Appendix B

Fundamentals of Experimental Design

The following information was developed by Steven Rothke, PhD, Department of Psychology, Rehabilitation Institute of Chicago and expanded by Mary F. Schmidt, PhD, Midwest Regional Head Injury Center for Rehabilitation and Prevention, RIC. Used with permission of the authors. Edits were made by Jonathan VanGeest, PhD, Institute for Ethics, American Medical Association.

I. BASIC CONCEPTS

Validity: refers to the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration.

Reliability: a matter of whether a particular technique applied repeatedly to the same object, would yield the same result each time.

Variable: A quantitative or qualitative entity that can take on different values. In research, variables are manipulated or measured and represent the concepts being studied.

There are at least two types of variables in an experiment.

1) **Independent Variable (IV):** Manipulated by the experimenter to produce changes in behavior to bring about some effect. Theoretically, we seek to show that the independent variable and its manipulation is responsible for the observed effect or change.

2) **Dependent Variable (DV):** The behavior or effect which is measured and presumed to change as a consequence of the experimenter's manipulation of the independent variable.

3) **Control Variable:** (Not required in an experiment) An additional independent variable which is not of direct interest to the hypothesis, but is included as a way of factoring out its effect on the dependent variable so that it is not mistakenly attributed to the main independent variable.

Example: We wish to experimentally test the effect of different dosages of baclofen (independent variable) on severity and reduction of spasticity (measured by EMG, the dependent variable). If we believe that the type of injury or level of injury (para vs. quad) will make a difference in terms of receptivity to baclofen, we can include this information as a control variable.

Another form of **experimental control** is a **control group** (or comparison group) which permits control for many possible contaminations of results (threats to validity). In the above example, a control group which gets no medication or a different medication could be included in order to be more confident that the presence of baclofen led to the reduction of spasticity.

II. THREATS TO INTERNAL VALIDITY

(i.e. sources of experimental error)

This is a list of factors which can confound the results of an experiment. They should be taken into account when designing and conducting any experiment.

1) **History:** Specific events occurring during the time of the study in addition to the manipulation of the independent variable.

2) **Maturation:** Processes within the subjects operating as a function of the passage of time which are not specific to the events of manipulations of the independent variable.

3) **Testing:** The effect of taking a test upon the scores at a second testing (practice effect, as well as making subjects aware of what you are interested in measuring).

4) **Instrumentation:** Changes in the calibration of measurement or changes in observers or scores over time (unreliability of measurement).

5) **Statistical regression:** (Regression toward the mean) Operating where groups have been obtained on the basis of extreme scores.

6) **Selection bias** of subjects.

7) **Experimental mortality**: Differential loss of subjects from groups.

8) **Selection-Maturation Interaction:** Where groups differ at the outset, such as naturally occurring groups.

The threats to internal validity are examples of a **confound** to hypothesis testing. A **confound** is something that varies systematically with the manipulations of the independent variable and may be the actual cause of the observed effect.

For example, suppose a researcher feels that a particular type of therapy leads to improved reading skills among children. To test this hypothesis, the researcher recruits all children from a clinic to be part of the study. Children range in age from 3 to 12 years. Some are in special education; others are in regular classrooms. Children who come to the clinic on Mondays and Wednesdays are assigned to the reading group; children seen on other days are assigned to the control group.

After therapy, some children improved in their reading skills, and others did not. The researcher concludes that the therapy does not work. Is this a valid assumption about the results of the study? It is difficult to judge because there was no mechanism to control for the effects of maturation, selection and history. Any of these factors could have led to these results, regardless of the kind of reading training provided.

Thus, it is important to include ways to **control** for likely confounds, such as subject differences, in an experimental design.

***

The way subjects are **assigned** to groups may determine how well controlled the design is.

Methods of subject assignment to groups:

1) **Random:** Each subject has an equal probability of being assigned to any group. This is comparable to pulling names from a hat.

2) **Block Randomization:** If there are three groups (or levels of a variable to which subjects are assigned), then every three subjects are randomly assigned to one of the three groups.

3) **Stratified Sampling:** a probability sampling procedure in which the population is divided into strata and independent random samples are drawn from each stratum.

4) **Matching:** A way to control for potentially confounding variables such as age or race or education level. Subjects are selected to "match" on potentially confounding variables and are assigned to groups in a way that minimizes the differences between pairs (or sets) of subjects on that variable.

III. SINGLE SUBJECT EXPERIMENTAL DESIGNS

Designs with one treatment (<u>independent variable</u>):

1. **Reversal Designs:** Allows for the reversal of the effects of a treatment as it is applied and removed.

   Where **A** = baseline condition, and **B, C** = treatments (independent variables):

   a) **A→B→A** Baseline is recorded, then treatment is applied, then baseline is obtained again after treatment is removed.

   b) **A→B→A→B** This design is useful in clinical investigations where patients need to be kept on medications.

   c) **B→A→B** Useful for studies that are initiated when treatment is already in progress. You can also use this to determine if a particular treatment remains useful to a patient.

Designs with more than one independent variable:

d) A→B→A→C→A→B & C→A
The individual treatment phases are comparable to the main effects in a factorial design. The "B & C" phase is the interaction effect.

The order of the phases can be set arbitrarily or specifically according to the theoretical approach. The length of each phase depends on the stability of the data and thus may not be able to be determined ahead of time.

IV. MULTIPLE BASELINE DESIGNS

These are useful for studies of treatments or experimental manipulations whose effects are not reversible (e.g., surgery) and which therefore cannot be studied using reversal designs.

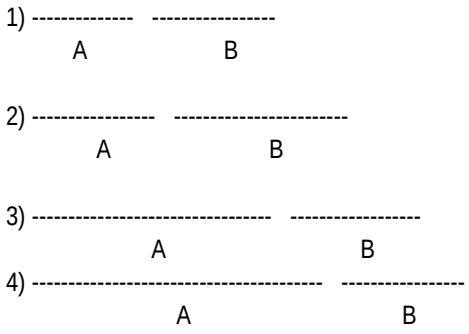
Multiple baseline designs can be applied in several ways:

1) **Across several behaviors (or conditions)** in the same individual. For example, the same treatment can be applied to several problems which are independent of one another in a systematic way such that each treatment begins at a different time to allow for baselines of varying lengths.

2) **Across several situations (or times)** in the same individual. Similar to the above description, however, the treatment is applied to only one problem at different times or in different situations, with each treatment phase beginning at a different time to allow for varying baseline phases.

3) **Across several subjects with the same problem or condition**. This may be more appropriate than reversal designs for medical research. Again, baseline phases would vary across individuals to permit for controls of possible confounds such as history or maturation.

All three types of multiple baseline designs can be set up as follows (the four levels can be four different patients, four different problems within the same person, or four different situations for the same subject):



where **A** = baseline, **B** = treatment, and the length of each line reflects the length of time allocated to each phase of the design.

V. BETWEEN GROUPS DESIGNS

In between group (or between subjects) designs, comparisons are made between groups. Subjects are only exposed to one level of each independent variable. If more than one independent variable is involved (e.g., a two-level variable and a three-level variable), the design would be diagramed as a 2 x 3 cell factorial design.

Suppose the researcher wishes to test the difference between two kinds of dressings (dry versus moist) by level of antibiotic medication (low, moderate, high) on healing of knee abrasions. The study could be diagramed like this:

Dressing (B) **Antibiotic Dose (A)**

	Low	Moderate	High
Dry			
Moist			

Each subject would be exposed to only one cell of the design (e.g., high antibiotics and dry dressing, or low antibiotics and moist dressing).

What are we trying to find in the above example?

1) **Main effect of the independent variable A:** difference in mean scores between three levels of A.

2) **Main effect of the independent variable B:** differences between the means of the two levels of B.

3) **Interaction:** The pattern of scores between the three levels of A is different at each level of B, i.e., the effect of one variable is not the same for each level of the other.

When performing statistical analyses, we would test for an interaction effect first and main effects second. We would try to show statistically that the differences obtained (variance of the scores) in the independent variable were due to this ratio:

$$\frac{\text{Unique influence of the independent variable}}{\text{Random error, subject differences}}$$

The larger this ratio, the more likely we are to find statistical significance.