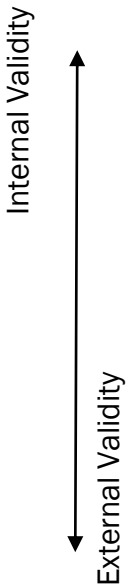
Experimental and Quasi-Experimental designs

True experimental designs are characterized by three "criteria for causality." These are: 1) The cause (independent variable) must precede the effect (dependent variable); 2) there must be a theoretical relationship (aka your conceptual map and its relationships) between the cause and the effect; and, 3) the casual relationship cannot be accounted for by some other variable (aka the environment and research situation must be controlled). Clearly, all three criteria are nearly impossible to guarantee in studies using human beings (they work much better with rats). So, the closest we have to this level of control is through the well-designed clinical trial that uses random assignment.

The most common **type of experimental design** employed in clinical research is the **randomized clinical trial** *(*RCT). Clinical trials are experiments with human beings. RCTs are true experiments with human beings designed to answer specific questions about the effects of biomedical or behavioral interventions. RCTs are categorized by phases, generally defined as follows:



Phase 1 Clinical Trial: clinical studies performed to evaluate the safety of diagnostic, therapeutic, prophylactic drugs, devices, or techniques in a small number of volunteer subjects

Phase II Clinical Trial: preliminary tests of treatment efficacy, side effects, and dose effects. May not necessarily employ a comparison group.

Phase III Clinical Trial: new treatments are experimentally compared with standard treatments, no treatment, or a placebo. Usually large randomized blinded clinical intervention studies to evaluate the efficacy of a biomedical or behavioral intervention in a large group of human subjects (typically several hundred to several thousand to ensure sufficient power to detect effects).

Phase IV Clinical Trial: large scale effectiveness trials of a new treatment, typically the final step in the evaluation of a new therapy that is intended to be marketed widely (e.g., vaccines, surgeries, drugs).

Note: The National Institutes of Health (NIH) require that all clinical trials must now have a plan for a Data Safety and Monitoring Board (DSMB) to oversee the safety of the intervention. However, only Phase III clinical trials must employ an independent DSMB (i.e., the board cannot include members of the research team.

SOME IMPORTANT TERMS USED IN DISCUSSION OF CLINICAL TRIALS:

1. ADHERENCE: degree to which subjects comply with treatment assignment or protocols.

2. BLINDING: subjects are unaware of which treatment they are receiving. Sometimes called *masking*.

3. CROSS-OVER TRIAL: trials in which subjects receive all the study treatments during different periods separated by a *washout* period. This permits estimating the **within-patient treatment differences.**

4. DROP-IN: in a comparative trial, a study subject who takes another treatment in the trial instead of the one to which he/she was assigned and remains available for follow up.

5. DROP-OUT: study subjects who stop taking the treatment to which they were assigned. In some clinical trial, these subjects are not removed from the study if they remain available for follow-up (see intention-to-treat). If they do not remain in the trial, they are then referred to as "lost to follow-up".

6. ENDPOINT: the indentified outcome for a clinical trial (dependent variable).

7. INTENTION-TO –TREAT: the idea that subjects assigned to treatments in a RCT should be analyzed according to the assigned treatment group rather than according to the treatment actually received. ("once randomized, analyze")

8. PROTOCOL: the written plan for conducting an experiment, clinical study, or clinical trial.

The **random assignment** of the participants in an experiment can be delineated by the type of design. For example, a **randomized block design** uses the two-group, pretest-posttest or the two group posttest only design with using a blocking variable to rank order participants. **Factorial designs** examine multiple causality. In the basic factorial design, two treatments and factors are involved and within each factor, two levels are manipulated. The designs can become more complicated with multiple factors, treatments and levels within the factors i.e. a 2 x 2 or 3 x 3 design. These designs are used when you want to test two types of intervention or the intensity of an intervention at different levels. **Nested designs** include independent variables like gender and education that defines the participants or can be geographical locations like different counties, clinical patient units. These designs examine interventions within the context of these independent variables. **Crossover/counterbalanced designs** involve experiments where the participants receive both interventions, but they are delivered sequentially rather than at the same time to two groups after a "washout" or resting period. This minimizes the any effects from fatigue, practice frequency, etc. because by using the same subject, any effects will occur for each intervention.

Quasi-experimental designs are experiments that do not meet the strict criteria for a true experiment either because they lack control over the sample or over extraneous variables. True experiments are ideal because the level of control increases confidence in the validity of the results. However, quasi-experiments are more feasible and can lead to results that are more generalizable since real life cannot mimic the controlled environment of a true experiment (see required readings by Fogg and Gross for more in-depth discussion). The researcher attempts to control as many threats to validity as possible when one of the three "causes of causality" cannot be met; the greatest amount of control possible is planned within the design to decrease the variance. There are three types of quasi-experimental designs: Comparison groups, Nonequivalent comparison groups, and Interrupted Time-Series Designs.

Below describe in your own words, each of the types from the Grove, Burns, & Gray (2013)

Quasi-experimental Study Designs

NONEQUIVALENT COMPARISON GROUP DESIGNS

One-group posttest-only design

Posttest-only design with comparison group

One-group pretest-posttest design

Pretest and posttest design with a comparison group

Pretest and posttest design with two comparison treatments

Pretest and posttest design with two comparison treatments and a standard or routine care group

Pretest and posttest design with a removed treatment

Pretest and posttest design with a reversed treatment

INTERRUPTED TIME-SERIES DESIGNS

Simple interrupted time-series designs

Interrupted time-series design with a no-treatment comparison group

Interrupted time-series design with multiple treatment replications

Experimental Study Designs

Classic experimental design

Experimental posttest-only comparison group design

Randomized block design

Factorial design

Nested design

Crossover or counterbalanced design

Clinical trials

Randomized controlled trials

Internal validity concerns the extent to which the design and conduct of the study eliminate the possibility of error or bias so that the effects detected in a study are a true reflection of reality and not the result of error caused by extraneous variables we have not measured. In a true experiment, internal validity reflects the likelihood that the relationship between two variables is a **causal** one; that is, whether observed changes in the dependent variable (outcome) can be attributed to the intervention (i.e., the cause or independent variable) and ***not*** to other possible causes (sometimes described as "alternative explanations" for the outcome).

THREATS TO INTERNAL VALIDITY

Did the planned independent variables really cause the dependent variable??

1. HISTORY: a common event or experience outside the experiment alters the outcome

2. MATURATION: changes that occur because the subjects are getting older or more experienced which then alters the outcome.

3. TESTING: the act of measuring the dependent variable alters the outcome.

4. INSTRUMENTATION: a measuring device changes either in its calibration or meaning between the pretest and the posttest [we will explore this much more in unit 2].

    a. INSTRUMENT DECAY: instruments become systematically less accurate over time
        - FLOOR EFFECTS: time 1 baseline levels are so low, no effects can be detected at Time 2. Pertains to research studies in which the dependent variable is hypothesized to decrease.
        - CEILING EFFECTS: time 1 baseline levels are so high, no effects can be detected at Time 2. Pertains to research studies in which the dependent variable is hypothesized to increase.

5. STATISTICAL REGRESSION TOWARD THE MEAN: the mathematical tendency for an extreme measurement on any variable to be followed by a second measurement which falls closer to the mean

6. SELECTION: pre-existing differences in subjects affect the experimental outcome

7. EXPERIMENTAL MORTALITY: differential attrition from a sample which biases the outcome

8. SELECTION-MATURATION INTERACTION: when respondents in one group grow more experienced, more frustrated, etc. than the respondents in another group.

9. SELECTION-HISTORY INTERACTION: when respondents in one group are exposed to an event (also called “local history”) that the respondents in another group are not exposed to.

10. SELECTION-INSTRUMENTATION INTERACTION: when different groups score at different mean positions on a test whose intervals are not equal (e.g., differential “ceiling” or “floor” effects).

11. RESENTFUL DEMORALIZATION OF RESPONDENTS RECEIVING LESS DESIRABLE TREATMENTS: when the outcomes of the no-treatment control group can be attributed to feeling deprived, demoralized, or resentful because they are not receiving a treatment which is perceived to be beneficial.

12. DIFFUSION OR IMITATION OF TREATMENTS: when treatments involve informational programs and when the experimental and control groups can communicate with each other, respondents in one group may learn the information intended for the other group (i.e., contamination effects).

13. COMPENSATORY EQUALIZATION OF TREATMENTS: when the experimental treatment provides goods or services generally believed to be desirable, there may emerge administrative reluctance to tolerate the focused inequality and cause them to overcompensate for the inequality.

14. COMPENSATROY RIVALRY BY RESPONDENTS RECEIVING LESS DESIRABLE TREATMENTS: when the control group is aware that the experimental group is receiving something positive, the control group becomes motivated to compensate for not receiving the treatment.

External validity is concerned with the extent to which study findings can be generalized beyond the sample used in the study. With the most serious threat, the findings would be meaningful only for the group being studied.

THREATS TO EXTERNAL VALIDITY
To what populations and settings can I generalize my findings??

1. INTERACTION OF SELECTION AND TREATMENT: characteristics of the subjects who received the treatments are not representative of the population.

2. INTERACTION OF SETTING AND TREATMENT: characteristics of the study location (e.g., lab) are not representative of the setting in which the study results would be generalized.

3. INTERACTION OF HISTORY AND TREATMENT: environmental events affect treatment efficacy in ways that would not be typical of the environments to which the findings would be generalized.

The relationship between **internal** and **external validity:** the greater the internal validity of a study, the poorer the external validity