I. Introduction

1. The between-subjects designs discussed are by no means the only experimental designs available. I will talk about several additional designs including within-subjects designs, mixed designs (which combine aspects of the within-subject and between-subjects designs), and nested designs.

2. In the between-subjects designs, you randomly assign subjects to groups and then expose each group to a single experimental treatment. You measure each subject’s performance on the dependent variable, calculate the average score for each group, and then compare the means to determine whether the independent variable or variables had any apparent influence on the dependent variable.

3. You then subject the data to a statistical analysis to assess the reliability of your conclusions.

4. Within-subjects designs follow a logic similar to that of between-subjects designs. As in between-subjects designs, you expose subjects to different experimental treatments. You average the scores within each treatment, compare the means to determine the influence of the independent variable on performance, and statistically evaluate the differences in performance for reliability.

5. The two types of design differ in the way you expose subjects to the experimental treatments. Instead of exposing different groups of subjects each to different treatment, you expose a single group of subjects to all the treatments.

24 April 2006
6. The within-subject design is called within-subject because the comparison of treatment effects involves looking at changes in performance within each subject across treatments. Each subject's performance is repeatedly measured, and within-subjects designs therefore are sometimes called "repeated measures" designs.

7. Within-subjects designs offer some powerful advantages over the equivalent between-subjects designs if certain conditions can be met. They also introduce problems whose solutions adds complexity to the basic designs, and offer other disadvantages as well.

II. Advantages of the Within-Subjects Design

1. Earlier, I introduced the concept of error variance: differences in subject scores that have nothing to do with the effect of the independent variable.

2. You can see error variance most clearly by comparing of subjects within a particular treatment. Because all subjects are exposed to supposedly identical conditions, their performances ideally should be identical. Yet scores do differ.

3. These differences arise from the effects of extraneous variables, which include relatively stable subject-related factors as well as momentary fluctuations that change each subject's performance from moment to moment.

4. Error variance can be a serious problem, as it tends to mask any effect of your independent variable and can also make it appear that your independent variable had an effect when it did not.

24 April 2006
5. Between-subjects designs, by assigning subjects at random to treatment groups, attempts to spread error variance evenly across the treatment. This approach does not reduce error variance but does make it possible to assess the likelihood that these randomized effects would by chance produce the observed differences in treatment means.

6. If this likelihood is small enough, you conclude that the difference did not result from error variance and therefore that they must have resulted from the effect of your independent variable.

7. Unfortunately, when subject-related factors exert a large influence on performance, error variance may be large enough to obscure any effect of your independent variable. Earlier, I described an approach for dealing with this problem, namely matching subjects on variables thought to have a strong influence on their performances.

8. Matched groups designs help to control error variance by directly comparing the differences in performance among the matched subjects. If matching has been successful, these subjects will perform similarly except for any effect of the independent variable. In other words, there will be less error variance to contribute to performance differences, and the effect of the independent variable will be easier to detect.

9. Within-subjects designs push the logic of the matched groups design to the limit. Each subject is "matched" with other subjects who are virtual clones of each other, because they are all the same subject! This is the ultimate in matching. All subject-related factors are literally identical across treatments.
10. It is as though you had been able to measure and match subjects on every subject factor that could affect the dependent variable. Any performance differences observed between treatments therefore cannot be due to error variance arising from such differences.

11. Because of this ideal matching of subject factors, the within-subjects design tends to be more powerful than an equivalent between-subjects design. A powerful design is one that does a good job of detecting real effects of the independent variables. The increased power of the within-subjects design may allow you to use fewer subjects than you would otherwise need to obtain statistically reliable differences among your treatments.

III. Disadvantages of the Within-Subjects Design

1. Although increased power and reduced subject numbers are strong advantages, the within-subjects design also carries some potentially serious disadvantages, some of which may even preclude its use in certain areas.

2. One disadvantage is that within-subjects designs require each subject to spend more time in the experiment than would the equivalent between-subject design. This is because each subject must participate in every experimental treatment. A complex design involving nine treatments could require a great deal of time to complete. If you find it difficult to get subjects to spend half an hour participating in your experiments, image what will happen when you ask them to spend 4 1/2 hours!

3. You can decrease the potential for fatigue and boredom by administering only one or two treatments per session and spreading the sessions over the week, but if you take this
approach you are likely to experience subject attrition. If the subject fails to show up for Session 4, you will have to discard that subject’s data and much time will have been wasted.

4. A second, and potentially far more serious, disadvantage of within-subjects designs is their ability to produce carryover effects. **Carryover effects** occur when a previous treatment alters the observed behavior in a subsequent treatment.

5. Carryover effects can be a serious problem in any within-subjects design. Between-subjects designs do not suffer from carryover effects, simple because there are no previous conditions from which effects can carry over.

IV. Sources of Carryover

1. **Learning.** If a subject learns how to perform a task in the first treatment, performance is likely to be better if the same or similar tasks are in subsequent treatments.

2. **Fatigue.** If performance in earlier treatments leads to fatigue, then performance in later treatments may deteriorate, regardless of any effect of the independent variable.

3. **Habituation.** Under some conditions, repeated exposure to a stimulus leads to reduced responsiveness to that stimulus. This reduction is termed **habituation.** Your subjects may jump the first time you surprise them with a sudden loud noise, but may not do so after repeated presentations of the noise.
4. **Sensitization.** Sometimes exposure to one stimulus can cause subjects to respond more strongly to another stimulus. In a phenomenon called "potentiated startle," for example, a rat will show an exaggerated startle response to a sudden noise if the rat has recently received a brief foot shock in the same situation.

5. **Contrast.** Because of contrast, exposure to one condition may alter the response of subjects in other conditions. If you pay your subjects a relatively large amount for successful performance on one task, and then pay them less in subsequent task, they may feel underpaid. Consequently, they may work less than they otherwise might have. This change occurs because subjects can compare (contrast) the treatments.

6. **Adaptation.** If subjects go through a period of adaptation (becoming adjusted to the dark, for example), then earlier results may differ from later results because of the adaptive changes. If adaptation to the drug causes reduced response, the change is called **tolerance**.

V. **Dealing with Carryover Effects**

1. You can deal with carryover effects in three ways: (1) you can use counterbalancing to even out carryover effects across treatments, (2) you can take steps to minimize carryover, and (3) you can separate carryover effects from treatment effects by making treatment order an independent variable.

2. **Counterbalancing.** In counterbalancing, you assign the various treatments of the experiment in a different order for different subjects. The goal is to distribute any
carryover equally across treatments so that it does not produce differences in treatment means that could be mistaken for an effect of the independent variable.

3. Two counterbalancing options are complete counterbalancing and partial counterbalancing.

A. **Complete counterbalancing** provides every possible ordering of treatments and assigns at least one subject to each ordering. For example a completely counterbalanced single-factor design that includes three treatment level: ABC, BCA, CAB, CBA, ACB, BAC.

B. The minimum number of subjects required for complete counterbalancing is equal to the number of different orderings of the treatments. N treatments have exactly N! (N factorial) orders. In the above example the number of orders and minimum number of subjects would be (3) (2) (1) = 6.

C. Complete counterbalancing is practical for experiments with a small number of treatments, but this approach becomes increasingly burdensome as the number of treatments grows.

D. **Partial counterbalancing** includes only some of the possible treatment orders. The orders to be retained are chosen randomly from the total set, with the restriction that each treatment appears equally often in each position.
E. If you choose to make the number of treatment orders in your partially counterbalanced design equal to the number of treatment levels, you can use a Latin square design to assure that each treatment appears an equal number of times at each ordinal position.

F. Counterbalancing (whether complete or partial) can be counted on to control order effects only if the order effects induced by different orders are of the same approximate magnitude. In contrast, when the magnitude of the carryover effect differs for different treatment presentation, counterbalancing may be ineffective. This problem is known as differential carryover effects.

G. The most serious asymmetry in carryover effects occurs when a treatment produces irreversible changes. The classic type of irreversible change is that produced by a treatment such as brain lesioning.

3. Taking Steps to Minimize Carryover. The second way to deal with carryover effects is to try to minimize or eliminate them. Of course, you would want to do this only if the carryover effects were not themselves the object of study. Minimizing carryover effects reduces error variance and increases the power of the design.

4. An example of minimizing carryover effect is to pretrain your subjects on a task. For example, you may pretrain subjects on an eye hand coordination task. Get all subject performing at the same level. This will work also for adaptation and habituation.

24 April 2006
5. Another way to deal with habituation, adaptation, and fatigue is to allow breaks between treatments. If sufficiently long, the breaks allow subjects to recover from any habituation, adaptation, or fatigue induced by the previous treatment.

6. **Making Treatment Order an Independent Variable.** A third way to deal with the problem of carryover is to make treatment order an independent variable. For example, if you were going to conduct a one-factor experiment to compare the effect of two memorization strategies on recall, you could design the experiment to include the order of testing as a second independent variable.

7. The main advantage of making order of treatments an independent variable is that you can measure the size of any present carryover effects. You can then take these effects into account for future experiments. If you find that carryover is about equal in magnitude regardless of the order of treatments, for example, then you can be confident that counterbalancing will eliminate any carryover induced bias.

8. Although making treatment order a factor in your experiment can provide important information about the size of carryover effects and can pinpoint the source of differences between findings obtained from within-subjects versus between subjects experiments, the technique does have disadvantages.

9. Every treatment order requires a separate group of subjects. These subjects must be tested under every treatment conditions. The result is a complex, demanding experiment that is costly in terms of numbers of subjects and time of to test them.
10. Furthermore, these demands escalate rapidly as the number of treatments (and therefore number of treatment orders) increases. This latter problem is the same one encountered when using completely counterbalanced designs. For these reasons, the approach is practical only with a small number of treatments.

VI. Types of Within-Subjects Designs

1. Just as with the between-subjects design, the within-subjects design is really a family of designs that incorporate the same basic structure. I will discuss several variations on the within-subjects designs. These variations include the single-factor, multilevel within-subjects designs (in both parametric and nonparametric versions); the multifactor within-subjects design; and multivariate within-subjects designs.

2. The **single-factor two-level design** is the simplest form of the within-subjects design and includes just two levels of a single independent variable. All subjects receive both levels of the variable, but half the subjects receive the treatment in one order and half in the opposite order.

3. The scores within each treatment are then averaged (ignoring the order in which the treatments were given) and the two treatment means are compared.

4. This design is directly comparable to the two-group between-subjects design, including the general advantages and disadvantages of the within-subjects approach. If order effects are not severe and are approximately equal for both orders, then counterbalancing will control the order effects without introducing excessive error variance.
5. Data resulting from a two-treatment within-subjects design can be analyzed by means of the t-test for correlated samples if the data meet the assumptions of a parametric test. A within-subjects single-factor analysis of variance may also be used. If the data does not meet the assumptions of a parametric statistic and the data is nominal, then the McNemar test is used. If the data is ordinal the Sign test or the Wilcoxon matched pairs is used.

6. Just as with the between-subjects design, the within-subjects design can include more than two levels of the independent variable. In the **single-factor, multilevel within-subjects design**, a single group of subjects is exposed to three or more levels of a single independent variable.

7. Just like the between-subjects design, the within-subjects design also has a single-factor parametric design and a single-factor nonparametric design. Let me remind you the parametric in this sense has nothing to do with parametric statistic but with the independent variable.

8. The statistical analysis is the single-factor within-subjects analysis of variance.

9. **Multifactor Within-Subject Designs**, each subject is exposed to every combination of levels of all the factors (independent variables). USE A 4 X 3 FACTORIAL DESIGN AS AN EXAMPLE.

10. A disadvantage of this design is as the number of levels of each factor and/or number of factors increases, so also does the time required to run each subject through all the required treatments. The number of possible orderings of treatments becomes
astronomical with only a few levels per factor and a few factors, thus making complete counterbalancing impractical.

11. Statistical analysis if your data meet the assumptions required for parametric statistics you would use a multifactor within-subjects (or repeated measures) analysis of variance.

12. As was the case with between-subjects designs, within-subjects designs can be extended to included two or more dependent variables. Designs that make use of more than one dependent variable are called **multivariate designs**. Example of multivariate designs are **Factor Analysis**, **Multiple Regression**, and **Multivariate Analysis of Variance** (MANOVA).

VII. Combining Between-Subjects and Within-Subjects Designs

1. There are certain time that it may be appropriate to combine both between-subjects and within-subjects designs. There are two basic designs that do that: the mixed design and the nested design.

2. A **mixed design** is sometimes also called a **split plot design**. The term comes from agricultural research, where the design was first developed (it referred to a plot of land).

3. In the split plot design, a field was divided into several plots. Different plots received different levels of a given treatment (different pesticides). Each plot was then split into subplots, and each subplot received a different level of a second treatment (different fertilizer). Thus each plot received all the levels of fertilizer, but only one level of pesticide.

24 April 2006
4. In psychological research, each plot is a group of subjects who receive the same level of the between-subject variable, whereas each subplot is an individual subject within the group who receives all the levels of the within-subjects variable.

5. For example, Factor A is a between-subject variable and Factor B is a within-subjects factor.

6. An example from psychology, Lorge (1930) was interested in determining whether massed practice ("cramming") or distributed practice (spacing the learning out over several sessions) produced faster learning in a mirror-tracing task. The between-subject variable was massed and distributed practice. The within-subject variable was 20 tracing trials. The dependent variable tracing time in seconds.

7. For data that fit this model a split plot or mixed analysis of variance would be the appropriate analytic tool.

8. Another design that combines within-subjects and between-subjects components is the nested design. An example of a nested design may best explain this design. For example there may be three levels of the between-subjects factor (A₁, A₂, and A₃). Under each of the levels of A are nested three levels of B. Notice that the levels of B found under different levels of A are not the same. For example, B₁, B₂, and B₃ appear under A₁, whereas, B₄, B₅, and B₆ appear under A₂. Each level of Factor A thus includes a within-subjects manipulation of Factor B, although the levels of B included in the manipulation differ with the A level.

24 April 2006
9. Nested designs are useful when you want to include more than one task under a level of an independent variable (nesting tasks). For example, imagine that you were conducting an experiment on the ability to solve anagrams. The between-subject factor might be anagram difficulty (low, moderate, and high). Under each level of anagram difficulty, you might want to include two sets of anagrams, with all subjects in each level of difficulty completing both sets of anagrams.

10. The major advantage of a nested design like this one is that you increase the generality of your results. By demonstrating the effects of items difficulty with several different tasks, you can be more certain that your effect is not limited to a particular type of problem.

11. A nested design can also be useful when you must test subjects in large groups, rather than individually (nesting groups of subjects). For example, you may find it necessary to test subjects during their regularly scheduled class hours. If you tested three classes under each of your experimental conditions, you would then have a class nested under each level of your independent variable.

12. The appropriate analysis for a nested design is a nested analysis of variance.