FIELD EXPERIMENTS: HOW THEY ARE MADE AND WHAT THEY ARE.

SIR JOHN RUSSELL, D.Sc., F.R.S.,
Director, Rothamsted Experimental Station.

ALL farmers are interested in the results of field experiments, while those who are directly or indirectly concerned with the activities of the County Agricultural Education Committees, the management of Farm Institutes or the work of the County Organisers are also interested in the methods of carrying them out.

Looking back on the mass of experiments made by the county authorities, the farm institutes and the organisers since agricultural education began in its present form in 1894, one is struck by the enormous amount of labour that has been put into them. Many of the reports bear unmistakable evidence of painstaking observations and accurate weighings; many, however, do not give as much information as might have been hoped, considering the cost. The object of this article is to show how experiments can be arranged to give the maximum return for the work and money spent on them.

Essentials.—The first essential is to frame a perfectly clear idea of what is expected from the experiment. An experiment is simply a question put to nature in the hope of discovering some secret. Even in the best-planned experiments, the answer can usually be only " yes " or " no," while, if the experiment is badly planned, no answer can be given, and much of the labour and expense are wasted. An agricultural investigation is like the old game of clumps in which a person who has left the room returns to discover, if he can, something that has been agreed upon in his absence by those who remained behind. He may ask any question and will obtain the answer " yes " or " no "; if his questions are well enough framed, and if he can interpret the answers, he can always expect to guess the secret. But if the questions are badly framed, neither yes or no can be given as the answer, and in consequence no information is forthcoming.

A committee or an investigator considering a scheme of experiments should first look critically at the plan of the experiments, considering it as a series of questions, and ask whether each experiment or question is framed in such a way that a definite answer can be given. The chief requirement is simplicity: only one question should be asked at a time. For example, if one wishes to know whether basic slag and kainit would be a better dressing for grassland than superphosphate and muriate of potash, it would not be sufficient to put down two plots, one treated with slag and kainit and the other with superphosphate and muriate. It is true that a difference in weight of herbage might be obtained, but the knowledge would be of no practical use because a different result might easily be obtained on the next field, or on the same field in the next year. The information cannot be interpreted because the question is not simple. It is really made up of two: is superphosphate better than basic slag? Is kainit better than muriate? The plan of the experiment must therefore include these tests.

It is obvious that no experiment can be properly planned without some knowledge of the subject. The experiment is to be an excursion into the unknown, but it must start from a basis of ascertained fact. Any child can ask a question that cannot be answered, but it requires a skilful and intelligent person with knowledge of his subject to ask a question that admits of a clear and unambiguous answer. The preliminary knowledge has usually to be obtained in some other way; often in the laboratory by scientific investigation, and this indeed is one of the reasons why an experimental station, such as Rothamsted, must be furnished with well-equipped laboratories. The purpose of the work is to obtain the knowledge with which the field experimenter must start before he can frame a clear question or devise a field experiment that will give a definite answer.

Interpretation of Results.—Having drawn up the simple question and carried out the work carefully, there arises the problem of interpreting the results. Interpretation is usually more difficult than devising the experiment, and much more difficult than making it. There may be no dispute about the facts, but considerable dispute about their meaning. Scientific controversies are often numerous and long-continued about the interpretation to be put upon recognised facts. In field experiments there is the difficulty that the result may not be due to the treatment, but to something entirely different. A famous fifth century writer on agriculture, Palladius, gives as a preventive against hail that the farmer should walk round the outside of the field carrying a tortoise upside down and laying it on the ground at each of the corners of the field. It is quite possible that the first man who did this escaped damage; the error lies in connecting the two events. Two agencies lie outside the plan of field experiments, and may cause a result quite independent of the experimenter's efforts—the weather, and soil variations.

Three Desiderata.—In planning field experiments or studying the results to see what information they will yield, three important considerations have constantly to be kept in view; the experiment must be simple and definite and capable of giving a clear "yes" or "no" answer: it must be based on knowledge that is trustworthy as far as it goes; and the results must be interpreted if they are to be of wide use to farmers. The interpretation may well lead to a good deal of discussion, indeed it is usually a mark of a good investigation that it does give rise to controversy.

Results Obtained.—A well-planned field experiment often gives more information than its designer anticipated. A classical example is the Broadbalk wheat field which was laid out in the first instance to ascertain the manurial requirements of the wheat crop. Stated in this form the problem is too indefinite for field experiment, but the genius of Lawes and Gilbert reduced it to three simple questions. They started out (as we have seen is always essential) from certain definitely known facts, viz.:—

(1) That farmyard manure is an excellent manure for wheat.

(2) That farmyard manure contains three groups of constituents: (*a*) ash constituents, (*b*) nitrogen compounds, (*c*) organic matter, and they proceeded to ascertain which of these were effective for wheat. Their plan of experiment was as follows: four plots were measured out on Broadbalk field; one was left without manure; a second received farmyard manure at the rate of 14 tons per acre; a third received the ash constituents of farmyard manure at the same rate; while a fourth received the ash constituents together with the nitrogen compounds (actually ammonium salts obtained from gas works). The plots were then sown simultaneously with wheat, and the resulting crops were weighed; the results were as follows:—

Produce of Wheat per Acre. Broadbalk Field, Rothamsted, 1844.

	Yield per Acre.	
	Grain.	Straw.
No manure	16 bus.	1,120 lb.
Farmyard manure (14 tons per acre)	22 ,,	1,476 ,,
Ashes of 14 tons of farmyard manure	16 ,,	1,104 ,,
Ash constituents + a nitrogen compound (ammonium sulphate) up to	26¼ ,,	1,772 ,,

Comparison of the yields from the first and second plots showed the effect of farmyard manure; this was the standard against which the other yields had to be measured. The third plot put the question whether the ash constituents were the effective fertiliser; the answer is clearly " no," since the yield is no better than that on the unmanured plot. The fourth plot put the question whether the ash constituents *plus* nitrogen compounds were the true agents, and the answer is " yes "; it was made more definite by further experiments which showed that nitrogen alone was not as effective as nitrogen *plus* ash constituents. Lawes and Gilbert drew the immediate conclusion that these substances could take the place of farmyard manure, and Lawes set up a factory at Deptford to prepare them on the large scale. The experiment was one of the foundations of the artificial fertiliser industry and might therefore have been regarded as finished. It was, however, continued by Lawes and Gilbert, and also by the subsequent directors, first Sir Daniel Hall and then the writer, and there is little doubt that future directors will do so for the very sufficient reason that it continues to give useful information. The plots have been found very valuable for studies in soil chemistry; microbiology; soil physics; for important field problems connected with drainage, draft of tillage implements, etc.; they have helped considerably in studying problems connected with the cultivation of the prairie lands of Canada and the United States, especially the very difficult problem of the rapid loss of nitrogen when the prairie is first broken up; and it seems improbable that a time will ever come when the experiment can be described as finished. It was started in 1843 and maintains its value as the years go by.

Carrying out a Field Experiment.—We must turn now to the consideration of the way in which a field experiment should be carried out, assuming the plan is satisfactory.

I.—The first method used was the side-by-side arrangement of plots familiar to those who know the Broadbalk field. There were single plots only for each treatment, but the experiment was repeated year after year on the same ground, *i.e.*, repetition in time though not in the field. Lawes and Gilbert published their results after a period of 20 years, though, as a matter of fact, substantially the same conclusions can be drawn from the first five years' results.

The method has the drawback that it takes no account of variations in the soil. Had the plot receiving ash constituents and nitrogen compounds been inherently less fertile than the rest so that it yielded no better than the unmanured plot, Lawes and Gilbert would have concluded that the organic matter was the essential fertilising ingredient, and the discovery of the great value of artificial fertilisers might have been delayed. They recognised the difficulty about soil variation, and repeated the experiment at certain other centres; at Holkham in Norfolk; at Rodmersham in Kent; and after many years at Woburn.

Admitting this weakness in the method, its advantages are that it is the simplest of all arrangements for field work, so simple that it can easily be carried out on a commercial farm, and it lends itself easily to demonstration, especially when the plots can be arranged end-on to a road so that farmers can assemble to see the results. Yet it can only be recommended :—

(1) Where the experimenter knows pretty well what result he will get, and it is reasonably sure that the differences will be visible to the eye (usually the difference must be 15 to 20 per cent. to be visible, and preferably more to make a good demonstration);

(2) Where the experiment must of necessity be carried out on a commercial farm and it is impossible to arrange for more than the minimum number of plots.

In this case the experiment must be made on uniform lines, at a number of centres with fairly large plots, on fields known to the farmer as being uniform. But it must also be made at the same time at one or more College or Research Station Farms properly replicated as described below. These form the " key " experiments from which the critical information is derived; the commercial farm experiments can be interpreted only in the light of the information they give.

An example of this method of experiment is furnished by the investigations on malting barley carried out by Rothamsted as part of the Institute of Brewing Research Scheme. The purpose of the experiment is to ascertain the influence of manures on the yield and malting quality of barley. It involves testing the effect of a complete mixture of artificial manures and of the mixture without nitrogen, without phosphate and without potash respectively. At Rothamsted the experiment is fully replicated, and also, though to a less extent, at Woburn, but it is an essential part of the scheme that it should be repeated by a number of commercial farmers *known to be good barley growers*, so that their observations and those of their barley buyers should be obtained. On these farms it is impracticable to have more than a few plots; five can be properly looked after and fairly accurately weighed, but 10 or 15 would be impossible. Five single plots of an acre each are therefore used; the seed for all centres is the same, as also are the manures; all are sent out from the central supply. On each farm the scheme is repeated without change for three or four years, the barley coming in its proper place in the rotation. In addition, at one of the more interesting centres (Wellingore) it has been found possible to repeat on the same field for three years. Worked in this way on a uniform basis for a period of four years, with the "key" experiments at the Research Station, the method gives good results, some of which are summarised below:—

DECREASE (—), OR INCREASE (+), PER ACRE DUE TO OMISSION OF CERTAIN FERTILISERS FROM THE COMPLETE MANURE.

Fertiliser omitted from complete manure.	*After a Straw Crop.*	*After Roots fed off.*	*After Potatoes or Beet (well manured).*	*Mean of all Experiments.*
	Bus.	Bus.	Bus.	Bus.
1 cwt. sulphate of ammonia ...	— 5.8	— 3.9	— 6.7	— 5.4
3 cwt. superphosphate	— 0.9	+ 0.5	- 1.2	— 0.5
1½ cwt. sulphate of potash ...	+ 1.1	- 1.3	— 1.1	— 0.3

On the basis of these experiments it would be safe to tell a class of students or a body of farmers in a lecture that an increase of some 4 or 5 bushels of grain could be expected as a result of applying 1 cwt. sulphate of ammonia and that the Table of Valuations shows that no reduction in value per quarter need be feared; that neither superphosphate nor potash would generally increase either yield or value per quarter to any important extent. Yet it would not be safe on this basis to advise an individual farmer on his own farm unless one happened to know the farm; the direct test would always be desirable.

II.—Lawes and Gilbert devised a second lay-out for field experiments in which one set of treatments is applied in one direction and the other set in the cross direction, so giving a number of rectangular plots arranged like a chess-board. (Fig. 1.) This method was used for the Hoos barley experiments begun in 1852. It is more compact than the side-by-side arrangement and is therefore more trustworthy, but it is not well suited to demonstration to large parties of farmers. It

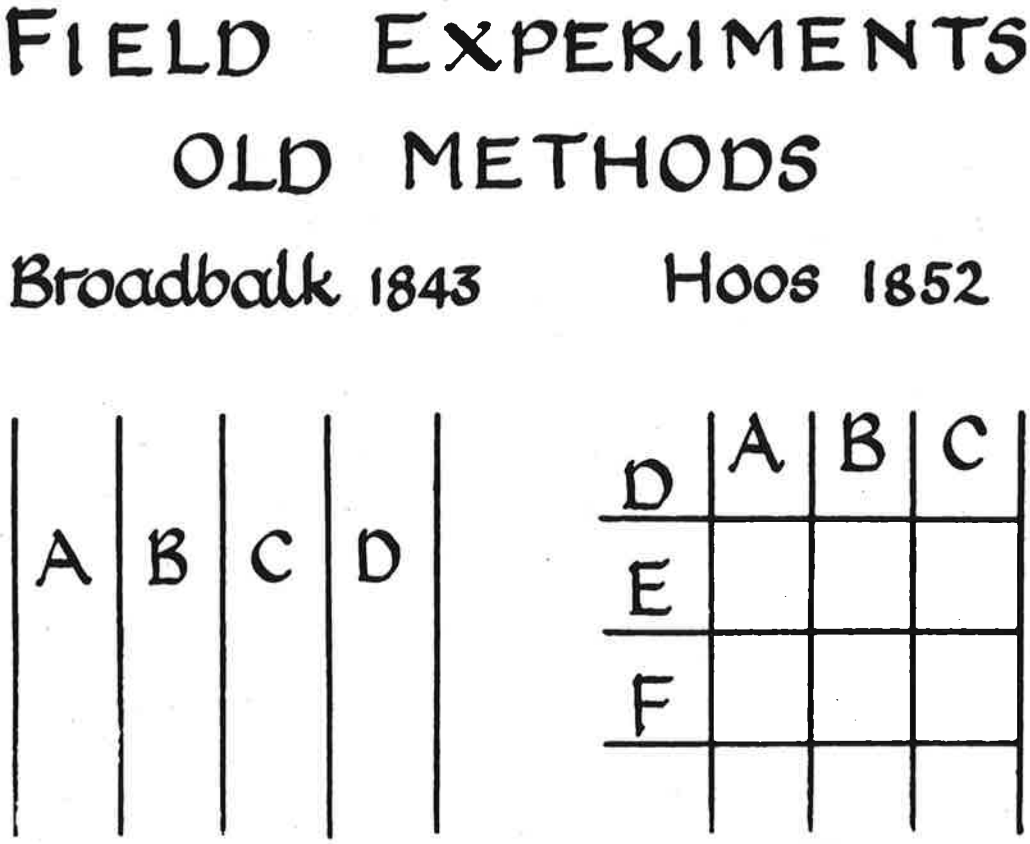

FIG. 1.—Side-by-side and Chess-board arrangement of Field Plots.

involves the danger that the experimenter may put in too many treatments; he may think it quite easy to test 16 or even 20 schemes; actually no experiment should involve more than four or five. A useful type of experiment on this model is to test the effect of fertilisers on different varieties of the same crop. The different varieties can be set along the strips, A, B, C, while the manurial treatments are given along the cross strips, D, E, F. The set needs to be repeated without change for several years, and on several fields or parts of the same field.

The Hoos barley plots, like those of wheat, are repeated year after year on the same ground, and so the variations due to season can be allowed for, but not the variations due to soil. These give an element of uncertainty which no length of time of continuance ever quite removes. Variations in soil can be overcome only by repeating the experiment on the same field at the same time. This is now well recognised, and duplicate experiments have long been the rule. There is, however, one important point about duplication. The duplicate plots must not follow in the same order as the first set. If two treatments or varieties which we may call "A" and " B " are being compared, it is not sufficient to arrange them alternately thus:

A B A B A B etc.

For the A's are always to the left of the B's, and will always come out better if the fertility of the land is falling off from the left to the right of the plots, or worse if the fertility slope runs the other way. Several instances could be quoted from published reports where plots so arranged have given misleading results: A was pronounced better than B, and the repetitions made the result look true, but in truth A was no better than B and the result was due to a difference in fertility.

The proper way to arrange a comparison between two treatments is to arrange them on a balanced plan, thus:

A B B A A B B A

Here each treatment is compared with itself on one side and the other treatment on the other side, and the plots are equally balanced about the centre. Whichever way the fertility may be varying the comparisons can still be made. This is the method used first by Dr. Beaven and now by the Institute of Agricultural Botany for testing varieties of crops; it is often called the half-strip drill method because half the drill carries seed of variety A and the other half carries seed of variety B; as the drill goes up and down the field, it sows strips as shown above. The difference in yield between one variety and another of the same crop may be as little as 5 per cent.; this could not be detected by any single-plot method nor by repetitions on alternate plots; it can, however, be shown on this balanced method.

III.—For manurial trials it is commonly necessary to have more than two treatments; usually five are required. The plots can still be arranged on a balanced plan as shown in Fig. 2. They are grouped round a central plot so that the distances from the centre of the plots under each treatment when added up are the same both on the left and on the right side. Thus, in the diagram B is the centre (it is marked 0). Under treatment A there are two plots to the left, distant respectively 1 and 6: total 7: one plot to the right distant 7. Under treatment C there is one to the left distant 7; and two to the right distant 2 and 5: total 7. Under D there is one to the left distant 5 and two to the right distant 1 and 4: total 5. This method overcomes some of the difficulty of soil variation, especially the small steady changes imperceptible to the eye.

The plan can be simplified by using a cross dressing. An instance is afforded by the experiment designed by Mr. R. A.

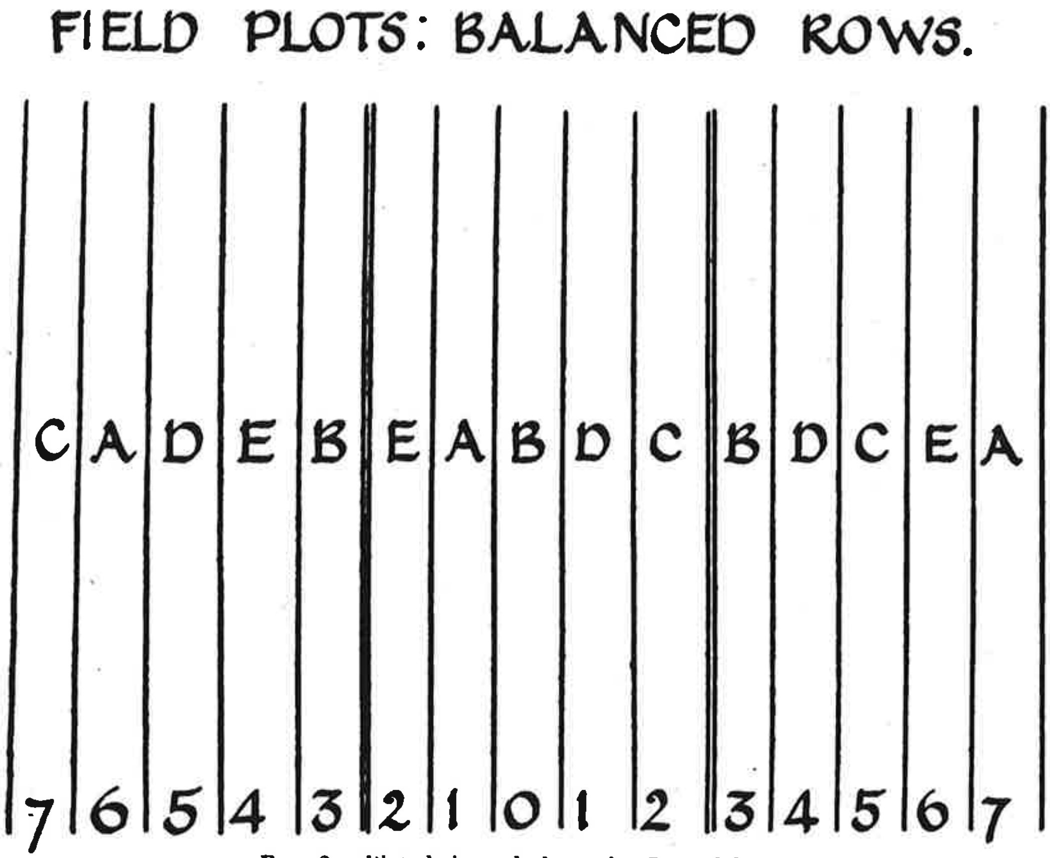

FIG. 2.—Plots balanced about the Central Strip 0.

Fisher for a detailed study of the effect of phosphates and of nitrogen on crop yield. Considerable accuracy is necessary because the data are to be used for studying the influence of the weather on the effectiveness of the manures. The plan is shown in Fig. 3. The whole area receives a potash fertiliser; the influence of potash cannot in this particular experiment be studied. The area is divided into "balanced" strips for the phosphate dressings, then it is cross-divided into two; the opposite corner groups receive nitrogenous manures, the others do not. The set involves 16 plots, but the agricultural operations can be managed without much difficulty.

IV.—A still better plan would have been to arrange the nitrogenous manures also in "balanced" strips similar to those used for the phosphate, but this would have necessitated 64 plots, an impracticable number for this particular experiment.

V.—Another modification consists in arranging the triplicates not side-by-side as in Fig. 2, but chess-board fashion as in Fig. 4. The numbers on the plots represent the various treatments. It will be observed that they "balance" about a line

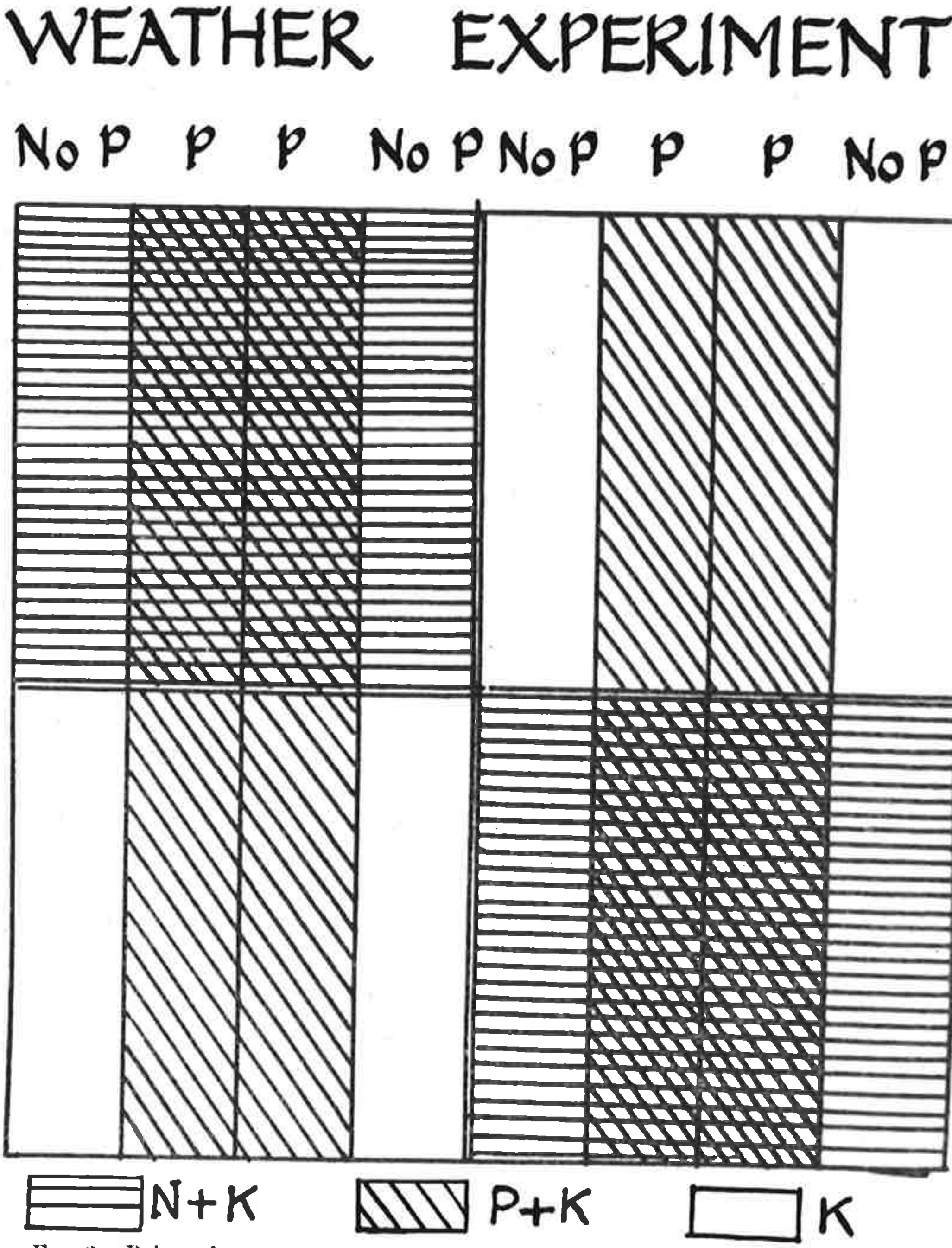

FIG. 3.—Balanced arrangement of Plots for studying the Influence of Weather on the Effectiveness of Phosphates and Nitrogen.

drawn across the middle. The shading of the plots shows the fertility variations as revealed by a "uniformity" study of the figures. The mean of the triplicates shows an increase of 7.8 bus. per acre for the use of 1 cwt. sulphate of ammonia, no appreciable result from the use of superphosphate, and a gain of 5.4 bus. per acre by omitting potash from the manurial

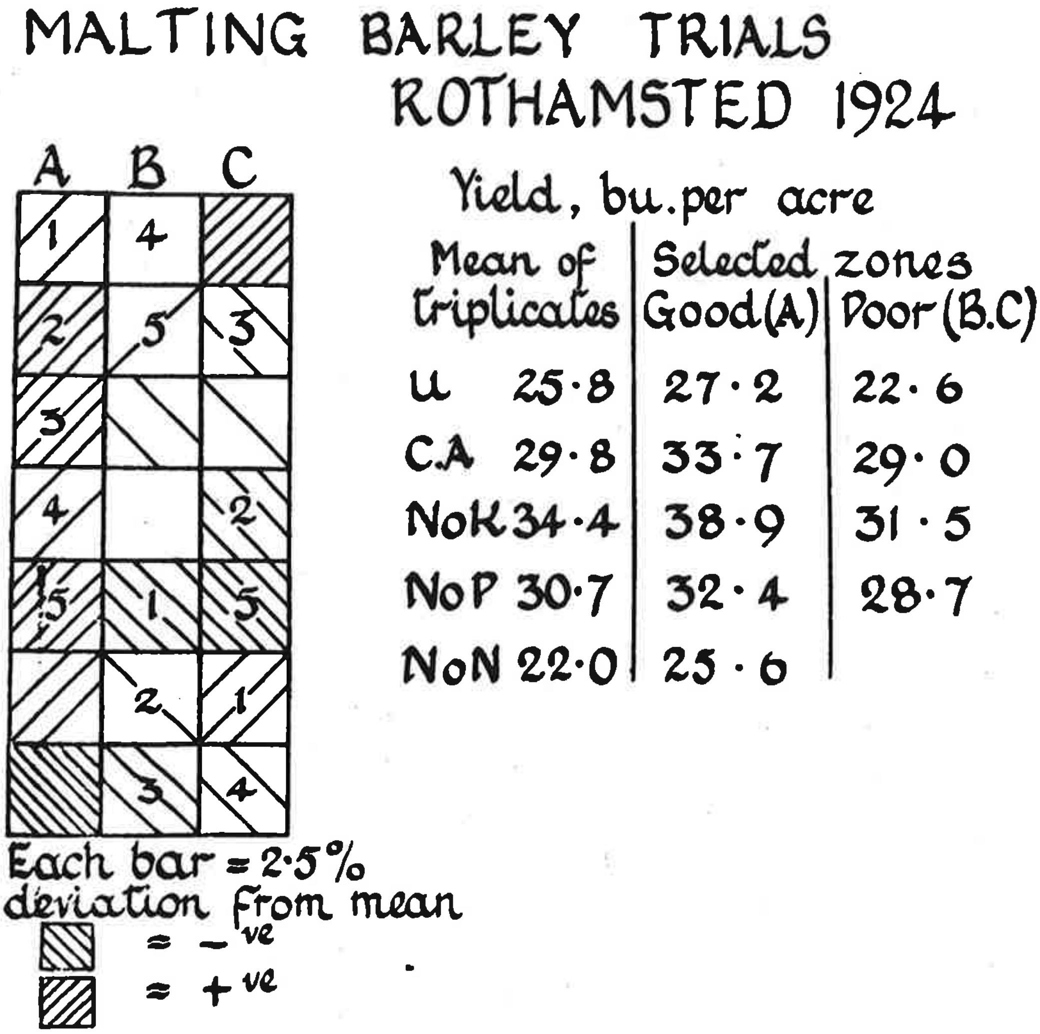

FIG. 4.—Balanced Plots on Chess-board Plan.

dressings. This last result is very interesting: it is obtained if one studies the whole group of plots, or if one considers only the best plots or only the worst; the chances are very remote that it is due to some accident such as soil variation. It is not obtained every year, and is not therefore a regular behaviour of soil or crop, but some peculiarity of the season. The experiment gives no information as to *why* this happened, but shows clearly and certainly that it *did* happen. Laboratory investigations are needed to discover the causes at work; until these are known it is impossible to predict when the result is likely to happen again.

The "balancing" of the strips has the great advantage that it very considerably reduces the errors due to irregularities of the soil. With single plots it is difficult on one year's trial to speak confidently of a difference of 10 per cent. in the yield, though if the same result is obtained for three or four years it becomes more certain; but with replicated "balanced" strips the results have a much greater value even in one season.

VI.—A further refinement is now being introduced at Rothamsted in consequence of the investigations of Messrs. R. A. Fisher, T. Eden and E. J. Maskell. The strength of the balanced strip method is that it reduces to a minimum the errors due to soil variation. Its weakness lies in the fact that the errors, though certainly small, are not definitely known; they cannot be calculated accurately. The investigator desires not only to minimise his errors, but to know how big they are. The amount of error, or rather the probable amount, can be calculated, but the calculation assumes that there has been no adjusting of the figures or selection of the ground for a particular plot; everything must have been left to chance. In practice this is impossible and a certain amount of selection is necessary; a compromise has to be made between what is desirable and what is practicable. The best practicable arrangement is to have as many repetitions as there are treatments, to set the plots out in chess-board fashion, but arranged so that no two of the same kind come in the same column or in the same row; the arrangement is called a "Latin Square": an example is as follows:—

A B C
C A B .
B C A

For a manurial experiment with five plots there are no fewer than 1,344 possible arrangements. (Fig. 5.)

In laying out an experiment on these lines the fact that no treatment is repeated on any one row or column gives all the advantages of the "balanced" strips. The fact that there are 1,344 ways of arranging the plots within the square allows ample play to the laws of chance. For the investigator does not himself choose which of all these ways he will have: each arrangement is written on a separate card, the pack of 1,344 cards is shuffled and one chosen at random; this is the arrangement adopted. As a still greater refinement three or four are chosen and all are used.

Obviously the method requires a considerable number of plots. Its use at Rothamsted necessitates special arrangements for harvesting, thrashing, weighing and recording, which, however, are too intricate to be dealt with here. The advantage

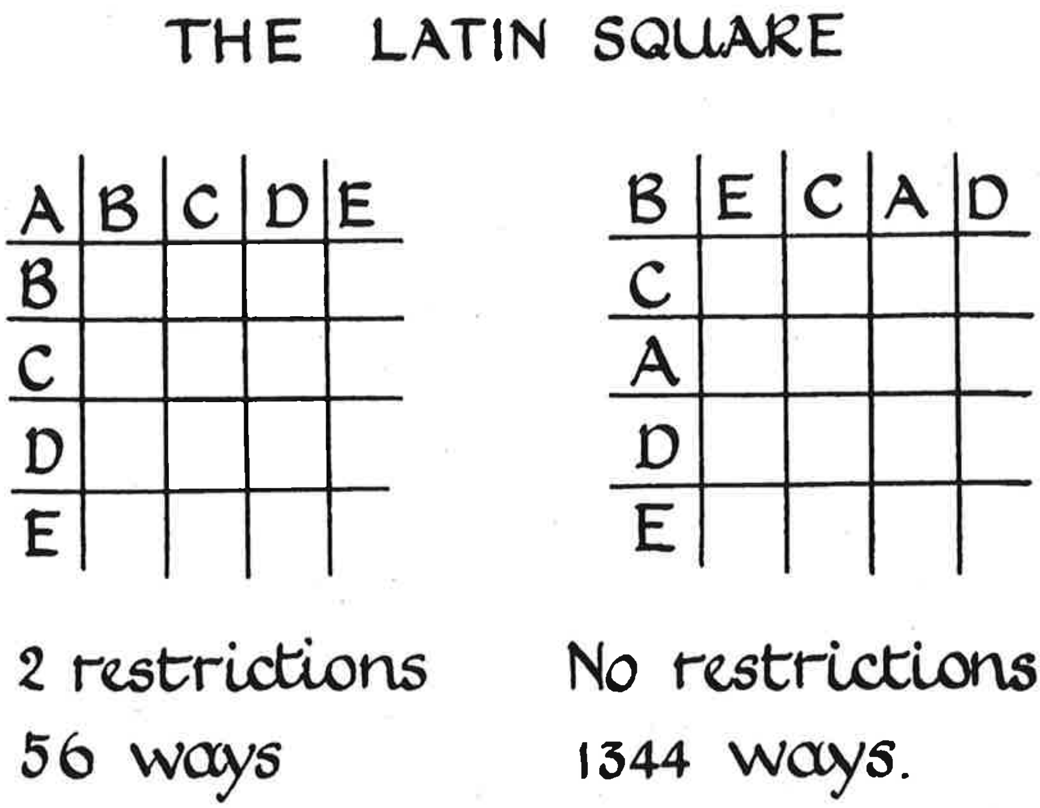

FIG. 5.—The best arrangement known at present for Testing 5 Treatments.

of the method is that the errors are reduced to a very small quantity, and that quantity can be calculated so that the statistician can apply proper statistical methods to the treatment of the results.

The planning of a field experiment which is intended to yield new knowledge is obviously no easy matter. The staffs of the Rothamsted Statistical and Field Experiments Departments are always ready to discuss experimental schemes with Organisers and College Lecturers who wish to carry them out with a view to increasing the value of the experiment without detracting from its practicability.

• • • • • •