be done, from the extremes of being applied to all animals in a population to none (though this would be relevant only if the animals were being measured for some other trait). For an effective and cost efficient performance recording system it is necessary that the recorded animals be chosen from those available in such a way as to serve the purposes for which the data are collected without unnecessary trouble and expense. To satisfy such criteria we must consider the purpose of recording, the section of the population to be recorded, how many animals should be recorded, and the returns from recording in relation to their costs.

4.1 Characterisation of a population

On occasions performance records will be required in order to obtain an indication of the average productivity of a population, whether that population is a breed, a strain within a breed or a particular herd or flock within a strain. Such a characterisation may be desired because there are a number of such populations and we want to consider making choices among them, or because the population is being used in an improvement program and we want to monitor changes in traits which are not being measured as part of the selection criteria for the breeding program. If the intent is then to assess the average quality of the population it is clearly important that the animals measured should be representative of the population, and therefore they should be chosen by a random process for that purpose.

Let us first suppose that we are concerned to characterise the mean breeding value of a population for a specific trait. How shall we choose which animals to record? Usually a population will not be a homogeneous group but will be divided into subgroups more or less equally related among themselves. In the simplest case where we have no other pedigree information we may know the sire of every animal available for measurement, but know no more about relationships among sires, or who the dams are. In this case we know that if we choose to measure a total of T animals such that each of s sires has n progeny measured, and T = sn, then the sampling variance of the estimated mean is given by

$$\sigma_S^2/s+\sigma_W^2/T=(n\sigma_S^2+\sigma_W^2)/T$$

where σ_S^2 is the variance between sire breeding values and σ_W^2 is the variance between performance records of progeny of the same sire. Thus for a given value of T the sampling variance is smallest when n is smallest, that is, when all available sires are equally represented among the progeny. If it is assumed that the cost of measurement is not affected by the sire of an animal it follows that all sires should be represented when T>s, if necessary by only one offspring. In fact the optimum is to have T sires with one offspring each.

Similar principles apply when the population has a more complex structure, such as being divided into herds, with sires being nested within herds and having half sib progeny. If σ_H^2 is the between herd variance component then the sampling variance of the estimated mean is given by

$$(sn\sigma_H^2+n\sigma_S^2+\sigma_W^2)/T$$

where now T=hsn and h is the number of herds sampled. Again it is clear that as many herds as possible should be sampled and as many sires per herd as possible should be taken if the total cost is determined by the number of animals. In this second case it is more likely that there will be extra costs associated with measuring animals from different herds, and a proper solution would involve using the expression above to compute the accuracy attainable for different sampling plans with the same total cost.

The picture presented above will be oversimplified in practice. For various reasons it may not be possible to sample equally from all subgroups, but this is not of major importance, and the general principles are easy to apply. A more difficult problem if we wish to compare different genetic groups is likely to be that the genetic groups are not kept in the same environments. It was stressed above that management groups should always be taken into account, and if different populations are kept entirely separately it follows that a proper comparison between them is impossible. Only when there is some comparison among animals of different genetic groups in the same management group can a meaningful estimate of genetic difference be made. For this purpose it may be possible to use a control group of animals in several management groups to provide an indirect comparison of genetic groups which do not themselves share the same management group.

For a simple illustration, suppose we have two strains of a particular breed, and have t management groups of each strain, none of which include animals of the other strain, but each of which contains n progeny of each of s sires of the strain, all ts sires of each strain being unrelated. Each management group also contains a set of m progeny of a single sire of a different strain produced by mating the sire, probably by artificial insemination, to a random sample of females of the strain in the management group. The records of animals in all management groups are then linked by the occurrence of the common or link sire. We assume that any heterosis occurring is the same for progeny of the link sire with dams of both of the strains to be compared, and that interaction of genotype and environment can be ignored. If these assumptions fail, then the use of indirect comparisons is not possible using this procedure. The problem of heterosis could be overcome if animals of a third pure strain were used instead of progeny of a sire from such a strain mated to dams of the two strains to be compared, but the problem of genotype - environment interaction cannot be solved without direct comparison in the same management group.

It can be shown that for this experimental design the variance of the difference between the estimates of the two strain means is

$$2\sigma_S^2[mn+(sn+m)\alpha]/tsnm$$

where σ_S^2 is the component of variance between sires (assumed the same in all strains) and α is the ratio of within sire to between sire variance. As would be intuitively expected, the variance is reduced as t increases if the same structure is used for every management group. Similarly, the precision of the comparison increases as the number of sires tested in each management group increases. It is likely that the number of groups and the number of sires per group would be limited, so an important question about the design of such a comparison is the number of progeny of the link sire which will give the greatest precision. This is easily seen to occur when half the progeny in each management group are from the link sire. This design problem is related to that considered by Miraei

Ashtiani and James (1991) but differs in that they were concerned with the average accuracy of comparisons among individual sires, whereas here comparison is of two genetic group means.

In practice it would very seldom be possible to use a design such as that outlined here, but the optima found for this simple case should provide useful guidelines which can be aimed for in more complex situations.

4.2 Monitoring

If we are primarily concerned with monitoring, the situation is similar to that described above in the sense that we are concerned to characterise a population, with the difference that the interest is not in making a choice between populations but to assess the mean value of the population which is being improved. For example, we may have decided that it will be possible to improve the growth rate of a population without increasing its feed conversion ratio by use of an index which does not depend on the measurement of feed consumption. However, it will be desirable to check on the validity of our prediction by periodic measuring of feed consumption in the population. The question to be decided is the manner in which such measurements ought to be made.

The primary problem when difficult measurements are to be made is that allowance must be made for possible environmental changes with time. Some type of genetic control is therefore necessary. It will be very rarely that an unselected control population will be available, and so other methods, such as the use of frozen semen from males born a number of years ago to produce offspring for comparison with progeny of the current breeding population will be necessary. There are several variants of this system which have been discussed by Smith (1977). All require that advance planning be done, and that the appropriate expertise is available to apply them. A somewhat simpler method to apply is the comparison of progeny of sires of different ages using field records (Smith, 1962). This approach depends on the existence of a sufficiently long record file, but otherwise is not very demanding. On the other hand, the accuracy is not very high, and its use for estimation of genetic change has limited value when the trait investigated is of major importance, but for a trait which is being monitored to check for gross changes it may be of real value. Of course, if a fully pedigreed population is involved and the data are available, it would be sensible to use a suitable BLUP analysis rather than the type of analysis considered by Smith before BLUP was a realistic option, but the precision depends on the data structure rather than the method of analysis, which can extract no more information than is in the data, though it may not extract all information which is present. Perhaps the use of a Bayesian analysis could be taken as not consistent with this statement, since Bayesian analysis depends on the introduction of prior information, but while this does add information it does not add information to the data itself. In this sense it is analogous to the assumption of a normal distribution for the data in conventional statistical analyses, which does allow extraction of more information from data than if we can assume only (say) that the data are symmetrically distributed, but not in a known form.

4.3 Candidates for selection

When performance recording is intended solely for the purpose of making genetic improvement, as distinct from being used as a general management tool, it follows that there is no reason to record performance on an animal unless it is a candidate for selection. Of course if one wishes to measure genetic progress or conduct a research program the situation is different, but from the point of view of making genetic progress, a performance record is of value only if it contributes to decisions as to which animals will be selected or what matings will be made. A measurement on an animal which will not be considered as a possible parent will make no difference to breeding program decisions and is therefore unnecessary. Thus animals which are culled as being unacceptable for some reason need not have performance records made. Similarly, an animal which has already been chosen as a breeding animal need have no further measurements made, since these will not affect whether it will be selected or not.

The principle enunciated above is oversimplified for application in modern breeding programs, especially those in which information on relatives is an important contributor to selection decisions. For example, if lambs which are extremely small at weaning are culled without being weighed then a progeny test of sires for yearling weight based on the weights of surviving lambs will be biased in favour of the sires whose lambs were culled, providing that the culling was dependent at least in part on genetic factors, and not solely on large environmental effects. A similar consideration will apply in other cases such as animal model BLUP evaluation where biased samples of relatives would be involved. With animal model BLUP the absence of information does not lead to bias provided that the information on which selection decisions is made is included in the analysis. Nevertheless, if the suggestion above that no further measurements be made on animals which have already been selected is accepted, then information on the traits which were not measured for that animal is not available for the assessment of its relatives. Thus when information from relatives is used in making selection decisions it is much more difficult to justify the failure to record performance on any animals, since all animals have some relatives, and therefore their performance could be relevant to the genetic improvement being made.

However, when selection is being made on individual performance, it is true that performance records are only of value when made on candidates for selection. Thus it is often possible to reduce costs through use of independent culling levels, when decisions are made progressively, and animals culled at any stage of the selection process need not be measured for further traits. This has long been recognised as an advantage of

independent culling level selection over index selection (e.g. Turner and Young, 1969) especially as in some cases the genetic improvement achievable through index selection may not be very much greater than that obtainable by shrewd use of independent culling. In principle this concept can be extended, so that at any stage of selection animals can be divided into three classes:
1. those which will definitely be culled;
2. those which will definitely be used for breeding;
3. those whose fate is still undecided.

Clearly animals in groups 1 and 2 need no further measurements, and performance recording can be concentrated on those animals for which a decision has yet to be reached. This approach has been investigated by Wade (1989) who found that there was in fact usually little to be gained by this division into three classes over what could be achieved by conventional independent culling. In effect the approach can be seen as a combination of independent culling and selection of extremes (Abplanalp, 1972). A similar problem was dealt with by James (1979) in considering optimum progeny group size for selecting among sires when prior information was available. He showed that depending on the relative importance of existing and future information, it might be best to obtain more progeny from sires whose current estimates of breeding value were intermediate and fewer progeny of sires with either high or low current estimated breeding values. The justification for this was similar to that in the case dealt with by Wade, namely that if a sire had a low EBV, or a high EBV, the probability that further information would greatly alter its chance of selection was low, and so it was not worthwhile to expend effort on improving the accuracy of its EBV, the effort being better spent on resolving differences among those sires with EBVs which were closer to the truncation point.

These are examples of the general principle that for genetic progress attention should be concentrated on making better selection decisions, and that information which does not contribute to that purpose is less valuable than possible alternative data which does.

4.4 Number of candidates to measure

When a given number of breeding animals is required to be chosen it is clear that the more candidates that are evaluated, the greater will be the selection differential achieved, though the increase in selection differential will follow a diminishing returns curve. For example the small table below shows how the selection differential changes as the ratio of number tested to number required increases.

Ratio	2	4	8	16	32
Selection Differential	0.7979	1.2711	1.6469	1.9678	2.2523

Clearly, as the amount of measurement doubles, the selection differential increases at a slower rate, and doubling the number measured leads to progressively smaller increments in the selection differential. It is then obvious that at some point the cost of measuring more animals will not be repaid by a corresponding increase in the value of genetic improvement made. Of course in practice there will usually be a limit on the number of animals available for measurement, and it may be that the point at which further measurement becomes unprofitable is at a greater number of animals than there are present, so that all available animals should be measured. It is particularly when measurements are very expensive that a restriction on measurement is likely to be desirable, for the obvious reason that it is then more difficult to recoup the expenditure in greater genetic gain.

When not all candidates are measured, there is an important point to be taken into account. It is always better to select an unmeasured animal than a measured animal which is below the average, because the EBV for an animal which is unmeasured is zero, as we have no information on it, whereas an animal whose performance is below average has a negative EBV (Smith, 1959). This means that unmeasured animals may still be candidates for selection, depending on the number of breeding animals needed and the number of animals measured. We need to distinguish between the concepts of candidates for selection and animals which are performance recorded.

4.5 Allocation between sexes

As the sexes are usually selected with different intensities and often play different roles in the production process the measurements made on males and females will commonly be different.

Even if the same traits are measured in both sexes it is possible that the heritabilities of the traits will differ between males and females, or that the genetic correlation between performances of males and females is less than unity. These factors are likely to mean that a different amount of testing in males and females is optimal. In principle we may look at the problem as follows. If H represents the breeding objective or overall genetic merit, the rate of genetic gain can be written as

$$R_H=\sigma_H[i_M r_{HI\,(M)} + i_F r_{HI\,(F)}]/[L_M+L_F]$$

where i is the standardised selection differential, r_{HI} is the correlation between breeding objective and selection criterion, and L is the generation length, with the subscript denoting the appropriate sex. We can approach the question of optimum use of performance testing by assuming that there is a total budget which is to be allocated. This budget can be spent to different extents on males and females and in each sex the allocation can be spread over different traits. This chosen allocation will result in a particular accuracy of selection for each sex, depending on the traits recorded, a particular selection differential which depends on the number of animals of that sex tested, and a particular generation length depending on the time scale over which recording is conducted. Upon inserting these values in the equation for genetic gain we find the response to be expected for that strategy. Other strategies can then be investigated and the one which gives the best response can be identified. In any particular case some sort of trial and error process is probably necessary, but some general ideas can be obtained by making some assumptions about the way in which the accuracy of estimation of breeding value depends on the amount of money spent. This has been done by Jackson *et al.,* (1986) and Wade and James (1996) with particular emphasis on the division of the total testing budget between the sexes.

4.6 Avoidance of recording

Even in a very simple situation, selection on performance records may be complicated. As an example consider selection on fleece weight in a sheep flock. Sheep must be shorn, the fleeces must be weighed, the weights must be recorded and then the best animals identified by comparison of the records. Then the selected animals must be separated from those to be culled. This process requires at least a temporary form of identification to be attached to every animal so that it can be associated with the performance record and the appropriate animal then chosen on the basis of the records. It also means that the animals must all be handled at least twice, once when the shearing is being done, and once when the sheep selected for breeding are being separated from those culled. In some circumstances this extra work will be unwelcome and may act as a disincentive to use performance measurements. It would be simpler if one could set a standard before the shearing began, and as each sheep's fleece weight was measured it could be compared with the standard and a decision could be made at once. This would eliminate the need for individual identification and for a second handling, though it would probably slow down the shearing and weighing somewhat. However, there would be a substantial saving. The difficulty would be in setting the standard, since the mean and standard deviation of fleece weight are not known in advance, because of large environmental differences between years. If the standard is set in advance it may be far too high or far too low. If the standard is set too high, then too few animals will be kept, and since the number required for breeding will be fixed, it will be necessary to make up the numbers from those culled, but this must of necessity be done at random, since there is no recorded information on which to base the choice. If the standard is set too low then too many animals will be kept, and their mean will be lower than if the correct number had been kept. Some can be discarded in order to reduce the number to that required, but again this must be done at random. The result in either case is that the mean of those kept for breeding will be lower than it would have been had the exactly correct standard been set.

Of course, the exactly correct standard can be set only after all the animals have been measured, but Tallis (1961) has shown that a culling standard can be set by sampling a small number of the animals and using the measurements to estimate the truncation point. He has also considered an alternative strategy of forming three groups, one of which consists of animals to be culled, one of animals to be kept, and a third intermediate group of animals to be kept in reserve as spares in case the group of animals to be kept turns out to be too small. Tallis gives an example where 60% of the animals measured are to be retained for breeding. If a sample of 75 animals is taken, the fleece weight of the 38th is used to set the standard for animals retained, all those with fleece weights between those of the 38th and 52nd are put in reserve, then the efficiency of the procedure is 97%. In doing these calculations he assumed that the sheep in the sample were tagged for later selection to avoid loss of efficiency. While the technique developed by Tallis may have limited application, it is well worth consideration when a minimal cost procedure is required. There are tables provided in Tallis (1961) so that calculations need not be made by the practitioner.

5.0 System design

So far we have considered a number of factors which are of importance in the development of a performance recording system. This background now enables us to look at the overall design of such a system. The design of a system should begin with the objective, since the system is intended to serve the purpose of the genetic improvement program. If the goal of the program is not clearly defined then a logical design of the measurement and recording system is not possible because there is no rational basis on which different designs can be compared.

In the case of a single breeding program such as might be controlled by a large breeding company the objective might be defined in simple economic terms as the maximisation of company profit over a certain time period, perhaps reduced to net present value. However, in other cases the goal will be to provide a service to many breeders whose objectives will not be identical, as is the case for many government-supported schemes, which aim to provide performance records for the use of many independent breeders, whose goals may differ appreciably. In such a case the provision of a service may be determined by the demand expressed by the breeders, perhaps expressed as a willingness to pay for the service. When there is a demand for performance records to the extent that their costs are willingly paid by their users, a government or commercial service can easily justify the provision of the information. Nevertheless, it will often be the case that in planning a service it will not be clear beyond doubt what will or will not be paid for ungrudgingly, and in the design stages any service provider will want to have some confidence that the system will have relevance to users' needs, and one important way of assessing this is to try to relate the information provided to the goal of a breeding program. Therefore, it will often be advisable to try to define breeding goals as the basis for design of a recording system.

5.1 Consideration of breeding goal

There have been several general discussions of defining breeding goals in recent years, and some of these have been mentioned earlier. This is not the place for a detailed discussion of how this should be done. There is evidence that the definition of the overall breeding value can be in error to some degree without greatly reducing the rate of genetic progress, although the errors may result in the redistribution of change across different characters (e.g. Smith 1983). The major problem is to ensure that all economically important characters are taken into account when the objective is defined, and are given the correct signs. It seems that a factor of two in the relative weights attached to different traits can be accommodated without serious loss of response in most situations. This is comforting when one is planning a performance recording system since it is likely that errors will be made, and also because in cases where the system has many users these users will wish to use different economic weights. The general procedure is first to consider a production system, and then to identify the inputs to the system and the outputs, where inputs include capital items and outputs include quality of product. It is important that all inputs and outputs should be identified, including those which are not formally traded. The unpaid work done by family members should be appropriately valued and included as an input, and products consumed by the farmer and family should also be valued and included as outputs. In some systems these components may be major parts of the input-output system and their omission could seriously distort the analysis.

The prices for inputs and for outputs must then be estimated, and these prices should logically be those which are expected to apply when the improved animals are incurring the costs or providing the returns. Unless there is good reason to expect the relative prices to change markedly in the future it is usually adequate to take an average of recent prices for this purpose. It is sometimes thought desirable to define a “biological” goal such as lean tissue feed conversion rate in meat animals, but there can seldom if ever be a real justification for this approach. To take lean tissue feed conversion rate as an example, it is assumed costs other than feed can be ignored, and that outputs other than lean tissue have no value, whereas they may have either positive or negative value. And certainly there will be other costs such as those for health care which will be relevant. It may be that variation in such other components is considered negligible in relation to variation in the chosen goal, but this assumption is not likely to be checked if the “biological” goal is preferred as a matter of principle. Whenever there are two or more inputs and two or more outputs (including quality) these must be reduced to a common unit of measure before they can be combined into an overall figure, and the simplest way to do this is in terms of money.

5.2 Alternative selection criteria

In the context of performance recording systems the point of having a breeding goal is to enable the comparison to be made of different combinations of selection criteria, while in the context of planning a breeding program the comparison of alternative selection criteria is done to enable the greatest progress to be made towards the goal. When a performance recording system is to serve a number of different breeding programs, those criteria which may be of value in one are not necessarily of value in another, and some compromises may need to be made.

As a first step the traits to be considered for recording should usually include all of the economically important traits, even if these may include characters which are regarded as too difficult for routine measurement, such as feed intake of grazing animals. If some important traits are excluded *a priori* there will never be a consideration of what their possible contribution to progress could be, and thus no chance to assess whether some effort to include such performance among selection criteria ought to be made. Naturally, in such a case it is probable that feed intake will quickly be eliminated from the list of criteria, but if it is thought about initially it is at least possible to assess what loss is incurred by not measuring it, and to investigate the extent to which other characters can be used to provide some of the information lost by ignoring it.

As a general rule, when considering criteria for selection, we should begin by being willing to evaluate any trait for which an argument can be advanced to support its use as an aid to selection. Type or conformation traits may be in this category, if they can be shown to lead to better estimation of overall breeding value, even if they are of no intrinsic value (though in some cases they may have value because of subjective preferences). However, while we should start by being willing to investigate the potential of any trait, we should also be willing to discard any trait which has been demonstrated to make a negligible contribution to the accuracy of genetic evaluation.

5.3 Prediction of change

As was pointed out previously, the rate of genetic improvement is proportional to the accuracy of estimation of breeding value, measured as the correlation between the true and estimated breeding values. Thus in order to predict response it is necessary to be able to calculate this correlation. Given that a breeding goal has been defined, the correlation can be calculated based on the parameters of the traits in the objective and the traits used as criteria for estimation of breeding value. The variance of overall merit can be computed as $a'Ga$ where a is the vector of economic weights and G is the matrix of genetic variances and covariances among the traits in the objective. The variance of a genetic evaluation is $b'Pb$ where b is the vector of coefficients for traits used to estimate breeding value and P is the matrix of (phenotypic) variances and covariances of criteria, which may include records on relatives, in which case some of the phenotypic covariances may be fractions of genetic covariances. The covariance of true and estimated breeding values is a'Qb where Q is the matrix of (genetic) covariances between the traits in the objective and the evaluation criteria. Then the correlation is

$$a'Qb/\sqrt{[(a'Ga)(b'Pb)]}$$

Now since the vector b is the solution of

$$P\ b = Qa$$

we can replace b in the expression for the correlation and obtain

$$a'QP^{-1}Qa\ /\ \sqrt{[(a'Ga)(a'Q'P^{-1}Qa)]}$$

That is, in order to calculate the correlation we must know a, P, Q and G. Note that this is more information than is actually needed to predict breeding values, since G is not needed for this. Only genetic covariances between objective traits and evaluation criteria are necessary to compute EBVs, genetic variances and covariances of objective traits being needed only for the variance of overall merit. Since this is a constant and unaffected by the criteria used, we can ignore it. This is easily seen if we look at the genetic superiority of selected individuals for overall merit H based on selection on an index I where the genetic superiority is $ir_{IH}\,\sigma_H$ with i as the standardised selection differential so that in the above notation it becomes

$$i\sqrt{(a'QP^{-1}Qa)} = i_{\sigma I}$$

Thus to evaluate the relative merits of different sets of criteria it is not necessary that G be known though the actual values of r_{IH} cannot be calculated without this knowledge. So the minimum necessary knowledge is a, Q and P.

It may be objected that in some cases this data may not be available, and that therefore the suggestion that decisions require the data cannot be correct. This is not so. Decisions can be made without all these data, but they cannot be based on a rational comparison of the outcomes of using different sets of criteria, since it is then impossible to predict the outcomes. If the outcomes are unknown, they cannot be compared. If the data are not known, then some guesses must be made as to their values before the various sets of criteria can be evaluated. Such an exercise may be particularly useful in showing the importance of gathering parameter estimates for objective traits or evaluation criteria. This would be done by assuming what appear to be reasonable values for unknown parameters, perhaps taking upper and lower limits, and seeing what the contributions of the criteria are over the range of assumed values. If a criterion appears likely to be of little value regardless of the true parameter values, it can be dismissed and no further attention need be paid. However, if the criterion appears to be possibly useful for an apparently reasonable set of parameters, then there may well be good reason to recommend a program of research to obtain the necessary estimates.

Of course, in practice the parameters are never known, but more or less accurate estimates may be available. If the parameter estimates are somewhat imprecise it will be wise to consider a range of possible values in making assessments. It should also be remembered that errors in parameter estimates can lead to inconsistencies, and the possibility of inconsistency ought to be checked. Another point which should always be considered is that if a number of criteria are considered, the probability of obtaining an inconsistent set of estimated parameters increases as the number rises, and so does the chance of finding an apparently useful criterion when it is in fact not helpful.

The result of these calculations will be a set of predicted relative rates of response to selection based on a range of sets of criteria. These predicted responses are then used to determine the contributions which different criteria can make to genetic improvement.

5.4 Program costs

If all information were to be made freely available, then clearly the criteria which gave maximum response would be chosen for use. When information is already collected for other purposes its cost will be zero for use in breeding programs, and so the cost can be ignored in costing. Other data may be collected for the purpose of the breeding program but be used as aids in other management problems. In such cases the cost of the data collection and recording can be discounted for the value it contributes to other management before being added as a cost of breeding. The main point is that the marginal cost should be debited to the breeding program.

In an earlier section the costing of performance recording was discussed in a general way. Here we do not need to consider the details of costs of particular schemes, but to concentrate on the manner in which an approach to the cost of the system should be made. There will usually be a basic system of a fairly simple nature which can be used as a starting point. For example, in wool sheep it is possible by inspection to gain a rough idea of the amount and fineness of wool in the fleece. The accuracy of such subjective assessment depends on how good a judge is used, which is always hard to know, but studies have been made which show that some judges are fairly good, and an average value of accuracy can be established. In such a baseline the costs would be restricted to the mustering of the sheep, payment of the judge and any assistants required in the judging process. Decisions can be made as animals are inspected, and further mustering is not needed, so the overall costs can be low, especially if the judge is the breeder and the judging costs are simply part of the whole operation. A system like this is in fact widely used in practice, and can make useful genetic progress, though at a slower rate than achievable through objective measurement. When a breeder is introducing objective measurement therefore, the cost of measurements should be assessed against the cost of the alternative subjective assessment method. If it is decided to add the measurements to the existing program then the full costs of performance recording would be added to the breeder's costs. However, for a proper evaluation of the objective measurements they should be also treated as replacing the subjective judging process, thus resulting in the saving of any costs associated with it. One may wish to consider three cases: the traditional system, replacement by objective measurement, and combination of traditional and objective assessments. Strictly, this is the correct thing to do, as all possibilities should be taken into account. But for reasons not connected with such rational analysis there may be a necessity to include some practices in a breeding program if it is to have credibility in the marketplace. At the stage of introduction of new methods it may be necessary to treat them as additions to existing practices rather than as replacements for them, although replacement may be the best long-term option.

Essentially, what is required is an analysis of costs within the framework of the farming operation, with all changes which will be introduced by performance recording having their costs (or savings) carefully accounted for. This will be very difficult to do with great precision, but should be done as well as possible. It would be a serious mistake to set up a performance recording system based on an analysis of costs which was badly wrong and find that the system was unused because the costs were much greater than had been thought. One aspect which can easily be overlooked is the cost of providing advice to users who are unfamiliar with the practice of performance recording. In many countries there has been a tradition of government support for agricultural advisers, but this support is being reduced, and if the costs of such support are omitted because they are considered to be zero in a marginal sense (because the advisers are paid whether or not they give advice on the recording scheme) this may have an important influence on the total costing of the scheme. This is only one example of the way in which subsidies can affect breeding programs.

Many performance recording systems have been set up and operated quite successfully without preliminary costing having been established for a range of options. This has been possible because not all options have been considered, and because it has been clear that a system will be useful. The approach discussed here has been based on trying to find the best system, whereas it will often be fairly easy to find a system which is better than the current one, even though it may not be the best possible. Nevertheless, it will always be the case that we are comparing several contemplated systems, not all possible ones, and choosing among those under consideration. Perhaps the choice will be obvious without analysis in some situations, but even in these a demonstration of the superiority of the preferred option has much to recommend it.

5.5 Evaluation of benefits

It is generally much easier to evaluate the costs of a breeding program than the returns, despite some of the difficulties mentioned in the last section. This is because the costs are for things done by the breeder and can therefore be accounted for in a comparatively straightforward manner. Returns are dependent on many factors, some of which are very difficult to predict, and some of which are outside the control of the breeder, such as changes in consumer demand for products. It is difficult enough to predict accurately what genetic changes will take place following introduction of a specific selection program, given uncertainties in parameter estimates and problems of predicting response in the long term. Even if this can be done satisfactorily, it is hard to decide how much benefit will accrue to the breeder. It is commonly accepted that the major beneficiary of genetic improvement is the public, since if better and/or cheaper products are available competition will prevent the breeder from keeping the profits from this improvement, and eventually any excess profit level will return to normal. And yet experience seems to show that in at least some cases breeders do retain benefits from genetic progress for long periods. The enormous genetic improvement in the poultry meat industry over the past half century has led to a great expansion in the consumption of chicken meat, and this has benefited poultry breeders in an expanded market. And within the industry those firms with successful programs have expanded and many have disappeared, whether from weakness in genetics or business sense.

The results of competition are so difficult to predict because they depend not only what the breeder does but what others do, both in the same industry and in other industries. In that sense the value of genetic progress to a breeder may simply be the maintenance of his present competitive position. In view of these uncertainties we must regard all calculations of returns with a degree of scepticism. Yet if we are to make rational decisions we must make such calculations. A rather clear account of how this can be done has been given by Amer and Fox (1992). A detailed example using a different procedure can be seen in Ponzoni (1986). Whatever way of calculating returns is chosen, there will be a value of the returns, and this can then have the costs of achieving these returns subtracted to give the economic benefit of the program.

Choices can then be made among programs based on their relative economic benefits. A performance recording system can conceivably serve several different breeding programs, so the choice of such a system may depend on the range of breeding programs put in operation by users of the system.

6.0 Choice of potential systems

The principle has been enunciated here that the selection of a recording system ought to be rationally based, and that the decision should rest on a cost-benefit analysis of the problem. Some uncertainty is unavoidable in such situations, and it may seem that in the light of these the methodical approach outlined can hardly be justified. It must be recognised, however, that the establishment of performance recording will have value only if it is put to proper use after it has been introduced, and this will normally involve its use for several years before its operation can be reviewed. There will thus be substantial costs incurred when a system is set up, and only if there is evidence that these costs can be recouped can an argument be advanced in favour of the initiation of the system.

This should not be interpreted to mean that every such scheme will need an independent complete analysis. There may be prior experience in a similar situation which can be taken as support for introduction of a similar scheme, perhaps with modifications to adapt it to the particular circumstances. Provided that the existing scheme is known to be worthwhile, it may be rather easy to justify starting a similar scheme, without going through all of the steps mentioned. The result may then be a good scheme, perhaps not the best, but one which will clearly improve on the current situation.

However, if a scheme is being considered it will always be advisable that the questions raised in the previous discussion be brought out into the open and considered. Even if it is then decided to adopt a "ready to wear" system the possible weaknesses of such a system will have been brought to our attention, and the way will have been cleared for the introduction of modifications at a later stage. There will also have been calculations made which will be helpful in demonstrating the value of the scheme to users and backers. The comparatively small amount of effort (in relation to the work of setting up a system) will be amply repaid.

7.0 Implementation strategies

A performance recording system can be organised on many levels:
1. it may be run by a government agency;
2. it may be run by an industry body (e.g. breed association);
3. it may be privately run (e.g. breeding company, individual breeder).

There may or not be an option to choose among these options in any particular case. Often the possibility of establishing a scheme is seen by a group at one of these levels, and the organisation of the scheme will be left to be carried out at that level if it is done at all. However, at any time there may be the chance to make such a choice or to make a change in organisation of an operational scheme. It is therefore important to consider the properties of all such types of system.

7.1 Government schemes

Many performance recording systems have been run by government agencies (including universities and other research groups as well as government agriculture departments in this class) because employees of these agencies have seen the need for such a scheme and have had the necessary expertise to conduct an analysis of the possibilities and devise methods of running performance recording. Schemes of this kind have generally been motivated by a desire to increase the value of animal production for the public benefit, whether the public has been seen primarily as the general consumer or the animal production section of the economy.

Such a scheme will usually have good technical quality, and will address what are seen by the agency as the important issues for improvement. The direct costs of such a system to the breeder may well be reduced because much of the input in operation of the scheme is paid from general funds. There will often be indirect subsidy also from research programs carried out to estimate parameters or design better selection programs which are not funded by the industry. Since the agency will have no direct stake in the results of the program the system will have a built-in lack of perceived bias.

Despite the high technical quality of such schemes they often have drawbacks. In the first place they are not seen as being “owned” by their users but as being imposed on them. Secondly they may turn out to be rather inflexible, since once a scheme is in place it may be bureaucratically difficult to alter it, though this difficulty may be faced in other situations. A third problem which has been noticeable in recent times is the financial problem raised by governments which are cutting spending. A scheme run by a government agency or a university which has even temporary money troubles may be in danger of closing down.

7.2 Industry schemes

Schemes run on a breed or an industry basis share many of the properties of a government scheme and may indeed be closely associated with a government agency, in the sense that perhaps the industry funds a great deal of the program, but the system is managed by government employees who perhaps use the data collected as research material. However, it is also possible that the industry retains complete control of the management of the scheme, employs the experts who manage the system and perhaps may fund research in outside bodies, or more rarely conduct it in-house. If the industry scheme is large enough, it will be possible to employ highly skilled staff and the scheme can be designed to meet the needs or desires of the industry. In this case there is unlikely to be great concern about the “ownership” of the scheme, and breeders are likely to have direct input to the decisions regarding the structure of the scheme, and this ought to make it easier to iron out problems which are discovered over time. On the other hand, when consensus must be reached and there are very different opinions held by breeders, the reaching of consensus may become virtually impossible and breeders may feel that they have been locked in to a scheme with which they are dissatisfied. When a scheme is run by a breed association it may not be trusted by others since it may be seen as a part of the promotion program of the association. But this need not prevent it serving its members well, if there is no perception that it is designed to serve the interests of some breeders more than others.

An industry controlled system is not likely to be suddenly discontinued because of decisions made for reasons unconnected with the worth of the system and therefore may be more secure than a government agency system. Provided it has a large enough base of users it can support a group of professional officers to run the scheme and contract out research to improve it. This is then a particularly valuable option when individual breeders cannot design and implement a scheme without support.

Another virtue of an industry or government scheme is that data collected from different breeders' populations can be combined and there may be then the possibility of across flock or herd genetic comparisons.

7.3 Individual breeders

For a large breeding company the establishment of a performance recording scheme for its own use will be straightforward and essential. The technical staff will have the necessary expertise to develop an appropriate breeding program and the measurement procedures on which it is based. For an individual breeder the situation is likely to be more difficult unless the measurements are few and simple. Even the basic statistical methods employed may be unfamiliar, and some assistance with technical problems is likely to be required. Often such assistance can be provided by a government extension service or by an agricultural management adviser. This will be especially the case when the breeder wishes to use a system which resembles those used by other breeders, because there is likely to be an "off the shelf" solution known to the adviser. In the past it has generally been possible, at least in many countries, for a breeder to acquire such assistance free or for nominal cost, and indeed the cost of learning to set up a simple system should not be high.

8.0 Data processing and analysis

The amount of data collected and the complexity of the analyses to which it is subjected can vary over a very wide range, from essentially no processing or analysis to the establishment of a national database and the analysis of the data with extremely complex statistical models. Beyond specifying that the scale of data processing should be appropriate to the application it is not possible to make general statements. Nevertheless there are some points which can usefully be discussed.

8.1 Simple analyses

At the simplest level where selection decisions are made immediately after performance has been measured and there is no further use for the data no permanent storage of records may be needed, as in the example of the technique proposed by Tallis (1961). This is an extreme case and nearly all applications will call for some data storage and some form of adjustment of raw data. This can be as simple as writing data in special books and then using a small calculator to make adjustments for effects such as age of dam or age at weaning. In these cases it will usually be best to have a set of standard corrections to apply, rather than estimating them from the data, although the standard values may be only roughly appropriate in a particular herd or flock. As emphasised previously, if there are different management groups there is no alternative to estimating their effects from the data, but this may mean no more than summing values and dividing by the number of animals to obtain the mean which can then be subtracted from each individual record. While these are simple operations they do introduce many opportunities for errors to be introduced, and if data processing is to have this form there is a strong reason to make only those calculations which are absolutely necessary.

With the increase of availability of computing power at low prices, there is now much greater likelihood that a simple analysis will be carried out by computer. For simple problems this can usually be done using a spreadsheet or a database program supplied with the basic operating software of the computer. Even a very simple spreadsheet program, some of which are freely available, can easily be used to make standard corrections, calculate means and deviations from means without requiring much expertise in computing. Such spreadsheet programs also include sorting facilities and it is then possible to sort the animal records in the order desired to make the job of locating the best animals simpler. At a slightly more complex level, it is possible to estimate regression corrections with a spreadsheet as well as use the simpler types of correction mentioned above. The advantage of a spreadsheet or a database program over a dedicated performance recording program written in (say) FORTRAN or C^{++} is that when a computer is in use on a farm it will generally be used for management with spreadsheets and/or databases so that the user will be familiar with these programs and will feel at home using them for other applications. Even users with no special computing skills often learn to develop such applications relatively easily. Should the development of a performance recording spreadsheet application be beyond the capability of the end user there are many people competent to set up a spreadsheet when given the specifications. Modification of a spreadsheet application should also be simple if changes are later introduced. Similar remarks apply to database programs but spreadsheets seem to be more commonly in use.

For the reasons given above it seems sensible to consider first when a simple performance recording system is being established whether a spreadsheet can be set up to handle the calculations. If not, then a special program will need to be developed. The actual calculations will form probably the easiest part of the program to construct, with most of the effort being devoted to ensuring that data input is simple for the user, that checks for obvious errors are included, and that the output is easily interpreted. An advantage of such a system is that it can be precisely tailored to the application and thus be more efficient in terms of time and space usage than a general spreadsheet, and can have its interface constructed to simplify interaction between user and computer. However, writing such a program calls for a much more highly skilled programmer, and will be correspondingly much more expensive to produce. The production of a custom-built performance recording program rather than use of a spreadsheet is then justifiable only when it will be widely used. It would certainly not be likely to be customised for each user, though it may be written in such a way that a number of options are available.

One advantage of a pre-compiled program is that unlike a spreadsheet it cannot be altered by the user and thus cannot have new bugs (those not inadvertently put there by the programmer) introduced by the user, and thus should be more stable and provide for easier troubleshooting when this becomes necessary.

8.2 Complex analyses

When complex analyses are to be done as part of the performance recording system, there seems no alternative to the use of specially written dedicated computer programs. Such programs may be highly specific, resident on a mainframe computer, and run centrally by a group who also monitor the data input and maintain the database on which the program relies for its operation. This is the approach taken by many national dairy cattle breeding programs throughout the world. Because the data which is needed for genetic evaluation is spread over many herds and the selection decisions (choice of proven sires) is not made by the people who actually record the data which is used to make the decisions, a central facility is needed to collect and combine the data, which is voluminous and therefore has stringent storage requirements. In addition the processing of such data is computationally demanding. Despite the amazing increase in computing power in recent years it is still not feasible to provide the full amount of information which is often desired, such as multitrait BLUP evaluations with their accuracies.

An alternative approach is the use of a fairly general program such as PEST (Groeneveld, Kovac and Wang 1990) which incorporates the capacity to carry out very complex analyses, but is flexible in the sense that it is not dedicated to a particular problem. For smaller scale systems use of such a program is to be preferred to development *de novo* unless there are good reasons for a new development (problems not covered by the available packages, wish to produce and market a new program with new features,...) because a program which has already been widely used will have a number of bugs already located and fixed, and there will also be a pool of expertise available through existing users of the program if difficulties arise. The writing, documentation and maintenance of such a computer program is a major effort and such effort may be much better invested in other ways of improving the breeding program.

Though large scale systems have important advantages they also have some significant weaknesses. In particular, quality control of data is much harder to guarantee when the data is collected by many different people under very different conditions. Experience shows that such matters as definition of management groups can be seriously mishandled, that measurement processes can be misunderstood and escape detection more easily that when a scheme is controlled by one unit. In large schemes too the opportunity may arise for some members of the system to exploit others by falsification of data, and although such fraud will probably be discovered in the medium term it can do harm in the short term. Incorporation of automatic data snooping routines in the processing program can help to locate suspect items, but positive identification of fraud is very difficult. Innocent errors may be just as serious as fraud and may be even harder to discover because no pattern is discernible.

9.0 Models

It is becoming increasingly common for the estimation of breeding values to be based on animal model BLUP, using multitrait models where feasible and single trait models otherwise. These methods are known to give the most accurate estimates of breeding values when the data are analysed using models appropriate for the data. Selection based on such EBVs may not give the greatest response in the long term and it is not at present generally agreed how the breeder should cope with the conflict between genetic progress in the short and long terms. It has been suggested that the estimation of breeding values should be biased in order to get a trade-off between short and long term gains, but others feel that deliberate mis-estimation of breeding value can hardly be the best way to proceed. In principle the balance between short and long term response can be reached by discounting future gain to present value, and in a simple case this has been used to find an optimum selection intensity (James, 1972). One difficulty with using biased estimation of breeding value is that in general the production of breeding value estimates and the making of selection decisions are separate activities unless the recording scheme is run by a single breeder or company. For example, in a national dairy cattle breeding program selection decisions are made independently by many different people, and national evaluations are also used in other countries, so that deliberate introduction of bias may have many undesirable consequences. Of course, if the bias is known then it can be allowed for.

More likely is the unintentional introduction of bias through the use of a model which is wrong in one or more of its assumptions or parameter values. Thus if a trait is affected by maternal effects but these are not included in the model used for estimation there will be biases in the estimated breeding values for direct genetic effects. Nevertheless it will usually be desirable to keep a model relatively simple, even at the risk of introduction of some bias. If the effects omitted from the model do not have large effects the estimates of these effects may have standard errors of comparable magnitude to the effects and the random errors introduced by the erroneous adjustments may cancel out the increase in accuracy obtained by removal of the bias. For example, the effect of age at weaning on weaning weight often varies between management groups, and so it would seem sensible to estimate the correction separately for each group. However, if the groups are small, and especially in groups with small ranges in age, the regression estimates will have substantial standard errors. Large deviations from the average regression may then be mainly caused by sampling error. The larger the apparent discrepancy from the average regression, the less attention should be paid to it when the discrepancies are due to sampling. But these are precisely the cases in which the effect of estimating the regression will appear to make the biggest improvement. This is analogous to the situation in fitting a linear model when the null hypothesis is true. Making adjustment for any estimated effect will then introduce only random error. But the corrections which appear to be most significant and to make the biggest improvement in accuracy are those which introduce the most error. It is more complicated when some of the effects do have an effect and some do not, but there are difficulties in using too complex a model.

Animal model BLUP is known to perform well when all pedigrees are known and trace back to the base population, when all data used to make selection decisions are available and used in the analysis, when the correct model is fitted to the data and when all parameters are known without error. These conditions will never apply exactly in practice, but this is not an argument against the use of such analyses, rather it is an argument for not placing blind trust in the results of the analyses.

10.0 Conclusion

In this document I have tried to outline the main factors to be taken into account in the design of a performance recording system, with particular emphasis on the statistical and quantitative genetic implications of the data collection and data analysis. In the design of any particular system there will be many highly specific matters to be taken into account, from the nature of individual identification methods to the detailed model of analysis of the data. In fact much of the hard work of planning a scheme is associated with these “nuts and bolts” questions, and with the difficult social interactions often necessary to arrange acceptance by breeders and producers of a proposed system. These hard details should not, however, be allowed to obscure the need to consider the general questions which have been considered in this document. If these questions are not considered it is possible that all the hard work done to organise the system and its acceptance will result in a flawed program which could have delivered much greater rewards for all of that work.

11.0 Acknowledgements

Some of the work on this document was done while on leave from the University of New South Wales to which I am grateful for the grant of leave. Prof.W.G.Hill kindly provided facilities at ICAPB for which I thank him.

A few specific computer programs have been mentioned in this document. I believe these to be excellent programs, but do not imply that other programs not mentioned are not also excellent.

12.0 References

Abplanalp, H. 1972. Selection of extremes. Animal Production, 14: 11-15.

Amer, P.R. & Fox, G.C. 1992 Estimation of economic weights in genetic improvement using neoclassical production theory : an alternative to rescaling. Animal Production, 54: 341-350.

Brascamp, E.W., Smith, C. & Guy, D.R. 1985 Derivation of economic weights from profit equations. Animal Production, 40: 175-179.

Goddard, M.E. 1997 Consensus and debate in the definition of breeding objectives. Symposium in honour of Charles Smith. Guelph, Canada.

Groeneveld, E., Kovac, M. & Wang, T. 1990 PEST, a general purpose BLUP package for multivariate prediction and estimation. Proceedings of the 4th World Congress on Genetics Applied to Livestock Production. 13, 488-491.

Groeneveld, E. & Kovac, M. 1990 A note on multiple solutions in multivariate restricted maximum likelihood covariance component estimation. Journal of Dairy Science, 73: 2221-2229.

Jackson, N., Lax, J. & Wilson, R.L. 1986 Sex and selection for fleece weight in Merino sheep. Journal of Animal Breeding and Genetics, 103: 97-115.

James, J.W. 1972 Optimum selection intensity in breeding programmes. Animal Production, 14: 1-9.

James, J.W. 1979 Optimum family size in progeny testing when prior information is available. Zeitschrift für Tierzüchtung und Züchtungsbiologie, 95: 194-203.

James, J.W. 1982 Economic aspects of developing breeding objectives: general considerations. in J.S.F.Barker, K.Hammond & A.E. McClintock Ed. Future Developments in the Genetic Improvement of Animals. Academic Press, Australia, pp. 107-116.

Meyer, K. 1989 Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative free algorithm. Genetics, Selection, Evolution, 21: 317-340.

Miraei Ashtiani, S.R. & James, J.W. 1991 Efficient use of link rams in Merino sire reference schemes. Proceedings of the Australian Association of Animal Breeding and Genetics, 9: 388-391.

Ponzoni, R.W. 1986 A profit equation for the definition of the breeding objective of Australian Merino sheep. Journal of Animal Breeding and Genetics, 103: 342-357.

Smith, C. 1959 A comparison of testing schemes for pigs. Animal Production, 1: 113-121.

Smith, C. 1962 Estimation of genetic change in farm livestock using field records. Animal Production, 4: 239-251.

Smith, C. 1977 Use of stored frozen semen and embryos to measure genetic trends in farm livestock. Zeitschrift für Tierzüchtung und Züchtungsbiologie, 94: 119-130.

Smith, C. 1983 Effects of changes in economic weights on the efficiency of index selection. Journal of Animal Science, 56: 1057-1064.

Smith, C., James, J.W. & Brascamp, E.W. 1986 On the derivation of economic weights in livestock improvement. Animal Production, 43: 545-551.

Tallis, G.M. 1961 An application of non-parametric statistics to truncated selection. Applied Statistics, 10: 77-82.

Turner, H.N. & Young, S.S.Y. 1969 Quantitative Genetics in Sheep Breeding. Macmillan, Australia.

Wade, C.M. 1989 Ph.D. thesis. The University of New South Wales.

Wade, C.M. & James, J.W. 1996 Optimal allocation of resources considering two sexes and selection in two stages. Genetics, Selection, Evolution 28: 3-21.