

16.5: Experimental Designs

Learning Objectives

- Distinguish between between-subject and within-subject designs
- State the advantages of within-subject designs
- Define "multi-factor design" and "factorial design"
- Identify the levels of a variable in an experimental design
- Describe when counterbalancing is used

There are many ways an experiment can be designed. For example, subjects can all be tested under each of the treatment conditions or a different group of subjects can be used for each treatment. An experiment might have just one independent variable or it might have several. This section describes basic experimental designs and their advantages and disadvantages.

Between-Subjects Designs

In a between-subjects design, the various experimental treatments are given to different groups of subjects. For example, in the "Teacher Ratings" case study, subjects were randomly divided into two groups. Subjects were all told they were going to see a video of an instructor's lecture after which they would rate the quality of the lecture. The groups differed in that the subjects in one group were told that prior teaching evaluations indicated that the instructor was charismatic whereas subjects in the other group were told that the evaluations indicated the instructor was punitive. In this experiment, the independent variable is "Condition" and has two levels (charismatic teacher and punitive teacher). It is a between-subjects variable because different subjects were used for the two levels of the independent variable: subjects were in either the "charismatic teacher" or the "punitive teacher" condition. Thus the comparison of the charismatic-teacher condition with the punitive-teacher condition is a comparison between the subjects in one condition with the subjects in the other condition.

The two conditions were treated exactly the same except for the instructions they received. Therefore, it would appear that any difference between conditions should be attributed to the treatments themselves. However, this ignores the possibility of chance differences between the groups. That is, by chance, the raters in one condition might have, on average, been more lenient than the raters in the other condition. Randomly assigning subjects to treatments ensures that all differences between conditions are chance differences; it does not ensure there will be no differences. The key question, then, is how to distinguish real differences from chance differences. The field of inferential statistics answers just this question. The inferential statistics applicable to testing the difference between the means of the two conditions can be found here. Analyzing the data from this experiment reveals that the ratings in the charismatic-teacher condition were higher than those in the punitive-teacher condition. Using inferential statistics, it can be calculated that the probability of finding a difference as large or larger than the one obtained if the treatment had no effect is only 0.018. Therefore it seems likely that the treatment had an effect and it is not the case that all differences were chance differences.

Independent variables often have several levels. For example, in the "Smiles and Leniency" case study the independent variable is "type of smile" and there are four levels of this independent variable:

1. false smile
2. felt smile
3. miserable smile
4. a neutral control

Keep in mind that although there are four levels, there is only one independent variable. Designs with more than one independent variable are considered next.

Multi-Factor Between-Subject Designs

In the "Bias Against Associates of the Obese" experiment, the qualifications of potential job applicants were judged. Each applicant was accompanied by an associate. The experiment had two independent variables: the weight of the associate (obese or average) and the applicant's relationship to the associate (girl friend or acquaintance). This design can be described as an Associate's Weight (2) x Associate's Relationship (2) factorial design. The numbers in parentheses represent the number of levels of the independent variable. The design was a factorial design because all four combinations of associate's weight and associate's relationship were included. The dependent variable was a rating of the applicant's qualifications (on a 9-point scale).

If two separate experiments had been conducted, one to test the effect of Associate's Weight and one to test the effect of Associate's Relationship then there would be no way to assess whether the effect of Associate's Weight depended on the Associate's Relationship. One might imagine that the Associate's Weight would have a larger effect if the associate were a girl friend rather than merely an acquaintance. A factorial design allows this question to be addressed. When the effect of one variable does differ depending on the level of the other variable then it is said that there is an interaction between the variables.

Factorial designs can have three or more independent variables. In order to be a between-subjects design there must be a separate group of subjects for each combination of the levels of the independent variables.

Within-Subjects Designs

A within-subjects design differs from a between-subjects design in that the same subjects perform at all levels of the independent variable. For example consider the "ADHD Treatment" case study. In this experiment, subjects diagnosed as having attention deficit disorder were each tested on a delay of gratification task after receiving methylphenidate (MPH). All subjects were tested four times, once after receiving one of the four doses. Since each subject was tested under *each* of the four levels of the independent variable "dose," the design is a within-subjects design and dose is a within-subjects variable. Within-subjects designs are sometimes called repeated-measures designs.

Counterbalancing

In a within-subject design it is important not to confound the order in which a task is performed with the experimental treatment. For example, consider the problem that would have occurred if, in the ADHD study, every subject had received the doses in the same order starting with the lowest and continuing to the highest. It is not unlikely that experience with the delay of gratification task would have an effect. If practice on this task leads to better performance, then it would appear that higher doses caused the better performance when, in fact, it was the practice that caused the better performance.

One way to address this problem is to counterbalance the order of presentations. In other words, subjects would be given the doses in different orders in such a way that each dose was given in each sequential position an equal number of times. An example of counterbalancing is shown in Table 16.5.1.

Table 16.5.1: Counterbalanced order for four subjects

Subject	0 mg/kg	0.15 mg/kg	0.30 mg/kg	0.60 mg/kg
1	First	Second	Third	Fourth
2	Second	Third	Fourth	First
3	Third	Fourth	First	Second
4	Fourth	First	Second	Third

It should be kept in mind that counterbalancing is not a satisfactory solution if there are complex dependencies between which treatment precedes which and the dependent variable. In these cases, it is usually better to use a between-subjects design than a within-subjects design.

Advantage of Within-Subjects Designs

An advantage of within-subjects designs is that individual differences in subjects' overall levels of performance are controlled. This is important because subjects invariably will differ greatly from one another. In an experiment on problem solving, some subjects will be better than others regardless of the condition they are in. Similarly, in a study of blood pressure some subjects will have higher blood pressure than others regardless of the condition. Within-subjects designs control these individual differences by comparing the scores of a subject in one condition to the scores of the same subject in other conditions. In this sense each subject serves as his or her own control. This typically gives within-subjects designs considerably more power than between-subjects designs. That is, this makes within-subjects designs more able to detect an effect of the independent variable than are between-subjects designs.

Within-subjects designs are often called "repeated-measures" designs since repeated measurements are taken for each subject. Similarly, a within-subject variable can be called a repeated-measures factor.

Complex Designs

Designs can contain combinations of between-subject and within-subject variables. For example, the "Weapons and Aggression" case study has one between-subject variable (gender) and two within-subject variables (the type of priming word and the type of word to be responded to).

This page titled [16.5: Experimental Designs](#) is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by [David Lane](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- **6.5: Experimental Designs** by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.