

Experimentation

An *experiment* deliberately imposes a *treatment* on a group of objects or subjects in the interest of observing the response. This differs from an *observational study*, which involves collecting and analyzing data without changing existing conditions. Because the validity of an experiment is directly affected by its construction and execution, attention to *experimental design* is extremely important.

Treatment

In experiments, a *treatment* is something that researchers administer to experimental units. For example, a corn field is divided into four, each part is 'treated' with a different fertiliser to see which produces the most corn; a teacher practices different teaching methods on different groups in her class to see which yields the best results; a doctor treats a patient with a skin condition with different creams to see which is most effective. Treatments are administered to experimental units by 'level', where level implies amount or magnitude. For example, if the experimental units were given 5mg, 10mg, 15mg of a medication, those amounts would be three levels of the treatment.

(Definition taken from Valerie J. Easton and John H. McColl's [Statistics Glossary v1.1](#))

Factor

A *factor* of an experiment is a controlled independent variable; a variable whose levels are set by the experimenter.

A factor is a general type or category of treatments. Different treatments constitute different levels of a factor. For example, three different groups of runners are subjected to different training methods. The runners are the experimental units, the training methods, the treatments, where the three types of training methods constitute three levels of the factor 'type of training'.

(Definition taken from Valerie J. Easton and John H. McColl's [Statistics Glossary v1.1](#))

Experimental Design

We are concerned with the analysis of data generated from an experiment. It is wise to take time and effort to organize the experiment properly to ensure that the right type of data, and enough of it, is available to answer the questions of interest as clearly and efficiently as possible. This process is called *experimental design*.

The specific questions that the experiment is intended to answer must be clearly identified before carrying out the experiment. We should also attempt to identify known or expected sources of variability in the experimental units since one of the main aims of a designed experiment is to reduce the effect of these sources of variability on the answers to questions of interest. That is, we design the experiment in order to improve the precision of our answers.

(Definition taken from Valerie J. Easton and John H. McColl's [Statistics Glossary v1.1](#))

Control

Suppose a farmer wishes to evaluate a new fertilizer. She uses the new fertilizer on one field of crops (A), while using her current fertilizer on another field of crops (B). The irrigation system on field A has recently been repaired and provides adequate water to all of the crops, while the system on field B will not be repaired until next season. She concludes that the new fertilizer is far superior.

The problem with this experiment is that the farmer has neglected to **control** for the effect of the differences in irrigation. This leads to **experimental bias**, the favoring of certain outcomes over others. To avoid this bias, the farmer should have tested the new fertilizer in identical conditions to the control group, which did not receive the treatment. Without controlling for outside variables, the farmer cannot conclude that it was the effect of the fertilizer, and not the irrigation system, that produced a better yield of crops.

Another type of bias that is most apparent in medical experiments is the **placebo effect**. Since many patients are confident that a treatment will positively affect them, they react to a control treatment which actually has no physical affect at all, such as a sugar pill. For this reason, it is important to include control, or placebo, groups in medical experiments to evaluate the difference between the placebo effect and the actual effect of the treatment.

The simple existence of placebo groups is sometimes not sufficient for avoiding bias in experiments. If members of the placebo group have any knowledge (or suspicion) that they are not being given an actual treatment, then the effect of the treatment cannot be accurately assessed. For this reason, **double-blind** experiments are generally preferable. In this case, neither the experimenters nor the subjects are aware of the subjects' group status. This eliminates the possibility that the experimenters will treat the placebo group differently from the treatment group, further reducing experimental bias.

Randomization

Because it is generally extremely difficult for experimenters to eliminate bias using only their expert judgment, the use of **randomization** in experiments is common practice. In a randomized experimental design, objects or individuals are randomly assigned (by chance) to an experimental group. Using randomization is the most reliable method of creating homogeneous treatment groups, without involving any potential biases or judgments. There are several variations of randomized experimental designs, two of which are briefly discussed below.

Completely Randomized Design

In a **completely randomized design**, objects or subjects are assigned to groups completely at random. One standard method for assigning subjects to treatment groups is to label each subject, then use a table of random numbers to select from the labelled subjects. This may also be accomplished using a computer. In MINITAB, the "SAMPLE" command will select a [random sample](#) of a specified size from a list of objects or numbers.

Randomized Block Design

If an experimenter is aware of specific differences among groups of subjects or objects within an experimental group, he or she may prefer a **randomized block design** to a completely randomized design. In a block design, experimental subjects are first divided into homogeneous blocks before they are randomly assigned to a treatment group. If, for instance, an experimenter had reason to believe that age might be a

significant factor in the effect of a given medication, he might choose to first divide the experimental subjects into age groups, such as under 30 years old, 30-60 years old, and over 60 years old. Then, within each age level, individuals would be assigned to treatment groups using a completely randomized design. In a block design, both *control* and *randomization* are considered.

Example

A researcher is carrying out a study of the effectiveness of four different skin creams for the treatment of a certain skin disease. He has eighty subjects and plans to divide them into 4 treatment groups of twenty subjects each. Using a randomized block design, the subjects are assessed and put in blocks of four according to how severe their skin condition is; the four most severe cases are the first block, the next four most severe cases are the second block, and so on to the twentieth block. The four members of each block are then randomly assigned, one to each of the four treatment groups.

(Example taken from Valerie J. Easton and John H. McColl's [Statistics Glossary v1.1](#))

Replication

Although randomization helps to insure that treatment groups are as similar as possible, the results of a single experiment, applied to a small number of objects or subjects, should not be accepted without question. Randomly selecting two individuals from a group of four and applying a treatment with "great success" generally will not impress the public or convince anyone of the effectiveness of the treatment. To improve the significance of an experimental result, *replication*, the repetition of an experiment on a large group of subjects, is required. If a treatment is truly effective, the long-term averaging effect of replication will reflect its experimental worth. If it is not effective, then the few members of the experimental population who may have reacted to the treatment will be negated by the large numbers of subjects who were unaffected by it. Replication reduces variability in experimental results, increasing their significance and the confidence level with which a researcher can draw conclusions about an experimental factor.