

## 4.1: Overview of the Control Variable

---

In the first chapter, we talked about many different threats to internal validity of a research design and one of the control techniques is to build the extraneous variable into our research design. In this chapter, we will extend between-subjects design by looking at different ways to add in an extraneous or a control variable. Why do we need to add control variables? And what criteria should we use when selecting control variables? The main reason we want to include control variables is that the control variables are having an effect on the dependent variable we are studying. Since control variables are not the independent variables in our research, they could potentially confound the results of the study if left unattended. In other words, they can impose threats to the internal validity of the research design. By taking some measures to include the control variables, we are minimizing their effect on the dependent variable, which gives us more confidence to claim it is the independent variable, not the control variable, that causes changes in the dependent variable.

Using the example from the previous chapter, let's say we are conducting an experiment on the effect of cell phone use (yes vs. no) on driving ability. The independent variable is cell phone use with two treatment conditions (yes or no) and the dependent variable is driving ability. A potential control variable would be driving experience as driving experience is most likely to have an impact on driving ability. In order to reduce the potential threat driving experience has on driving ability, we can add it into our study as a control variable. Although it is not the focus of the study, control variable IS a part of your study as we know it influences the outcome variable. By including driving experience into our study, we can minimizing its effect on our research design, and be more confident it is the cell phone use, not driving experience, that leads to changes in driving ability. Therefore adding control variables can increase the internal validity of the research design.

How do we select control variables? Any variables can be potential control variables as long as there is good theoretical or empirical evidence(s) to show they influence the outcome variables. The nature of the variable is not a concern. The control variable can be categorical or continuous. Using the same example above, to measure driving experience, we can ask participants to identify which following level of driving experience represents them the best - seasoned, intermediate, or inexperienced. Or we can ask participants to identify how many months they have driven. Or if you are concerned about the accuracy of participants' own estimation, you can ask participants the age they received their driving license and do the calculation yourself. Regardless how you measure it, as long as the control variable is solid, as in it indeed influences the outcome variable, it can be included in the research study.

Then how do we use different types of control variables? There are two major ways to use control variables. One is randomized block design, which uses control variables at the design stage when we actively set up the experiment. Randomized block design typically uses categorical control variables. The other one is analysis of covariance, which uses control variables at the data analysis stage when we analyze the statistical data. Analysis of covariance typically uses continuous variables. We will look at each of them closely in the following sections.

---

This page titled [4.1: Overview of the Control Variable](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Yang Lydia Yang](#).