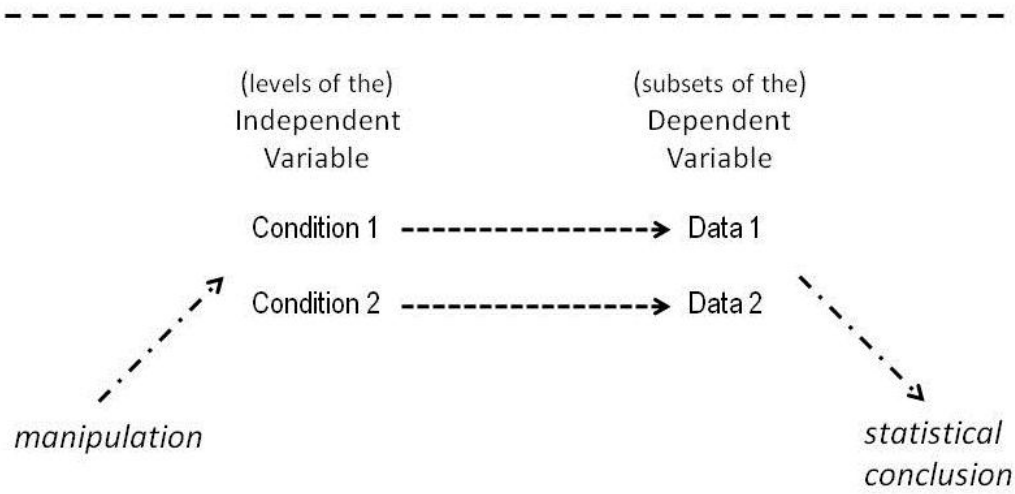
Pre-lecture Notes II.3 – Introduction to Inferential Statistics

Here, again, is a picture that summarizes the core components of an experiment:



Recall from II.1 that internal validity – which depends on the absence of confounds – is mostly related to the dot-dashed line on the left side (i.e., the arrow between *manipulation* and levels of the independent variable). If the manipulation creates one and only one difference between the conditions, then you are on your way to a high level of internal validity (although, as you'll see later, you're not all the way there yet). If your manipulation (by mistake) actually creates more than one difference between the conditions, then your internal validity will be relatively low.

Now note that there's another dot-dashed line (between subsets of the dependent variable and *statistical conclusion*). This arrow is related to the other (new) kind of validity for this part of the course:

> Statistical Conclusion Validity – *the extent to which inferences about the sampling population, based on a sample, are accurate*

This second kind of validity is important because we usually are not trying to come to a conclusion about only the subjects in our experiment, but would like to make a conclusion about everyone. For example, we probably aren't running the lighting/memory experiment because we want to know if this particular (and relatively small) group of people have memory systems that are affected by lighting; we're probably running this experiment because we want to know if people *in general* have memory systems that are affected by lighting. Therefore, we are going to want to go beyond our actual data and make some statement that includes people that we didn't actually measure. The extent to which we are accurate when we do this is what we mean by statistical conclusion validity.

Paralleling internal validity, there is one, major, general threat to statistical conclusion validity plus several other specific threats that depend on the situation or the outcome from the analysis. (As a reminder: the one, major, threat to internal validity is having one or more confounds.) The general threat to statistical conclusion validity is violating one or more assumptions of the statistical procedure being used. These assumptions are necessary because we are making an inference – that is: we are going beyond the data in hand to make some statement about the entire population from which we took our

sample. Inferences always involve assumptions which can, I'm sorry to say, turn out to be false. When we use a statistical procedure that relies on an assumption that happens to be false for our data, the inference that we make could easily turn out to be wrong. Therefore, violating one or more of the assumption that are required by our statistical procedure(s) is the major, general threat to statistical conclusion validity.

We'll get into the details of all this later, but, in case you're wondering (and can't wait), an example of an assumption that is required for a statistical procedure is that the data (across subjects, within a condition) are normally distributed. (This is why we include shape as one of the three descriptive stats; our statistical procedures care about shape, even when we don't.) If the shape of the distribution of scores (across subjects, within a condition) is not a nice, bell-shaped, normal distribution, then the conclusions that we might try to make based on the sample could be very wrong.