

Complex Experimental Designs

In this section, we will consider more complex experimental designs. Our topics include:

1. Designs with more than two levels of the independent variable
2. Designs with more than one independent variable - *Factorial Designs*.
3. Interpreting the results from factorial designs.
4. The pragmatics of doing complex designs.

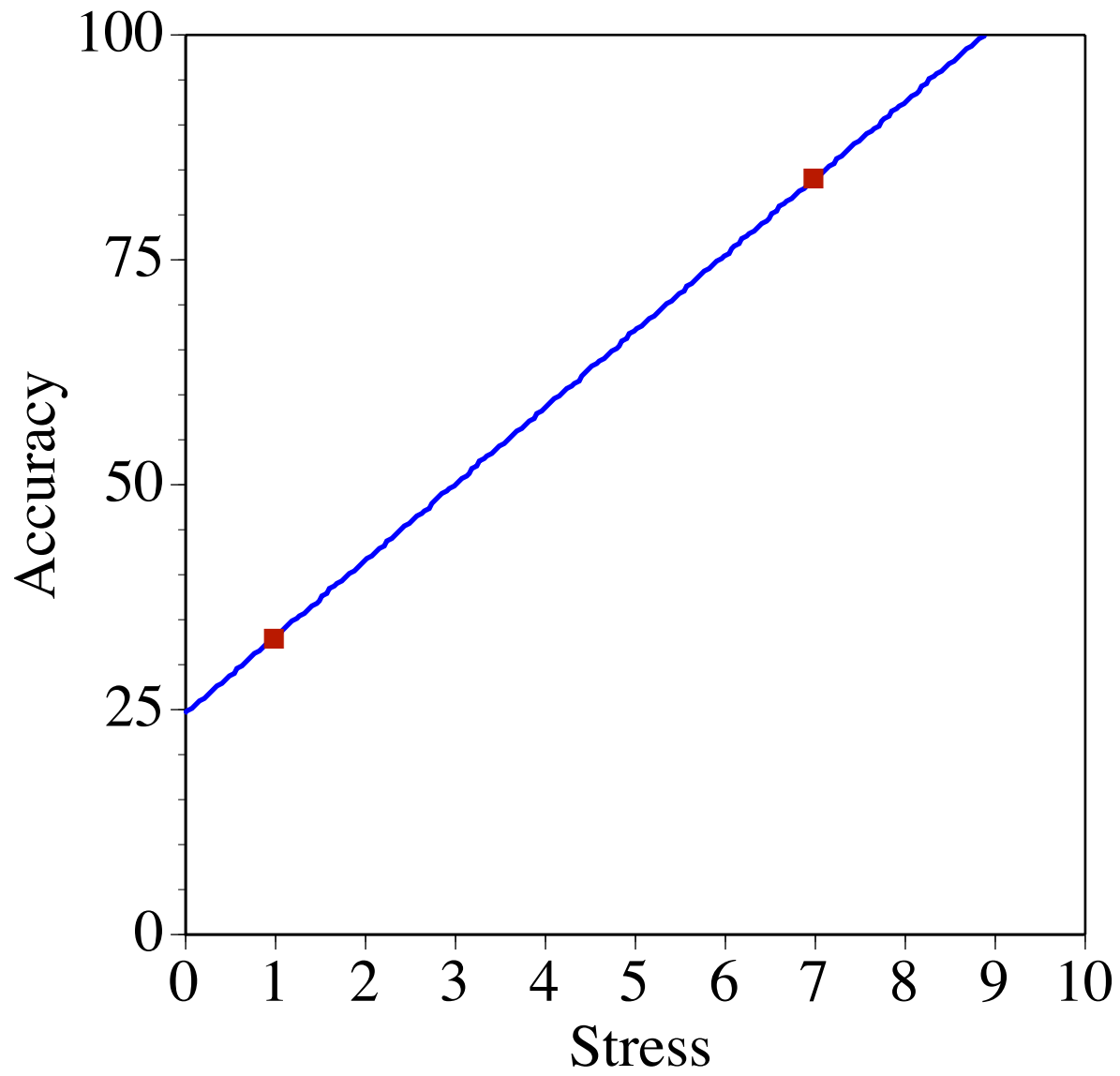
I. Three or More Levels in the Independent Variable

In the simplest experimental design, there are two levels of the independent variable. This allows us to assess the influence of the IV on behavior.

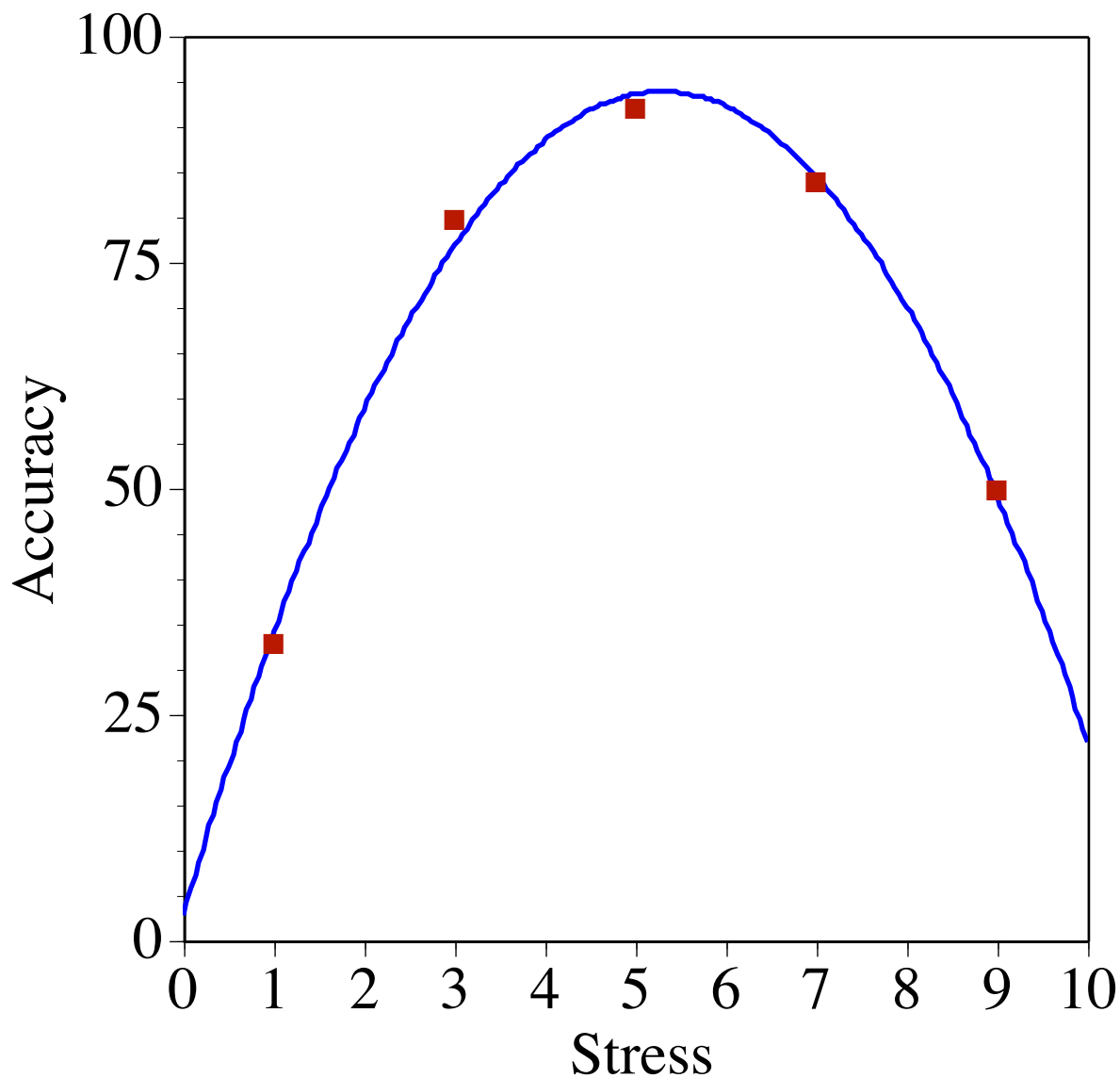
What if the influence of the IV on behavior is complex, and changes with the degree of the IV? Will we be able to tell this from a study with only two conditions?

Consider a study designed to look at the effect of stress upon performance on a demanding motor control task. If we generate just low and moderately high levels of stress in our study, our data might look like the next figure. It appears that the effect of stress on performance is positive.

Is it? What else could we have done?



Now, let us run the experiment with 5 levels of stress in our manipulation: low (1 on the previous graph), med-lo, med, med-high (7 on the previous graph), and high. Our data now look like the curve below.



Clearly, the manipulation involving only two levels of stress did not reveal the true nature of the effect of stress on performance.

By using more than two levels, we have found a curvilinear relationship. Performance is best at intermediate levels of stress.

When there is a range of possible levels for the independent variable, more than 2 are often necessary to obtain an accurate picture of how the IV influences behavior.

While an initial study may only use two levels with a large difference to see if there is any effect, further research would explore the relationship in more detail, especially if there is any reason to believe that the relationship is NOT linear. What would have been seen in the study looking at stress if only two levels, 4 and 7, had been used? What about 3 and 9?

Finally, there may be more than one reasonable operational definition for the treatment condition with your IV. In this case, you might have multiple treatment conditions for comparison to the control condition. Each different treatment would follow a different operational definition of the concept being investigated.

II. Factorial Designs

In the real world, behavior has multiple causes, and these can interact. When this is the case, then the ecological validity of experiments is increased if the experiments use multiple independent variables (corresponding to multiple possible causes) and examine how they interact.

A factorial design is one in which all *levels* on each independent variable occur (are combined with) with all *levels* on each other independent variable. Each combination is called a *condition*.

A) An Example: The Sleeper Effect

Memory about the persuasiveness of a message can change over time. That is, the message in an advertisement becomes more powerful after a delay. However, for this to happen, you also need to be exposed to some sort of "discounting information" that causes you to question the credibility of the original message.

In an experiment, we could present subjects with an advertisement, followed by a discounting cue, then ask them to rate the persuasiveness of the message. A month later, ask them to recall the original advertisement and rate the persuasiveness again. We get a *higher* rating on this second measurement. This is the "sleeper effect". The control group would be run in the same way but *without* the discounting cue. The control group does not get the discounting cue and ratings do not increase.

This simple experiment has two independent variables: presence or absence (control) of discounting cue and delay between presentation and rating (immediate test and delayed test).

Since there are two independent variables and each has two levels, there are 4 conditions in the experiment. This is 2 (immediate test and delayed test) x 2 (discounting cue and no discounting cue).

One of the explanations for the sleeper effect is that the association (in memory) that links the message and the discounting information gets weaker over time. Later recall of the message is less likely to also retrieve the discounting information, so the persuasiveness of the message increases.

Pratkanis et al. (1988) examined this using a 2 x 2 factorial design. Their independent variables:

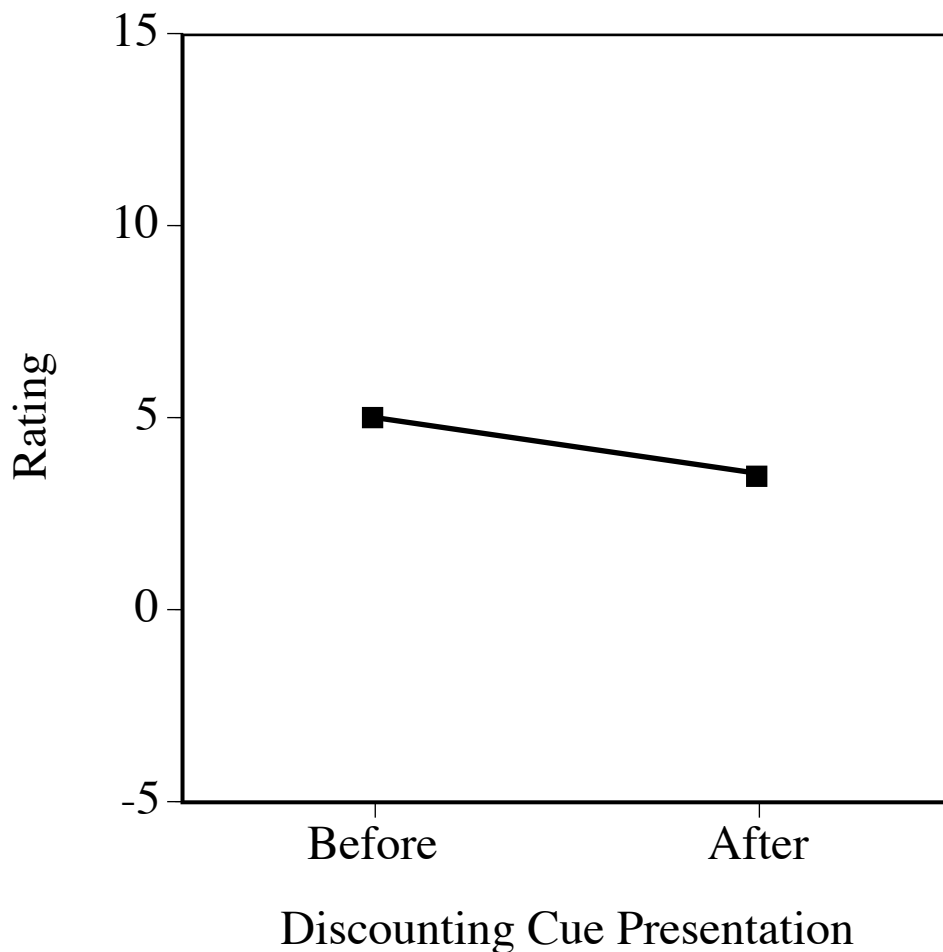
1. Delay between persuasive message and testing (none, immediate testing vs. 6 weeks).
2. Presentation of discounting cue (before message or after message).

These two variables (each has 2 levels) were combined in a 2 x 2 factorial design, which results in 4 experimental conditions.

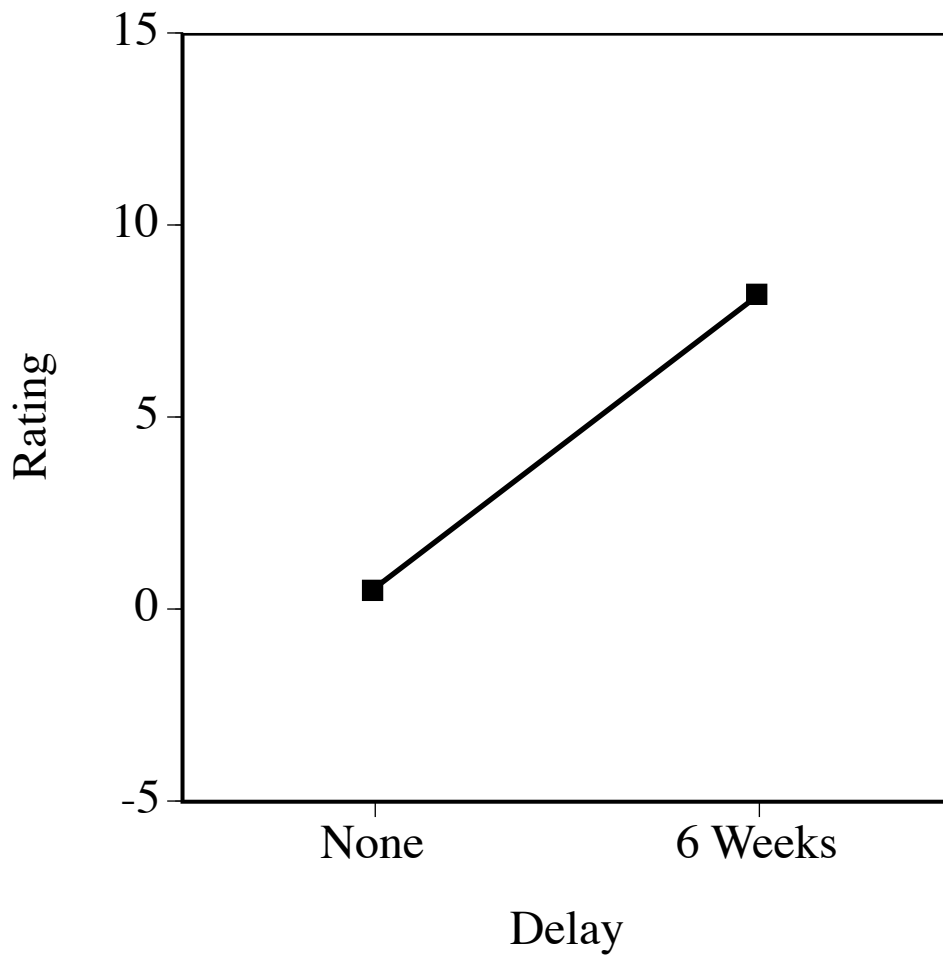
<u>Indep Var 2</u> <u>Delay of Rating</u>	<u>Indep Var 1</u> <u>Discounting Cue</u>	
	<u>Before</u>	<u>After</u>
<u>no delay</u>	before, no delay	after, no delay
<u>6 weeks</u>	before, 6 weeks	after, 6 weeks

First, we'll look at the "main effects", that is, the effect of each independent variable collapsed across the variation in the other variable.

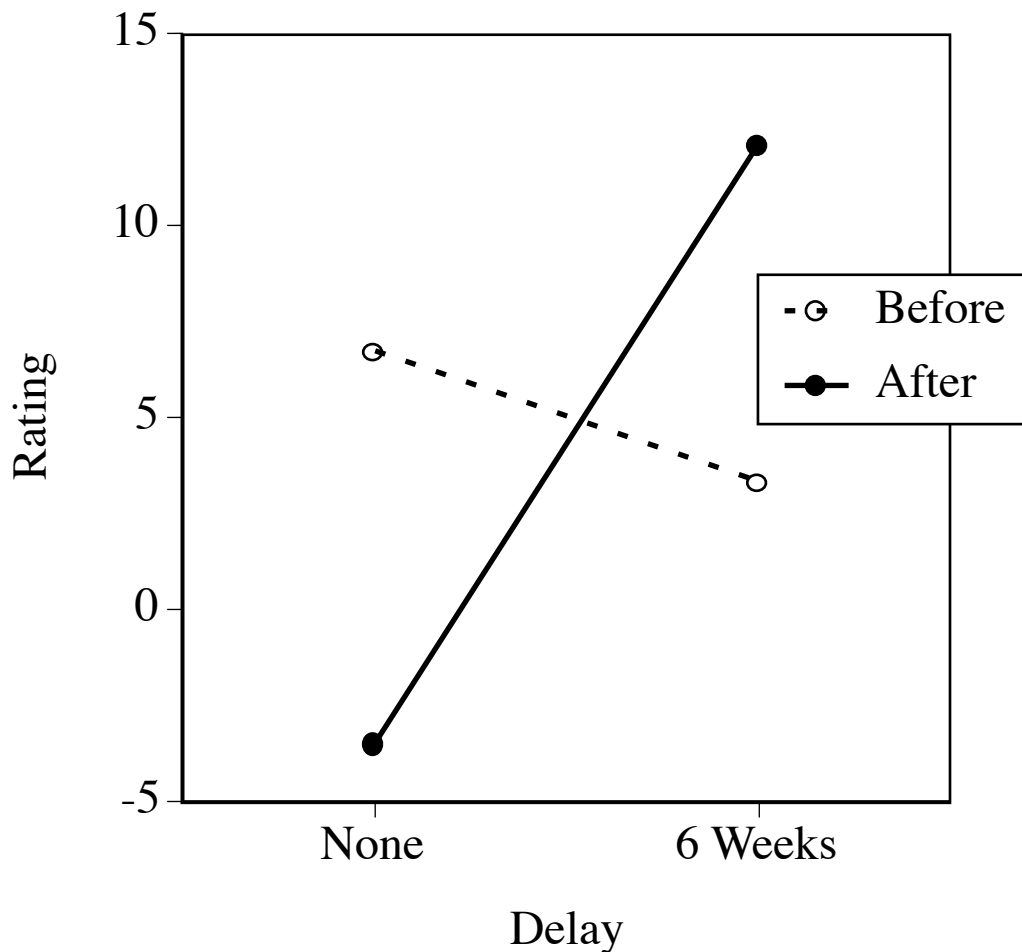
For the Presentation of the Discounting Cue, there was no overall effect of whether it occurred before or after the persuasive message. Only a small, non-significant change in the ratings occurred.



For the delay in rating the persuasiveness of the message, a robust effect was found. Ratings were significantly higher after the 6 week delay.



Most importantly, the two variables interacted. The sleeper effect was only found when the discounting information was presented *after* the persuasive message. When the discounting information was presented before the persuasive message, the rating of the message *went down* over time.



So, it is critical that the discounting information occur after the persuasive message for the sleeper effect to occur.

One explanation is that for the sleeper effect to occur, the persuasive message has to be stored in memory separately from the discounting cue. Then, over time, some forgetting occurs and the message can be remembered separately from the discounting cue. If the discounting cue is presented before the message, then the message is interpreted differently (integrated with the discounting cue) and breaking of the link between message and cue can not occur.

Thus, this experiment provides a valuable insight. However, questions still remain to be answered. For example, we do not know exactly how the subjects encoded the message in the discounting cue before and discounting cue after conditions.

III. Interactions

The basic purpose of a factorial design is to examine whether the independent variables interact. What are the possible patterns of interaction between two independent variables? To illustrate these, we will consider just the case of a 2 x 2 design.

The variables can either interact or not interact. When there is *no interaction*, we call this *additivity*.

There are also two basic types of interaction:

1. simple interaction
2. crossover interaction

That is, there are *three* basic patterns of data that can come out of a 2x2 factorial experiment: additivity, interaction (also called a simple interaction), and a crossover interaction.

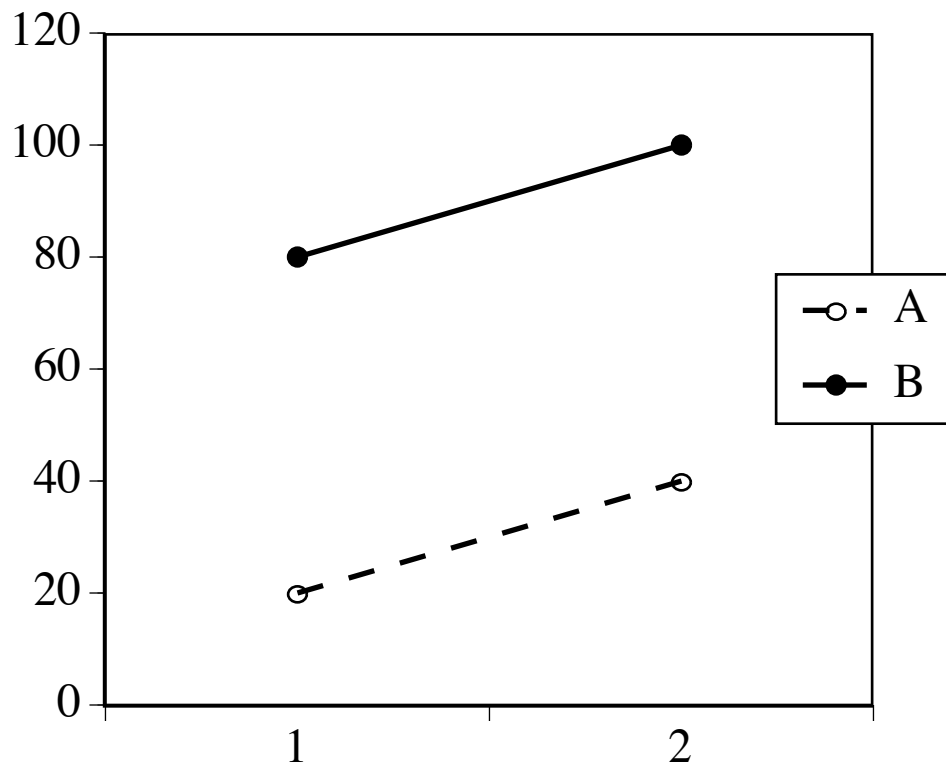
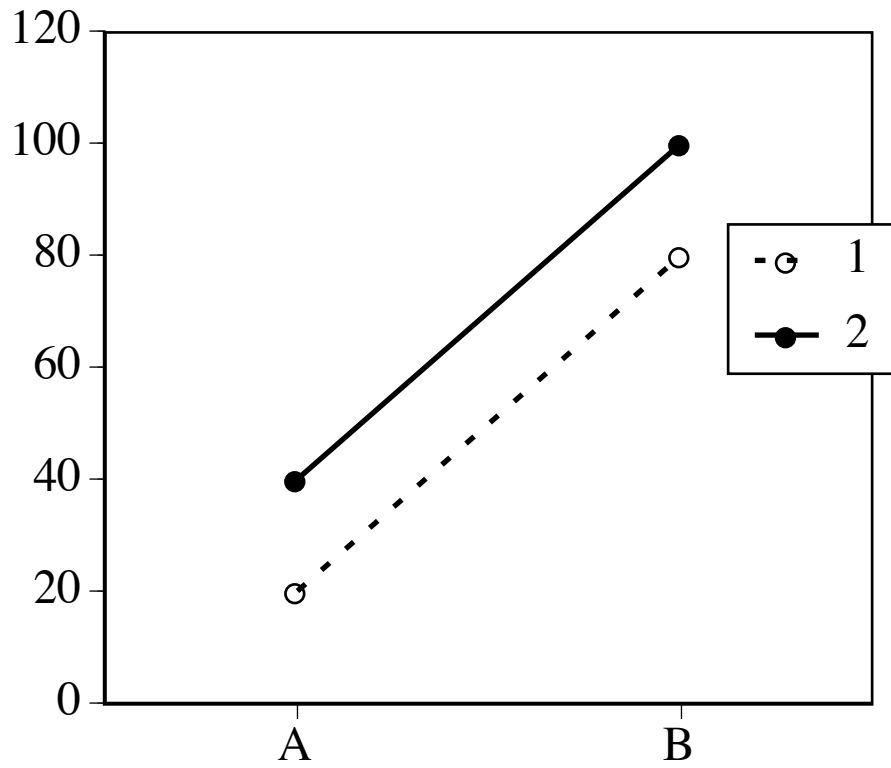
A) *Additivity* - The effects of two variables are additive if the level of one does not influence the results obtained when the other varies. That is, the effects of the two variables add together to produce the results. The influence of any combination is the sum of their individual influences.

Example Data:

Independent Variable (IV) 1

<u>IV2</u>	A	B	Mean
1	20	40	30
2	80	100	90
Mean	50	70	

Additivity



When data that are additive are graphed using a line graph, the lines are parallel. This is true regardless of which IV is used for the x-axis and which is used for the two lines.

However, the actual assessment of whether the data are additive or show one of the two types of interaction is done with statistical tests such as analysis of variance (ANOVA).

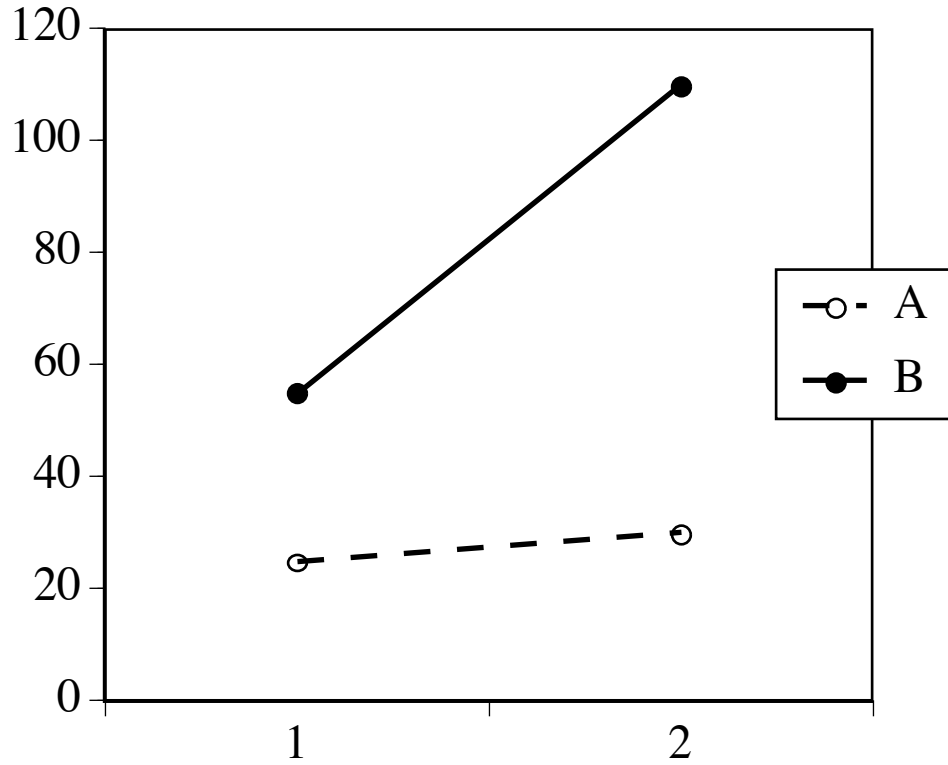
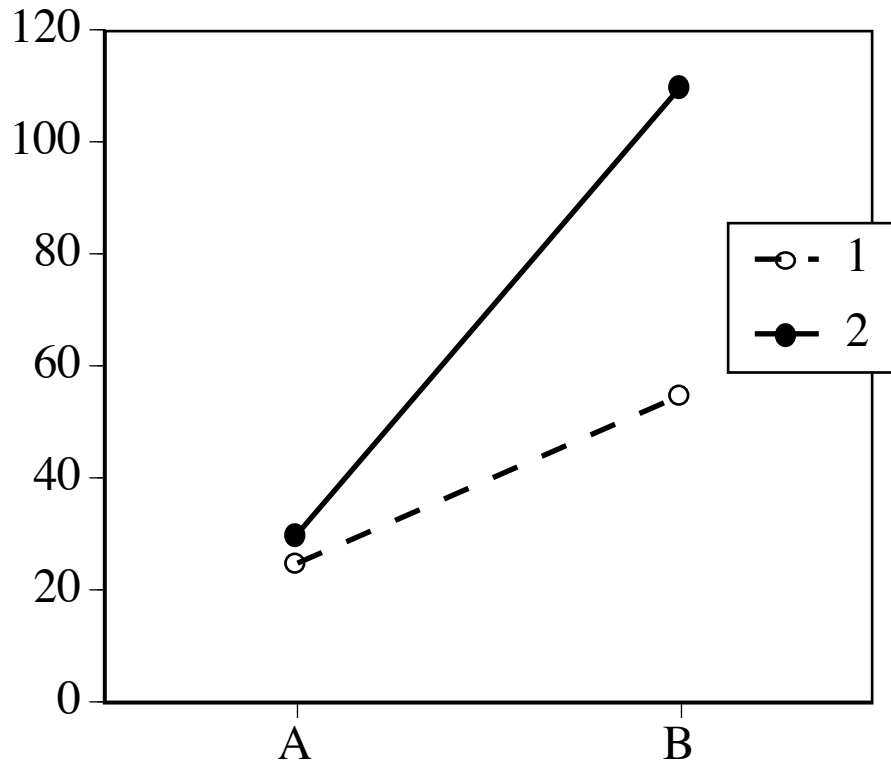
B) *Simple Interaction* - The effects of two variables interact if the level of one *does* influence the results obtained when the other varies. That is, the effects of the one variable are not predictable without knowing the level of the other independent variable.

The simple interaction is illustrated here:

Example Data:

	<u>Independent Variable 1</u>		
<u>IV2</u>	A	B	Mean
1	25	30	27.5
2	55	110	82.5
Mean	40	70	

Interaction



The pattern illustrated here is “over-additive”. That is, the condition B2 (level 2 on one variable and B on the other) has a larger effect than would have happened if the effects of the two independent variables had simply added together. This can be seen as the two lines diverging (spreading apart) from left to right.

The two IVs could also interact and be “under-additive”. In this case, the combined effects of the two IVs together is less than the sum of their separate effects. The two curves would converge (come closer together from left to right).

Both of these patterns are *simple interactions*.

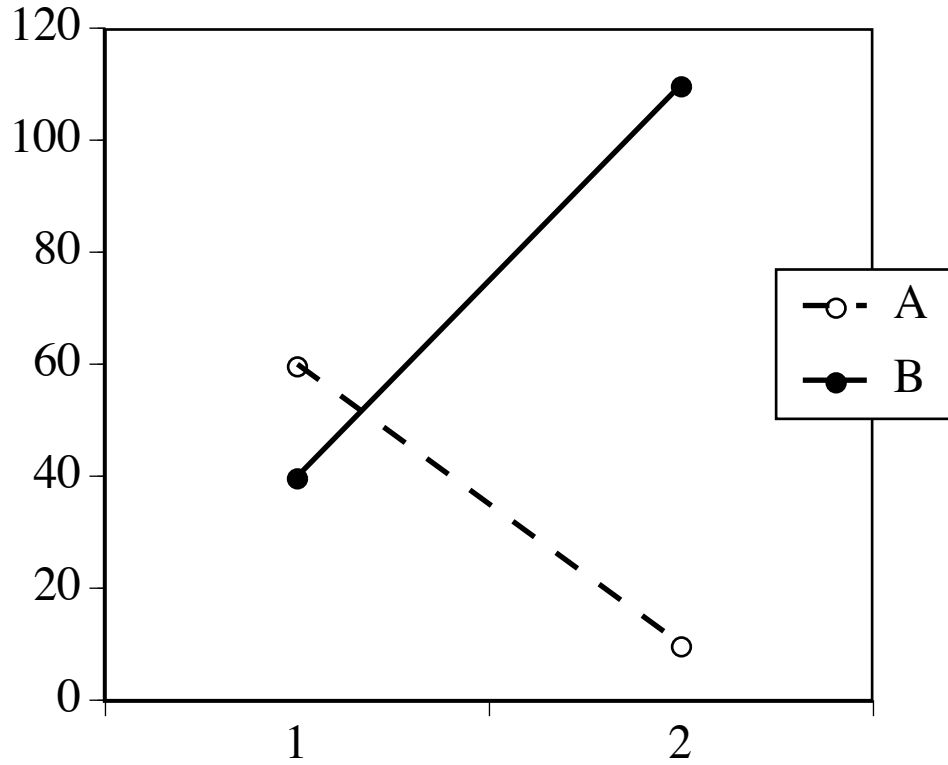
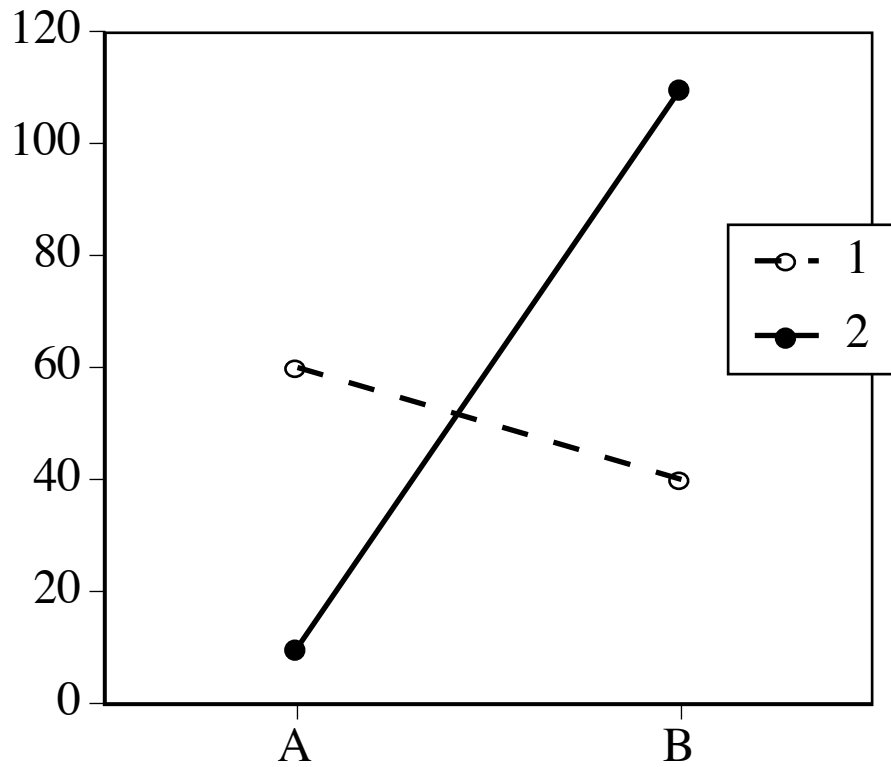
C) *Crossover Interaction* - The effects of two variables interact. Here, the effect of one independent variable on the dependent measure is opposite, depending upon the level of the second independent variable.

The crossover interaction is illustrated here:

Example Data:

	<u>Independent Variable 1</u>		
<u>IV2</u>	A	B	Mean
1	60	10	35
2	40	110	75
Mean	50	60	

Crossover Interaction



The name “crossover interaction” comes from the graph of the data. No matter which IV is used on the x-axis and which is used for the two lines, the two lines cross.

In this case, the effects of the two IVs are antagonistic to one another.

The experiments on the sleeper effect described earlier produced a crossover interaction.

A factorial design is run to understand how the independent variables combine to influence behavior. These three different patterns show that without doing the factorial design, we have no way to accurately predict outcomes in the real world where multiple factors (cause of behavior) are varying.

IV. Factorial Between Subjects Design

The key element to the between design is forming equivalent groups. If the groups are equivalent, then differences between subjects are unlikely to be the cause of any differences that are found on the dependent measure.

A) *The ideal.* A large number of subjects are assigned, randomly, to groups (conditions). All groups (participants) are run at once.

This means that no differences in what takes place in the environment, over time, can cause individuals to differ because they are all run at once. The Random Assignment assures, on average, equivalent groups at the start.

Problems can occur if the participants are not all run at once. If, somehow, subjects were exposed to different events in the environment systematically, in a way that co-varied with the independent variable(s), then we would have a confounding problem. This is a *history effect*.

Example of a History Effect:

The study examines peoples' estimates of the probability of occurrence of various types of events. One independent variable is whether or not subjects have just read an article on traffic accidents as part of another experiment (the control group read an article on inflation).

Are peoples' estimates strongly influenced by information that they recently encountered? Subjects are run one at a time. The Experimental group subjects are run first, then the Control group subjects.

During the course of the study, a big, chain reaction accident occurs on a nearby highway. News stories of the accident are on the internet, radio, TV, and in the paper. Consequently, subjects run after this event are likely to base their probability estimates (at least partly) on news of this event. Since the control group subjects were run second, this group has a greater exposure to this external event. This makes the subjects after the news story different from those before. Now, the two groups are NOT equivalent at the start of the study. This is a *history effect*.

B) Pragmatic design.

We can not always run all of the subjects at once. Sometimes, we can run only one at a time.

We need a method of running our participants that will minimize *history effects*. To do this, create a balanced Latin-square order to use in assigning subjects to groups. If you have n conditions (groups) in your experiment, you need an n item balanced Latin-square.

Assign the first subject to the group denoted by the first item in the first row. The second subject to the second item in the first row, etc. If you finish the Latin-square but still have subjects, start over at the beginning again.

Running participants in different conditions using this approach will distribute any history effects equally across conditions. That is, we control history effects by having them occur equally across all conditions.

V. Factorial Within Subject Design

This design will minimize the number of subjects needed in a factorial design. It also means that we don't have to worry about equivalent groups because each subject serves as her/his own control.

The essential element here is that we must distribute carry-over effects so that the order in which the conditions are run does not influence the pattern of results.

Also, to increase reliability in our dependent measure, we usually obtain multiple observations per subject for each condition. The results for the subject in each condition are then the average of the multiple observations.

Example: Solving Anagrams.

Anagrams are words where the order of the letters has been scrambled (gsenid). To solve the puzzle, you must unscramble the word (design).

Dewing & Hetherington (1974) investigated the influence of imagery and giving hints on solving anagrams.

High imagery words are easier to remember than low imagery words, so anagrams based on high imagery words may be easier to solve (because it is easier to retrieve the word from memory).

Hints usually help people solve anagrams, but the type of hint may interact with the content of the word (e.g. high or low imagery).

The two independent variables were:

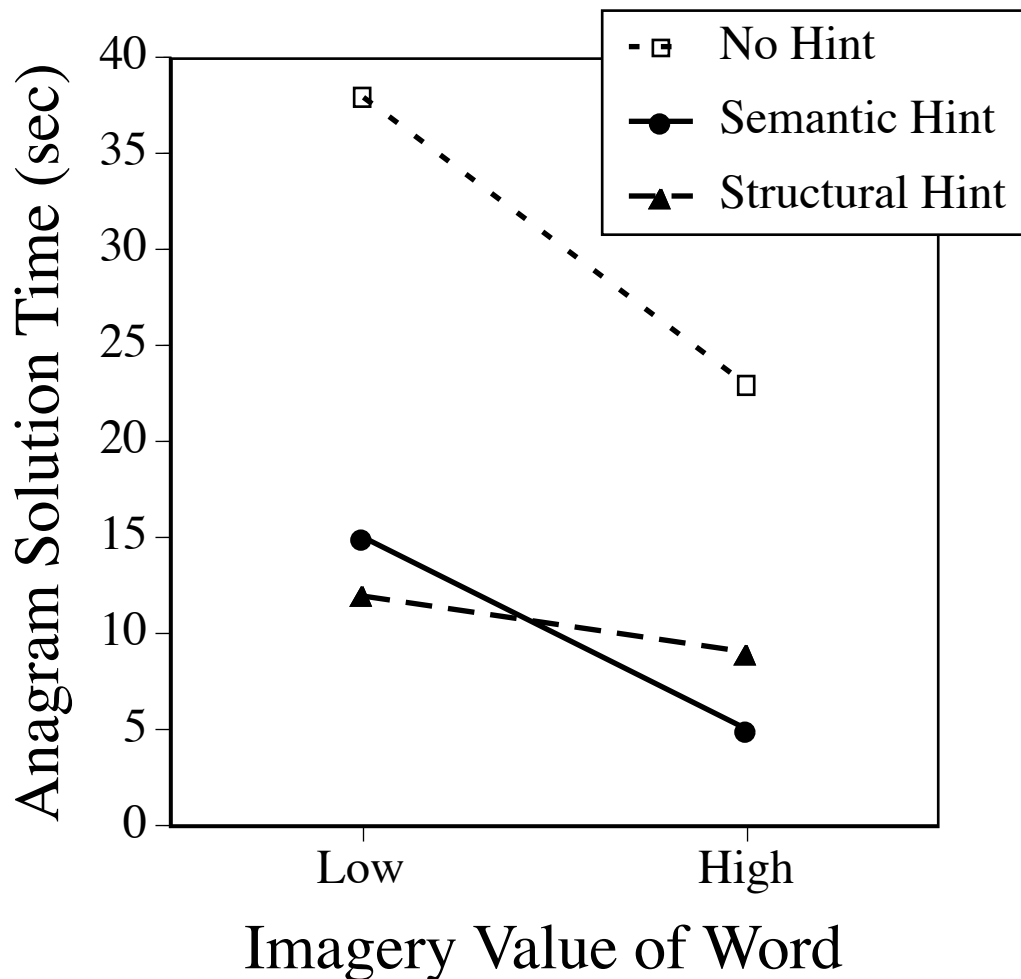
- 1) Imagery value of word (2 levels: high or low)
- 2) Type of hint (3 levels: none, semantic, structural/letter order)

They used a 2 x 3 within subject factorial. Each subject did 12 anagrams, 2 for each of the 6 conditions. The order of the anagrams was *block randomized*. Each of the 72 subjects got one anagram from each of the six conditions in random order. Then they got the second anagram from each condition, in random order.

Counterbalancing and control variables

1. Number of letters per word, all were 5 letter words. (held constant)
2. Word frequency. All words were of about the same frequency. (held constant)
3. Type of word. All were nouns. (held constant)
4. Anagrams. Multiple scramblings (anagrams) were developed for each word. Different subjects got different scramblings. (balanced)
5. Order of conditions. Blocked random counterbalancing.
6. Type of hint per anagram. Completely counterbalanced (3 hint types x 3 anagrams per word).

Results:



The type of hint interacted with the imagery of the word (crossover). Both hint conditions produced much faster solutions than no hint regardless of the imagery for the word, but semantic hints helped high imagery words more and structural hints helped low imagery words more.

VI. Mixed Factorial Designs

Designs with at least one between subjects variable and one within subjects variable are called mixed designs.

A mixed design is used in place of an all within design when:

1. The carry-over effects of an independent variable are too large or permanent to be controlled by counterbalancing.

2. The variable of interest is a subject characteristic which can not be done as a within subject factor. This subject characteristic is treated like an independent variable in the study design (formation of groups of participants).

3. A within design would require too much time from the participants.

VII. Factorial Design Summary

A) Benefits and Approach

Since real world behavior often has multiple factors or variables that influence it, we need an experimental design to investigate multiple factors or variables simultaneously.

Furthermore, these multiple factors can combine in different ways. Their effects could simply add together, they could amplify the effects of one-another (interaction) or they could antagonize or cancel one-another (cross-over interaction). In order to investigate these possibilities, we use a factorial design.

A factorial design has two (or more) independent variables and/or subject variables. All possible combinations of the conditions for each variable are included in the design.

Example - If a study had three independent variables with 2 levels of the first, 4 of the second and 2 of the third, this would result in $2 \times 4 \times 2$ or 16 conditions.

Each independent variable can be done either as a within subject variable or a between subject variable.

Within - This is preferred since each subject is her/his own control.

Between - Can be used for any variable. This is the only way to *assign* a subject variable. Used when carry-over effects preclude within design.

In both designs, we must be careful to distribute (control) any history effects.

B) Disadvantages

Factorial designs take much more work
(subjects, time, etc.)

Factorial designs can produce interactions which are difficult or impossible to interpret. This is especially true if there are five or more factors or if the combination of factors produces conditions that do NOT occur in the real world.

Chapter 10 Sample Exam Questions

- 1) If a study examined the preference for different flavors (bitter, sour, sweet, salty, tart) across individuals with different eye colors and found that people generally preferred sweet tastes over others and that this preference was far higher among people with blue eyes, then we have: a) an interaction b) null results c) three main effects d) a & c above

- 2) Which of the following could reduce our confidence in the results of a between subjects factorial experimental design? a) differential mortality across groups of subjects b) history effects c) carry-over effects d) a & b above

- 3) In a graph of data from a factorial design (with two independent variables), parallel lines would indicate that: a) there is no interaction b) there is at least one interaction present c) there is a significant interaction d) b & c above

4) An important reason for the use of factorial designs in behavioral research is that: a) they can be used to explore multiple potential causes of behavior in one experiment b) factorial designs can be used to explore how variables interact in influencing behavior c) factorial designs can represent a closer analog to the real world than single variable designs d) all of the above

5) In a mixed factorial design: a) one independent variable is tested within-subjects and another is tested between-subjects b) there is more than one dependent measure c) all subjects participate in every condition or treatment d) b & c above

6) If an experiment examined the recall of high and low imagery words following either elaborative or rote rehearsal, then the factors (independent variables) are: a) high and low imagery b) recall c) word imagery and type of rehearsal d) all of the above