CONTROLLING VARIABILITY[1]

P. Allen Hammer[2]
Department of Horticulture, Purdue University,
West Lafayette, IN 47907

Controlling variability is central to the principles of scientific experimentation. The researcher starts with a *written* statement of the question or questions and the hypotheses. The researcher uses "planned or controlled variability" (treatments) in an experiment to test these hypotheses. However, for valid conclusions, the researcher must also consider "non-planned or unwanted variability" when *designing* the experiment (Fig. 1). The following quote from the 1920s about field experiments graphically makes this point:

> *"As Fisher put it in correspondence, the experimenter games with the devil; he must be prepared by his layout to accommodate whatever pattern of soil fertilities the devil may have chosen in advance." (1)*

The main statistical tools for measuring and/or controlling variability are replication, randomization, and blocking.

Replication

Replication in the statistical literature means multiple experimental units per treatment (the repetition of the set of treatments). Replication along with randomization, which will be discussed later, are both necessary for valid estimates of experimental error. Although replication appears to be a simple concept, it is one requiring careful consideration. Subsamples are sometimes labeled replications (Fig. 2). In nearly all the horticultural research I have observed, the use of subsamples as replicates leads to an underestimate of experimental error, which means that we attach greater significance to treatment differences than we should. Let's look at some examples:

Field. Suppose we were studying the effect on fruit quality of various rates of soil application of potassium to apple trees, and in our study an individual tree was the experimental unit. Multiple apples from a tree would be subsamples rather than replicates, and only apples from different trees would be replicates, because I would argue that the tree is the experimental unit that received the treatment. The apples on a tree would probably be more uniform than apples from different trees; thus we would underestimate the variance (experimental error).

Greenhouse or Growth Chamber. When studying the effect of air temperature on plant growth in a greenhouse or in growth chambers, multiple plants within a single greenhouse or growth chamber would be subsamples and not replicates. The chamber or greenhouse is the experimental unit that should be replicated through the use of more than one greenhouse or chamber at each temperature. This is particularly important because of the significant between-chamber or between-greenhouse variation.

Laboratory. In laboratory analyses, subsamples are often misused as replicates of experimental units. If, for example, one were studying the effect on plant tissue nitrogen content of applying nitrogen to the soil, samples from the same plant (when one plant is the experimental unit) would estimate the variance within the plant and also contain a contribution from laboratory procedure but would not estimate the plant to plant variance which is the true replicate or experimental unit. Multiple analysis of the same tissue sample would provide an estimate of variance of the laboratory procedure but would probably greatly underestimate the plant to plant variance.

Fig. 1. Last year's experiments may be a source of "unwanted variability" for this year's experiments, as shown above. Small grains were strip-planted across the field in the previous year and carrots were planted in rows perpendicular to the strips the following year. Note the effect of allelopathy from a small grain on the carrots. (Photograph by Dr. Alan Putnam, Michigan State University).

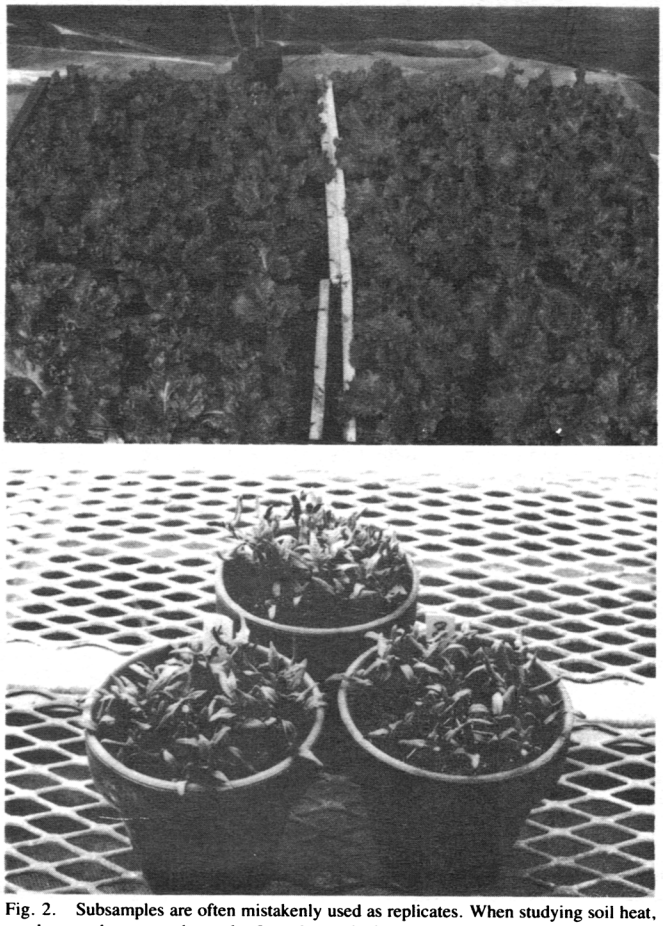

Fig. 2. Subsamples are often mistakenly used as replicates. When studying soil heat, lettuce plants are subsamples from the particular treatment plot but are not replicates on the treatment (top). Multiple seedlings within a pot are subsamples in most experiments (bottom).

[1]Journal Paper number 8547 of the Purdue Agricultural Experiment Station.
[2]Associate Professor.

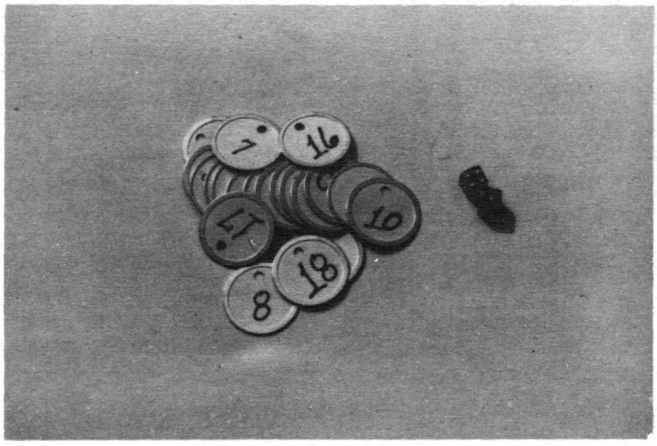

Fig. 3. Randomization should be accomplished with a random device or change mechanism. Remember that some numbers of dots occur with much higher frequency than others when using a pair of dice, thus they should not be used unless you completely understand the frequency.

Randomization

Randomization goes hand-in-hand with replication in making statistical statements about treatment means. Randomization and replication provide a valid estimate of experimental error (variance). If an experimenter would balance pretreatment plant differences (making sure large and small plants were balanced in each treatment group) then the estimate of variance would be inflated and the post-treatment differences will be reduced [see Federer p. 14 (4) for an example]. Nonrandom assignment of treatments to plants can and probably will lead to biased results.

Randomization should be accomplished with some random device or chance mechanism (Fig. 3). Numbered plastic discs or a random number table are the best and most widely used devices. Haphazard or unplanned assignments should not be used as a randomization method because it is not equivalent to randomization and may, in fact, add some very subtle bias to the arrangement.

We have already seen that randomization and replication are important in evaluating biological variation. Sometimes biological variability can be reduced by selecting more uniform plants at the pretreatment stage and then using the tools of randomization and replication for treatment applications. In many cases this will allow the experimenter to detect differences between treatments with fewer replicates. However, one must be careful with such a selection if the conclusions are to be used for statements about the population, because one assumes in such a case that the treatment equally affects the large, small, and medium plants. This is generally a safe assumption, but a researcher needs to be aware of the possibility of again adding a subtle bias to the results.

Blocking

Blocking, on the other hand, can be considered a deliberate or nonrandom assignment of treatments to experimental units to account for *known* sources of variation that are generally uninteresting to the researcher; uninteresting in the sense that they are not planned treatments. Blocks consist of sets of plants and/or microenvironments which are as homogeneous or internally similar as possible. In the field this may be contiguous areas (blocks) of different soil characteristics, while in the greenhouse or growth chamber it may be areas (blocks) of different radiation levels and/or temperature; within the block, conditions are very similar. Each treatment generally appears exactly the same number of times in each block. It is very important that the treatments within each block be randomized independently.

Environmental variability has been long recognized as an important source of "unwanted" variability in many experiments. It is common practice to use blocks in field studies to account for this source of variability. In horticultural research, blocks have generally been square or rectangular. However, blocks can be irregularly shaped if enough is known about the experimental area to justifiably say that the irregularly shaped area contains a more homogeneous environment than a square or rectangular area. Whole blocks also can be noncontiguous if, again, the areas within each block are made more uniform by such a choice.

In much horticultural research, blocking has not been applied to greenhouse and growth chamber studies. The environments were considered homogeneous, and thus did not require special consideration. This was particularly true when greenhouse or growth chamber studies were compared to field studies. This assumption may be unjustified in many cases because the environments, while more uniform, still will show significant variability and thus will justify the use of blocks. The very practical consequence of blocking in this case is a significant increase in precision, i.e., detecting smaller differences among treatments.

Sources of "unwanted" variability

Randomization, replication, and blocking are powerful tools to control variability in horticultural research, but like any tool, they should be used with care and knowledge. An experiment will not necessarily correctly test the hypotheses if these tools are not used with a clear understanding of 3 major sources of "unwanted" variability: biological, environmental, and experimenter-induced.

Experimenter-induced variability should be the *greatest* concern, because it is the source of "unwanted" variability the researcher has greatest control over. A discussion of this source is best approached with several examples.

The method of application of a treatment can influence the results of and conclusions for an experiment. For example, when a carrier gas is used to apply a chemical, the experimenter must be concerned not only with the effect of the chemical but also with the carrier gas. A water control with the carrier gas will not answer the question of chemical and gas interaction. Thus, in this example and in other studies, the method of application of a treatment makes it impossible to have a "true" control. The point is, understand the treatment structure completely before applying labels.

Handling or movement of plants should be accomplished uniformly over all plants and treatments. Daily plant measurements can be a very troublesome source of variability, because touching and moving can induce vibration stress in the plants (Fig. 4). If blocks are used in an experiment, all experimental operations should be by blocks (from applying treatments, to taking measurements, to harvesting). This will allow the researcher to account for this variability as a block effect thus not confounding it with the treatment effect. In fact, separate blocks could be harvested on different days and be accounted for as a block effect.

Sometimes variously colored containers are used to identify treatments, particularly in greenhouse fertilizer studies (Fig. 5). Container color can influence medium temperature, and thus root growth, which would become confounded with treatment effect.

Data recording is often easier when done systematically, but this should be an area of great concern for randomization. Systematic data collection can add very large systematic errors. The best example is

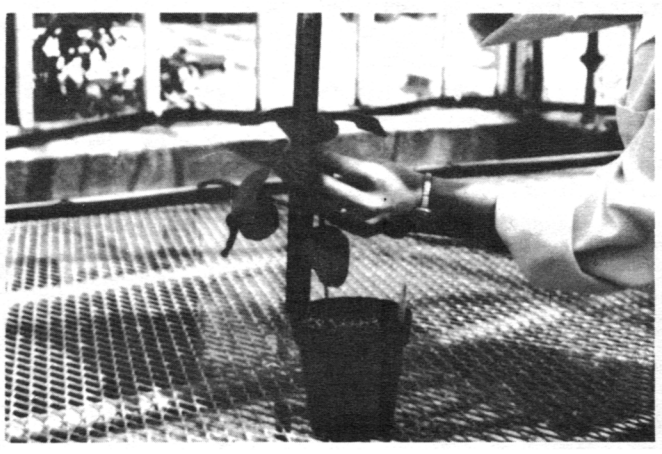

Fig. 4. Daily plant measurements can be a very troublesome source of variability from touching and movement of the plants.

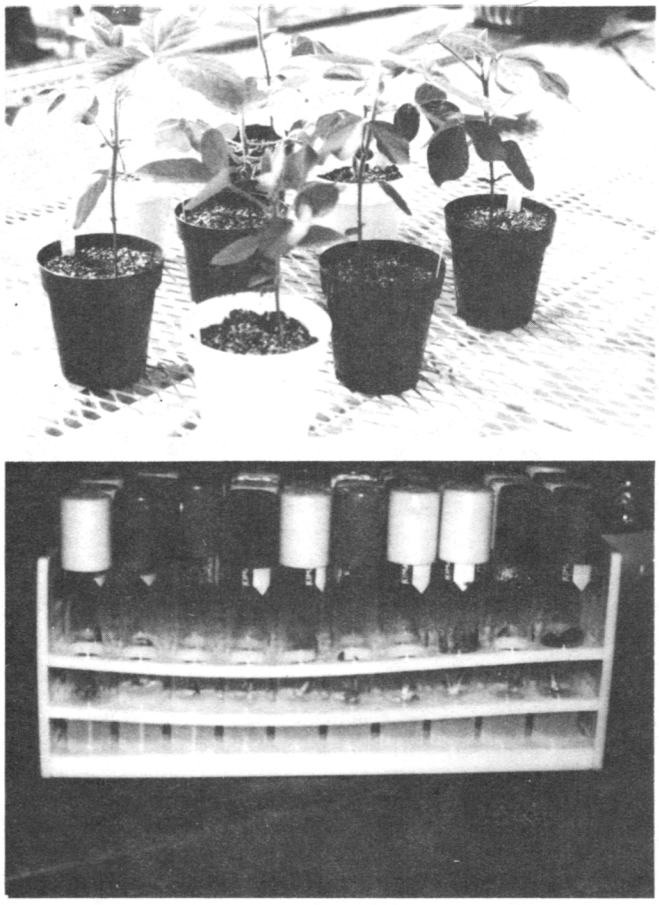

Fig. 5. Differently colored containers are often used to identify treatments. Color can influence temperature and radiation received by roots (top) or plants in tissue culture tubes (bottom).

Fig. 6. Systematic data recording can add large errors. This is a very serious problem when recording fresh or dry weight.

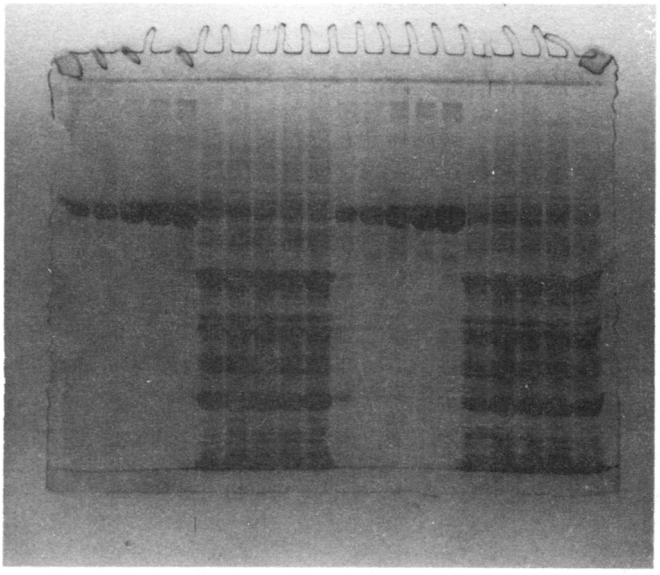

Fig. 7. Every track in gel electrophoresis is not necessarily uniform, as is apparent in the above photographs. Randomization is important in all experiments. (Photograph by David Eichholtz, Department of Horticulture, Purdue University).

recording dry weights. Plant materials removed from an oven should be cooled to room temperature in a desiccator before weighing. How often we have seen dry samples lined-up on a bench by treatment and replicate for weighing (Fig. 6). In such a case, the last samples measured will be heavier because of increased water absorption from the atmosphere when done in a room without temperature and humidity control.

Sometimes, we add systematic errors because a method may not only be easier to use but because we consider our measurement equipment as an absolute. For example, in gel electrophoresis it is easier to put the same sample in the same track for every run; this makes it much simpler to see the results. One may argue that all the tracks are uniform; however, how many have actually tested that assumption over several runs (Fig. 7)? Thus, back to the basics — randomize everything. It is better to assume "unwanted" variability as a problem and randomized every place possible, then to report data that can be suspected of systematic errors.

Another area of concern in controlling variability is repeatability of a study over a time or in a different laboratory. This involves a very careful description of experimental procedures and conditions. We have been extremely concerned about repeatability of growth chamber studies; however, repeatability is essential to all research. No matter how good your particular results are, they are of no value if they can not be repeated.

Again, let us use an example. Radiation can be measured with several different instruments, all sensing the same thing differently. Where do you locate the sensor? How do you know that it is working? The point is: 1) understand what you are measuring; 2) calibrate the instruments; and 3) carefully report measurements and how they were obtained.

I have not covered every possible source of variability in horticultural research, nor have I suggested every possible solution to controlling "unwanted variability." However, I hope this presentation has raised your level of consciousness to, or awareness of, the idea of controlling "unwanted variability." The improved experimental results could be well worth the extra effort.

Selected References

1. Box, J. F. 1980. R. A. Fisher and the design of experiments, 1922–1926. Amer. Stat. 34:1–7.
2. Cochran, W. G. and G. M. Cox. 1957. Experimental designs. Wiley, New York.
3. Davies, O. L. (ed.). 1954. The design and analysis of industrial experiments. Oliver & Boyd, London.
4. Federer, W. T. 1955. Experimental design. Macmillan, New York.
5. Fisher, R. A. 1960. The design of experiments, 7th ed. Oliver & Boyd, London.
6. Hammer, P. A. and R. W. Langhans. 1972. Experimental design consideration for growth chamber studies. HortScience 7:481–483.
7. Hammer, P. A. and N. S. Urquhart. 1979. Precision and replication: Critique II. p. 343–368. In: T. W. Tibbitts and T. T. Kozlowski (eds.) Controlled environment guidelines for plant research. Academic Press, New York.
8. Hruschka, H. W. and E. J. Koch. A reason for randomization within controlled environment chambers. J. Amer. Soc. Hort. Sci. 85:677–684.
9. Kempthorne, O. 1952. The Design and analysis of experiments. Wiley, New York.
10. Mitchell, C. A., C. J. Severson, J. A. Wott, and P. A. Hammer. 1975. Seismomorphogenic regulation of plant growth. J. Amer. Soc. Hort. Sci. 100:161–165.
11. Pearce, S. C. 1976. Field experimentation with fruit trees and other perennial plants. Techn. Comm. 23, 2nd ed. (Rev.). Commonwealth Bureau of Horticulture & Plantation Crops, East Malling, England.