THE VALUE OF SUMMARIZING RESULTS FROM INDIVIDUAL EXPERIMENTS

by

Walter T. Federer
Professor of Biological Statistics
Department of Plant Breeding
Cornell University, Ithaca, N. Y.

A Lecture Presented in the Department of Agriculture Auditorium

March 30, 1953

This is the ninth in a series of lectures sponsored by the Departmental Committee on Experimental Design, and the Agricultural Research Administration. The previous lectures of this series are listed below.

Dec. 14, 1949 - How Statistics Improve Physical, Chemical and Engineering Measurements.
  William J. Youden, Assistant Chief, Statistical Engineering Laboratory, National Bureau of Standards

Feb. 15, 1950 - Sampling Methods in Marketing Research
  Earl E. Houseman, Statistical Consultant, Bureau of Agricultural Economics.

Sept. 16, 1950 - Statistics and Research on Pasture and Grazing.
  Henry L. Lucas, Animal Science Statistician, North Carolina State College

Jan. 11, 1951 - The Value and Usefulness of Statistics in Research.
  Gertrude M. Cox, Director, Institute of Statistics, University of North Carolina.

May 16, 1951 - The Application of Punch Card Equipment to Statistical Processing.
  Lawrence W. Armstrong, Assistant Chief, Machine Tabulation Division, Bureau of the Census, United States Department of Commerce

Dec. 12, 1951 -- A Review of the Principles of Experimental Design and Some Applications
  Oscar Kempthorne, Professor of Statistics, Statistical Laboratory, Iowa State College, Ames, Iowa

Feb. 6, 1952 - Determination of the Size of Experiment and Samples
  William G. Cochran, School of Hygiene and Public Health The Johns Hopkins University

Feb. 26, 1953 - Some Lattice Designs
  Boyd Harshbarger, Virginia Agricultural Experiment Station of the Virginia Polytechnic Institute

Copies of the above lectures are available and may be obtained from:

  Office of the Administrator
  Agricultural Research Administration
  United States Department of Agriculture
  Washington 25, D. C.

THE VALUE OF SUMMARIZING RESULTS FROM INDIVIDUAL EXPERIMENTS[1]

W. T. Federer

Introduction

All too often the data from individual experiments are filed away or are discarded shortly after reduction of the data to a few pertinent summary statistics such as means, variances, standard errors, F values, etc. The fact that data from individual experiments still contain a considerable amount of information frequently escapes the attention of research personnel. Also, our budgetary structure is such that more emphasis is placed upon the individual experiment or project than upon the coordination of results and efforts. Our research men are not, in general, urged or encouraged to use every fifth or tenth year to summarize and to interpret past results. A critical evaluation of the results and of the methods of past experimentation is of considerable value in planning future work. Inadequate planning of future research leads to inefficient use of research personnel and facilities, and frequently to publication of inconsequential results.

The statistical problems associated with summarizing results from a group of experiments have not been completely solved, and statisticians will find this a fertile field for research. Some relatively simple and usable methods are, however, available in the published literature.

The purposes of this paper are:

(1) to illustrate certain types of information that are available from the summarization of results from a group of experiments,

(2) to indicate the changes made in two research programs as a result of a summarization of past results, and

(3) to present a method for summarizing experimental variances.

Examples of Types of Information Derived from a Summary of Results

Examples of the kinds of information that may be derived from summarizing the results from several experiments are given below. The list of illustrations is incomplete, but it indicates the diversity of information that may be obtained from summarizing results from experiments conducted primarily for other purposes. In this connection, experiments should be designed to serve more than a single purpose. For example, the individual experiment may be concerned with the effect of fertilizer treatments on the yield of grain, but the group of all such experiments may be used to investigate some aspect of the experimental technique.

(i) Loss in information due to subsampling the field plot.

Yates and Zacopanay (24) investigated the loss in information due to harvesting only a fraction of the plot relative to complete harvesting.

Thirty-three experiments on cereals were used in the investigation. The individual experiments were set up to compare the effect of various treatments on the yield of wheat, barley, and oats. Since the plots were large subsampling procedures were used to obtain a sample for threshing. The subsampling rate in the experiments averaged six per cent of the total plot area, and the loss in information relative to complete harvesting was 31.2 per cent. This means that 50 per cent more replicates would be required to attain the same degree of precision with a subsampling rate of six per cent as with complete harvesting. With subsampling rates of 12% and 18% the losses in information due to subsampling relative to complete harvesting were 18% and 13%, respectively. Yates and Zacopanay illustrate the procedure and the gains in information obtained when use is made of total unthreshed weight per plot. Also, certain sampling schemes and methods of stratifying the plot are more efficient than others (14, 24). Under some cost situations it is more advantageous to use additional replicates with subsampling than to use fewer replicates with complete harvesting. Yates and Zacopanay present the procedure for determining the optimum sampling rate and number of replicates for a given cost situation.

In each type of experimental work it is necessary to obtain factual evidence concerning the relative loss in information due to subsampling. Then, in light of cost and practical considerations, the experimenter is in a position to decide the appropriate subsampling rate.

(ii) Estimating a treatment response from a group of experiments.

The same treatments or varieties may be repeated in a number of experiments, and it may be desired to obtain a comparison over the group of experiments. For example, it may be desired to estimate the mean response for various nitrogen applications in a series of experiments conducted over a considerable area and period of time. The responses to nitrogen applications in the individual experiments may be subject to different errors. Cochran (2) and Yates and Cochran (23) describe the procedure for estimating the mean response under these conditions. They obtain a weighted mean response for the different fertilizers from a group of experiments.

The individual experiments may not be large enough to detect small responses whereas the combined set of exeriments may clearly indicate the presence of an effect. Also, the individual estimates of effects may be interesting but rather uninformative with regard to an entire region. For the latter case a combined estimate is desired. In addition to obtaining a mean response, an investigation of the responses from the individual experiments may be informative in ascertaining which experimental situations yield large responses and in observing the suitability of the treatments selected.

(iii) Relative precision of experimental designs.

The results from all experiments may be used to determine the relative efficiency of the various experimental designs used. Prior to obtaining an average efficiency over all designs the experiments should be grouped according to type. If the experimental error variances in each of two designs are denoted as E_1 with f_1 degrees of freedom and E_2 with f_2 degrees of freedom, if the corresponding numbers of replicates are denoted as r_1 and r_2, and if the corresponding costs of conducting the experiments are denoted as c_1 and c_2, then the efficiency per unit of cost of the first design (subscript 1) to the second design (subscript 2) is equal to

$$\left(\frac{r_1 c_1}{E_1}\right)\left(\frac{f_1 + 1}{f_1 + 3}\right) \Bigg/ \left(\frac{r_2 c_2}{E_2}\right)\left(\frac{f_2 + 1}{f_2 + 3}\right) . \tag{1}$$

If the first design is more efficient than the second the ratio in the above equation is larger than one. The cost of experimentation includes the costs associated with the design, the analysis and the conduct of an experiment. In many situations the costs for the two experiments are nearly equal, and c_1 is set equal to c_2.

In certain experimental designs it is possible to estimate the error variance for a design with less complete stratification. For example, it is possible to obtain an estimate of the error variance for a completely randomized design even though the experiment is designed as a latin square. This property allows the computation of efficiency of the actual design with one having fewer restrictions on the allocation of treatments over the experimental area.

Cochran (4) presents a summary of the results from field experiments carried out at Rothamsted and associated centers between the years 1927 to 1934. He found that ten replicates of a completely randomized design were required to give the same precision as six replicates of a randomized complete block. Four to five replicates of a latin square were as accurate as six replicates of the randomized complete block design or as ten replicates of the completely randomized design.

Further work by Cochran (5) illustrated that large gains in information were to be realized from using lattice designs rather than the randomized complete block design for corn varietal trials in Iowa. He found that the precision of the lattice designs tended to increase as the size of the block increased. Cox (6) found similar results from a summary of efficiencies of lattice designs carried out at the North Carolina station.

Such summarizations tend to point up the value of certain designs and the defects of others. Some designs are excellent in certain fields of research and of little use in others. A summary of the efficiency of various designs is suggested for each field of research and for each experimental set-up. It is recommended that at least 30 and preferably 40 to 50 experiments be included in each summary since an individual estimate of relative efficiency is subject to considerable variation.

(iv) Investigation of differential errors in fertilizer experiments.

Kempthorne (16) investigated the individual experimental errors for main effects and interactions from 128 experiments on sugar beets. He found that the effects were subject to the same error variances, and hence, the use of the pooled error in a factorial experiment and the use of confounded arrangements are warranted for fertilizer experiments on sugar beets.

(v) Effect of size of sampling unit on the relative variability in marketing studies.

Professor Max E. Brunk and his associates have conducted a number of controlled experiments in retail grocery stores to compare the effect of various marketing procedures on the sales of apples. The treatments (the various marketing procedures) remained in a store for different lengths of time. Some treatments remained in a particular store for one day while other treatments in another experiment remained in a store for longer periods. The coefficient of variation was computed for each of 24 latin square experiments in which the experimental unit, or the period of observation on a treatment, was one day. The average coefficient of variation was 37.7%. In another group of six latin square experiments with the period of observation being either one or two weeks the average coefficient of variation was 14.3% with none of the experiments having a coefficient of variation as high as the lowest one in the group of 24 experiments. The decidedly larger coefficient of variation associated with periods of observation of one day's duration points up the fact that more replicates are required to attain the degree of precision obtained with experiments in which the period of observation is one week in duration. If the treatments are difficult to change it may be more efficient to use a longer period of observation.

(vi) Estimate of heritability.

The estimates of heritability presented in the literature often pertain to animals reared under the particular experimental conditions associated with experiment stations. The results may or may not be applicable to animals reared by commercial breeders. King (17) summarized the estimates of heritability for various characteristics of poultry presented in the literature. He investigated the heritability of several traits on a commercial poultry flock and found good agreement of the estimates if he ignored the effect of hatching date, but not otherwise. The effect due to date of hatching was large and appeared to be random. Hence, it should not be ignored in computing heritability estimates since a change in these estimates may change the method of breeding employed in selecting for certain characteristics.

(vii) Investigation of results from corn experiments in Iowa

Dr. G. F. Sprague has conducted a relatively large number of experiments on corn each year since 1940. The individual experiments were designed primarily to compare grain yield and other characteristics, but they have been used for various other purposes. Some of the experiments were useful

for estimating the specific and general combining ability of various inbred
lines (15, 21). Since most of the experiments were conducted as lattice
designs Cochran (5) made use of the data to estimate the efficiency of
lattice designs relative to the randomized complete block design.

More recently an investigation (9, 10, 11, 20) of the 302 experiments on
corn conducted in Iowa between the years 1940-1947 was made to determine
the adequacy of the experimental technique. More specifically, information
was desired on the optimum combination of numbers of varieties, replicates,
locations, and years for selecting the highest yielding variety or varieties.

Some of the outstanding results obtained from this study are:

(1) The size of the genetic variation among doublecrosses, topcrosses,
and single crosses relative to the error variation was .52, .82,
and 1.14, respectively. The higher genetic variation associated
with singlecrosses points up the value of selecting among
singlecrosses.

(2) For a given number of plots in a single experiment and for certain
numbers of varieties the largest average genetic advance in yield
is made with two replicates for singlecrosses and with four
replicates for doublecrosses.

(3) The variety x year and the variety x location components of
variance are important sources of variation. The former com-
ponent is consistently larger than the variety x location
component.

(4) The use of a single tester parent to evaluate total combining
ability of inbred lines of corn is an inefficient procedure.
Also, the choice of the tester parents depends upon the
particular objectives of the breeding program.

(5) The initial work (11) on determining the optimum number of
tester parents led to further work (15, 19) which clarified the
whole problem of evaluating inbred lines.

(6) A relatively simple statistical procedure for combining the
ratios of variance components was evolved.

(7) The cost of experimentation is important in determining alloca-
tion of resources. A procedure was developed to take into
account the cost of experimentation.

Although other miscellaneous results were obtained from the investiga-
tion of the individual analyses the above results illustrate the diversity

of information that may be available for a summary of the results from individual experiments.

(viii) Investigation of results from forage crops experiments in New York.

Lowe (18) investigated the yields from 44 perennial and nine biennial forage crops experiments conducted in New York during the period 1945-1950. The purpose of the investigation was to determine the adequacy of the experimental technique and to determine what changes in the breeding program were indicated. The chief results of this study are:

(1) Due to the divergent results from the various forage crop species each species should be treated separately. Also, due to the high correlation between maturity and yields grass strains should be grouped into maturity groups before comparing them for yield.

(2) For perennial grass experiments the highest value for genetic progress was made with three to four replicates and two years of harvest.

(3) For alfalfa experiments the best combination was six to eight replicates with one year of harvest. Since information on persistence is usually desired a portion of the replicates will need to be retained for more than one year. The fact that stands of alfalfa do not become thicker in succeeding years leads to a high year to year correlation of yields. For this reason more replicates and fewer years are recommended for alfalfa strains than for grass strains.

(4) Although grass strains may exhibit considerable variation when grown alone this variation tends to disappear when grass strains are overseeded with a legume. Since weaker grass strains offer less competition to the accompanying legume the yields of the mixtures tend to equalize.

(5) For many grass species the yields of the first cutting (the hay cut) are already acceptable, but the yields of subsequent cuttings (the aftermath) are not. The recommended combination of number of replicates and years of harvest is different for aftermath yields than for total seasonal yields. The yearly variation in aftermath production in New York is quite large.

Changes in the Corn Research Program in Iowa

Dr. G. F. Sprague (written correspondence) lists the following changes in the Corn Research Program in Iowa resulting from the various investigations cited under section (vii):

(1) Reduction in number of replicates: In 1951 the number of replicates was reduced to two. The plot size was 2x10 hills. In 1952 four replicates of 2x5 hill plots were used. This appeared to be a satisfactory arrangement. There is a possibility that a further change will be made to three replicates of 2x5 hill plots.

(2) Change in procedure for testing new lines: In the past, this has been done by making and testing the $\frac{n(n-1)}{2}$ possible combinations. Because of the high variety x location or variety x year interaction, this did not appear to be a very practical approach. The plan used at present is to select a group of single crosses which are acceptable as single crosses and to use these as testers. New lines are therefore evaluated as three-way combinations. An example of the savings in area and effort can be illustrated by the following example. Twenty lines in all possible combinations equals 190 entries to test. Using 4 single cross testers, crossing each of the 20 lines to each tester makes 80 items to test; a saving of over 50% in planting and harvesting cost, as well as costs of seed production. Three way test crosses permit double cross predictions as well as single crosses, so there is no loss there. The gain in time and cost is, of course, related to number of lines. For 100 lines, all possible single crosses equal 4950 combinations; 100 x 4 = 400 three way test crosses.

Theoretically, through the use of only 4 testers, some decrease in efficiency of estimating general combining ability is involved. However, this seems to be of limited importance, since the new lines will almost certainly be used in some combination with the standard lines involved in the testers. Therefore, beyond a certain point, increased efficiency in estimating general combining ability through using more testers becomes a question of academic interest, primarily.

(3) Additional information on types of gene action involved: This is a somewhat intangible item, but the data accumulated on specific and general combining ability has been of value in clarifying the general problem of testers and has indicated that specific combining ability is made up to a large degree of genotype-environmental interaction. Although it is not apparent how this information would be expressed in terms of cost the findings have been very valuable from a theoretical point of view.

Changes in the Forage Crop Research Program in New York

Drs. C. C. Lowe and R. P. Murphy list the following changes that were made in the forage crops breeding program in New York:

(1) Changes in emphasis on breeding objectives and methods of evaluation: The total season yield of grass-legume mixtures has not been very effective in determining variety differences for grasses. Future tests of this type will be conducted primarily for studying legume persistence with the different grass varieties.

Greater emphasis in the breeding of grasses will be placed upon the improvement of more specific characteristics such as aftermath production, disease resistance, etc. rather than on total yield. The summarization of the data indicated that the evaluation procedure may vary considerably with these specific breeding objectives.

(2) Changes in techniques on the grasses: With greater emphasis on aftermath production in grasses, a change to fewer replicates and more years of harvest is planned.

(3) Changes in techniques on perennial legumes: More replications and fewer years of harvest are planned for future experiments to measure yield when stands are good. Experiments will be continued to determine persistence but yields will not be determined each year.

A Method for Combining Variance Components

Several methods for combining variance component estimates from a group of experiments could be used. The relative efficiency of the various methods has not been ascertained to date. The distribution for variance components is not known. Therefore, we have to rely on empirical evidence in choosing among the various methods of estimation.

One method for combining variance component estimates is to take a weighted or unweighted average of the variance components from the individual experiments (10, 19). A second method for combining variance components is to obtain an unbiased estimate of the ratio of variance components in the individual experiments, and then to average the ratios. The latter procedure is the one used by a number of workers (2, 9, 10, 11, 18, 20, 24) in summarizing variance component estimates. The ratio of variance components is of interest in estimates of heritability, genetic progress, relative loss in information, etc. The method is rather non-restrictive in that no assumption is made about the equality of variances in the individual experiments. If the assumption of equality of variances is tenable then more efficient procedures of estimation are available (9). Due to the very nature of field experiments, and perhaps other experiments, the individual errors are affected by different factors in the different locations and in the different years. Differential erosion, insect infestation, rodent damage, etc. are factors affecting the error variances in the individual experiments. Each individual experiment furnishes an estimated variance component associated with a specified parameter. The parameters may vary from experiment to experiment. If the variation among the parameters is large compared to the variation among estimates of the individual parameter then an unweighted mean may be more efficient in that it has a lower variance than the weighted mean (2, 9). The individual variances in each experiment are assumed to have a chi-square distribution.

The method for combining ratios of variance components from simple randomized complete block designs and from randomized complete block designs at several places is illustrated below. In addition, cost of experimentation is considered for the latter case. It should be remembered that the treatments or varieties may change from experiment to experiment, but that the group of experiments used should be similar in type. Also, small inequalities in the size of the experimental unit may be ignored since the ratio is independent of the units of measure.

The pertinent part of the analysis of variance for a randomized complete block of r replicates conducted at a single place and in a single year is:

Source of variation	df	Average value of mean square
Varieties	$v-1$	$\sigma_e^2 \neq r\sigma_v^2$
Error	f_e	σ_e^2

The unbiased estimate of the ratio of the estimated variety variance component, s_v^2, to the estimated error variance component, s_e^2, from the uth experiment is

$$\hat{a}_u = \frac{f_e - 2}{f_e} \left\{ \frac{s_v^2}{s_e^2} - \frac{2}{r(f_e-2)} \right\} . \qquad (2)$$

The unweighted mean of the ratios from n experiments is

$$\bar{a} = \sum_{u=1}^{n} \hat{a}_u / n, \qquad (3)$$

with an approximated variance equal to

$$\frac{\sum (\hat{a}_u - \bar{a})^2}{n(n-1)} \qquad (4)$$

The variance given in equation (4) is only an approximation since the $\hat{a}_u$ are not subject to the same variances.

The average genetic progress, expressed in standard deviation units, made by selecting the highest yielding variety from the v varieties grown in a randomized complete block experiment of r replicates is estimated to be:

$$G_1 = \frac{\bar{a}}{\sqrt{\bar{a} \neq \frac{1}{r}}} m_v , \qquad (5)$$

where m is the expected value of the largest deviate in a sample of size v from a normally distributed population with a mean of zero and a variance of one (10, 13). Computation of the value for equation (5) for various

values of r and v give an indication of the rate of genetic progress attained under the different schemes. Fify singlecrosses of corn in two replicates of a randomized complete block design gave a higher rate of progress in selecting for high yield than did 25 strains in four replicates. The reverse was true for corn doublecrosses since the value for $\bar{a}$ is much lower for doublecrosses than for singlecrosses. If the highest k varieties were always selected instead of the single highest yielding variety the optimum number of replicates and varieties would be different for the two situations.

For the second illustration of the method for combining estimates of variance components suppose that the individual experiments consisted of v varieties grown in r replicates of a randomized complete block design at each of p randomly selected places (20). The pertinent part of the analysis of variance is:

Source of variation	df	Average value of mean square
Varieties	v-1	$\sigma_e^2 + r\sigma_{vp}^2 + rp\sigma_v^2$
Variety x place	(v-1)(p-1)	$\sigma_e^2 + r\sigma_{vp}^2$
Error	f_e	σ_e^2

The unbiased estimates of the ratio of the estimated variety, s_v^2, and variety x place, s_{vp}^2, variance components to the estimated error variance component, s_e^2, are:

$$\hat{d} = \left(\frac{f_e - 2}{f_e}\right)\left(\frac{s_v^2}{s_e^2}\right) \qquad (6)$$

and

$$\hat{b} = \left(\frac{f_e - 2}{f_e}\right)\left(\frac{s_{vp}^2}{s_e^2} - \frac{2}{r(f_e - 2)}\right) \qquad (7)$$

If n such experiments are available the unweighted averages are:

$$\bar{b} = \sum_{u=1}^{n} \hat{b}_u / n \qquad (8)$$

and

$$\bar{d} = \sum_{u=1}^{n} \hat{d}_u / n . \qquad (9)$$

The average genetic progress made by selecting the highest yielding variety in each experiment is:

$$G_2 = \frac{\bar{dm}_v}{\sqrt{\bar{d} \not{/} \frac{\bar{b}}{p} \not{/} \frac{1}{rp}}} \cdot \qquad (10)$$

For given values of $\bar{b}$ and $\bar{d}$ the various values of v, r, and p result in different values for G_2. For a fixed number of plots, the combination giving the largest value is the desired one.

Since the cost of adding an additional place is often considerably larger than the cost of adding an additional replicate at a given place the rate of genetic progress per dollar spent is desired rather than the value for G_2. The cost relationship suitable for growing v varieties in Iowa is:

$$\text{Cost} = X = \left(\frac{p \not{/} 1}{2}\right) A \not{/} rpB, \qquad (11)$$

where A = all costs associated with one round trip to a place and B = all costs associated with planting and harvesting a single replicate of v varieties. Since the cost of going to an additional place is approximately one-half that of going to a single place the coefficient for A is $(1 \not{/} \frac{p-1}{2})$ $= (p \not{/} 1)/2$. If the cost for travelling to the p places is independent of the number of places then the coefficient for A is p. The particular value of the coefficient for A depends upon the experimental conditions.

Taking the cost equation into account and assuming that v varieties are to be tested the optimum number of replicates and places is given by the following two equations:

$$r = \sqrt{A/2B \ \bar{b}} \qquad (12)$$

and

$$p = \frac{2X - A}{A \not{/} r} \cdot \qquad (13)$$

To illustrate let $\bar{b} = .4$, $X = \$160$, $v = 25$, $A = \$39.38$, and $B = \$8.38$ (20). From equation (12) r is equal to 2.42, and from equation (13) p is equal to 3.51. In practice then one would use either four locations with two replicates per place or three locations with three replicates per place.

The above examples illustrate the procedure for combining ratios of variance components and for making use of the combined estimates for a given cost situation. The procedure for extending the results to other experimental situations is described in the references cited at the end of this paper.

Concluding Remarks

The foregoing was limited to a discussion of types of information that may be derived from a summary of past results and to the discussion of a method for summarizing estimates of variance components. No discussion

is given on the types of information derived from the efficient planning of long-term experiments, such as those discussed by Cochran (3), Cochran and Crowther (7), Dutton (8), Fisher (12), and Yates (22), and from well planned groups of experiments bearing on the same problem. Rapid progress in research depends upon a complete summarization and interpretation of past research as well as on the efficient planning of future experiments. A comprehensive summary and interpretation of past research highlight the gaps in our knowledge about certain areas and emphasize the importance or lack of importance of other areas.

It has been assumed that the summary of experimental data has proceeded further than the collection of the data and the reduction of the data to a few summary figures. If this assumption is not tenable then it is recommended that the results from the individual experiments be studied prior to summarizing the results from a group of experiments.

In closing, I would like to urge all experimenters to preserve the original data from experiments for future use and to make use of these data to accelerate the progress of research. There is still a large amount of golden information to be mined from the original data from past experiments.

Literature Cited

1. Brunk, Max E. and Federer, W. T.
   How marketing problems of the apple industry were attacked and the research results applied
   Methods of Research in Marketing, Paper No. 4, Mimeo. Cornell Univ., Ithaca, N. Y., January 1953

2. Cochran, W. G.
   Problems arising in the analysis of a series of similar experiments
   Jour. Roy. Stat. Soc., Suppl. 4:102-118, 1937

3. Cochran, W. G.
   Long-term agricultural experiments
   Jour. Roy. Stat. Soc., Suppl. 6:104-148, 1939

4. Cochran, W. G.
   A survey of experimental designs
   Mimeo. USDA Lecture Series, March 1940

5. Cochran, W. G.
   An examination of the accuracy of lattice and lattice square experiments on corn
   Iowa Agric. Expt. Sta. Res. Bul. 289:400-415, 1941

6. Cox, G. M.
   A survey of types of experimental designs
   Biometrics 6:301-302, 1950

7. Crowther, F. and Cochran, W. G.
Rotation experiments with cotton in the Sudan Gezira
Jour. Agric. Sci. 32:390-405, 1942

8. Dutton, A. M.
Statistical analysis of long-term agricultural experiments
Ph.D. thesis, Iowa State College Library, Ames, Iowa, 1951

9. Federer, W. T.
Evaluation of variance components from a group of experiments with multiple classifications
Ph.D. thesis, Iowa State College Library, Ames, 1948

10. Federer, W. T.
Evaluation of variance components from a group of experiments with multiple classifications
Iowa Agric. Expt. Sta. Res. Bul. 380:242-310, 1951

11. Federer, W. T. and Sprague, G. F.
A comparison of variance components in corn yield trials: I. Error, tester x line, and line components in top-cross experiments
Jour. Amer. Soc. Agron. 39:453-463, 1947

12. Fisher, R. A.
The influence of rainfall on the yield of wheat at Rothamsted
Phil. Trans. Roy. Soc. 213:89-142, 1924

13. Fisher, R. A. and Yates, F.
Statistical tables for biological, agricultural and medical research
Oliver and Boyd, Ltd., London, 1938

14. Homeyer, P. G. and Black, C. A.
Sampling replicated field experiments on oats for yield determinations
Soil Sci. Proc. 11:341-344, 1946

15. Keller, K. R.
A comparison involving number of, and relationship between, tests in evaluating inbred lines of maize
Agron. Jour. 41:323-331, 1949

16. Kempthorne, O.
A note on differential response in blocks
Jour. Agric. Sci. 37:245-248, 1947

17. King, S. C.
Heritability studies of egg production in the domestic fowl
Ph.D. thesis, Cornell Univ. Library, Ithaca, N.Y., 1953

18. Lowe, C. C.
Evaluation of variance components for forage crop yield trials
Ph.D. thesis, Cornell Univ. Library, Ithaca, N.Y., 1952

19. Rojas, B. and Sprague, G. F.
A comparison of variance components in corn yield trials: III. General and specific combining ability and their interaction with locations and years
Agron. Jour. 44:462-466, 1952

20. Sprague, G. F. and Federer, W. T.
A comparison of variance components in corn yield trials: II. Error, year x variety, location x variety, and variety components
Agron. Jour. 43:535-541, 1951

21. Sprague, G. F. and Tatum, L. A.
General vs. specific combining ability in single crosses of corn
Jour. Amer. Soc. Agron. 34:923-932, 1942

22. Yates, F.
Design of rotation experiments
Commonwealth Bur. Soil Sci., Tech. Comm. No. 46, 1949

23. Yates, F. and Cochran, W. G.
The analysis of groups of experiments
Jour. Agric. Sci. 28:556-580, 1938

24. Yates, F. and Zacopanay, I.
The estimation of the efficiency of sampling, with special reference to sampling for yield in cereal experiments
Jour. Agric. Sci. 25:545-577, 1935