EXTENSION EXPERIMENTS
THEIR NATURE AND MEANING

P. Kanapathipillai

The Institute will shortly be launching a series of experiments on estates. It is the purpose of this article to explain the objectives of these experiments and their relationship to the experiments at St Coombs. An explanation of some of the statistical terms commonly used is also given.

Why should estates go through all the bother of experiments when they have set up the Institute for this very purpose ? To answer this question, we have to discuss the objectives of field experiments. Field experiments serve two broad purposes. In the first place, they enable a research worker to gain an insight into some scientific phenomenon and, secondly, they help to decide on practical courses of action. It is possible that some experiments may serve both purposes. In general, the experiments at the Institute tend to favour the first type, namely the study of the fundamental relationship of various factors affecting the yield and quality of tea. The essence of this type experiment is that the specialist *deliberately creates artificial conditions* to enable him to track down the phenomenon he is investigating. It is obvious, therefore, that the results of these experiments cannot be used by estates. This does not mean that such experiments are unimportant. On the contrary, not only are they absolutely essential for a sound interpretation of experimental results, but progress in research itself depends on such experiments.

The second type of experiment, however, is of immediate interest to planters. Here, the Institute studies directly the effects of changes in the various factors on yield and quality, without concerning itself with the exact mechanisms leading to those effects. The conclusions reached from these experiments, however, have a serious limitation. They are obtained under the particular set of soil, climatic and factory conditions that prevailed during the experiments. Further, an experimental technique that works well when special attention is paid to it, may be quite unsuited to routine use. The more a particular estate differs from the research station in respect of the above conditions, the less applicable are the results of experiments conducted at the research station. It follows, therefore, that the most valuable recommendations for an estate will come from none other than its own experiments. When the results of these estate experiments are pooled together, it may be possible to discover patterns of response to certain treatments. In such an event, the experiment would also prove useful to the *zone* in which the estate is situated. Finally, the experiments will help to feed back to the Institute the information on how the recommendations fared under a wide range of conditions. The last two benefits are secondary considerations and are merely by-products of extension experiments.

Before discussing the basic concepts in field experiments, it is necessary to explain a few important terms.

Error

The term '*error*' refers to a deviation from an average value. This deviation is a perfectly natural one, and has nothing to do with a mistake. For example, if the average yield of high-country estates is, say, 1000 lb made tea per acre and that of St Coombs is 1400 lb made tea per acre, then we say that St Coombs is +400 lb in 'error' of the mean value. To put in another way, we can say that 400 lb is a measure of the *natural variability* in yield of one of the estates when compared with the average yield of all high-country estates. The totality of all such natural variations gives us an estimate of the variation in yields of high-country estates. In practice, these deviations are not expressed as arithmetic differences but instead we calculate a value called the '*standard deviation*' or '*standard error*', which in essence, measures the same thing.

Why is it necessary to calculate the standard error ? Because the standard error is a measure of the reliability of the mean value. In other words, if the variation in yields between estates is large, then the mean value is correspondingly unreliable. Let us take a hypothetical example to illustrate this concept. Suppose there are two groups of three estates each and their yields are as follows :

	Yields (lb made tea/acre)			
	1st Estate	2nd Estate	3rd Estate	Average yield
Group A	1000	2000	3000	$\frac{6000}{3}=2000$
Group B	1500	2000	2500	$\frac{6000}{3}=2000$

Although the mean yields of both groups of estates are the same, *viz* 2000 lb each, there is a fundamental difference. In Group A, the 'error' deviations are − 1000, zero and +1000 lb respectively. In Group B, they are −500, zero and +500 respectively. Group B is, therefore, half as variable in yield as Group A. That is, the mean value of Group B is twice as reliable as the mean value of Group A. This emphasizes the fact that the mean value by itself does not give any information about how variable the estates were in respect of their yields. Similarly, the variability of fields within an estate will be unknown if only the mean yield of the estate was given. It is, therefore, necessary to attach a '*standard error*' to each mean value as a kind of index of the variability of the original units. A mean value without a standard error would be like a map without a scale.

Probability

When a coin is tossed, it may fall either as a head or a tail. We say that the *probability* of throwing a head is ½, 50% or 50:50. This statement does not mean that exactly one throw out of every two will produce a head ; but, if the reader actually throws a coin say 100 times, then it is *likely* that he will obtain approximately 50 heads. Probability is usually measured on the scale 0 to 1. Thus, in the case of the tossed coin, we can say either that the probability of a head is ½, or that the probability of a head is 1 in 2 ; or $P=0.5$ where P stands for 'probability'.

If 2 million sweep tickets are sold and a person has bought 2 tickets, then the probability of winning a Mercedes Benz Car for that person would be 1 in a million or $P=0.000,001$ (a very low probability indeed) ; or, to put it the other way about, the probability of not winning is $P=0.999,999$ (a very high probability).

Significance

This refers to the *degree of confidence* in the conclusions. It is not a statement of the importance of the conclusions. An effect can be significant and yet unimportant as far as the industry is concerned. Thus, when we say that the difference in yield due to a certain treatment is significant at $P=0.05$, we mean that the evidence is strong enough for us to conclude that the effect was due to the treatment. The chance of this conclusion being wrong is only 1 in 20. Below are listed the conventional levels of probability, together with their interpretations.

Level of Probability	Statistical Expression	Interpretation
P=0.10	Not Significant	The difference in the yield may be due to pure chance
P=0.05	Significant	The evidence is strong enough to conclude that the difference in yield is due to the treatments imposed
P=0.01	Highly Significant	It is more certain that such a difference in yield is due to the treatments.

We are now in a position to discuss certain aspects of these trials. Why do we use small plots, each of about 1/20th acre ? Would not a large plot, ten times the size, give much greater information? It would be conceded that the more nearly equal, in respect of soil fertility and other factors, two plots are, the more accurate would be the conclusions drawn from experiments on them. It is also well known that the larger the plot size, the greater would be the soil variation and, therefore, the less reliable the inferences. Smaller plots tend to be more *uniform* and hence more efficient for experimental purposes.

There is another important reason why an experiment on, say, 10 small plots would give more reliable information than another on only one plot 10 times as large. In the case of a single large plot, we have no means of knowing the extent of soil variation within the plot ; but if this is divided into ten small plots and their yields recorded separately, then the variations in these yields would give an *estimate* of the uncontrollable variation or '*error variation*'. It is this estimate of variation that is eventually used as a yardstick to judge whether the differences in yield between certain plots are, in fact, due to the treatments or not.

Randomization

Why are the treatments randomized ? We have stated earlier that an 'error' is a *deviation* from an average value. If we consistently chose all those estates which gave a yield above the average value, then we would always get positive deviations from the mean value. In such a situation, we say that there is a '*systematic error*' or, in simple language, a bias. In the above case, we are quite aware of the bias. But there may be systematic sources of disturbance of which we are unaware. Randomization can be considered as a means of avoiding bias and overcoming any other systematic error like fertility gradients. This means that any effect not under the control of the experimenter would be fairly shared by the various experimental units. The technique of randomization consists of an adherence to some proper random process like the drawing of lots or the use of a table of random numbers. It does not mean a haphazard allocation of treatments to plots.

Replication

The repetition of a whole set of treatments is referred to as a *replication.* We have seen that several small plots can give us an estimate of the error variation. Replication also provides an estimate of error variation but its main purpose is to increase the *precision* of this estimate.

The principles of randomization and replication occur in one form or another in nearly all field experiments. There are, however, circumstances where one of the two principles may not be observed ; but these are exceptions, where the designs are sufficiently ingenious to be able to do without the aid of these principles.