Introduction to "Kinds" of 2-way Factorial Designs

- Incorporating within-groups comparisons
- ANOVA for BG, MG & WG designs
- Applying Pairwise Comparisons & LSD_{mmd} to kxk designs

When working with factorial designs we will consider two major types of IVs…

- those that are investigated using a between groups comparison
  - different participants in the conditions of the IV (BG)
- those that are investigated using a within-groups comparison
  - the same participants are in all conditions of the IV (WG)

Thus, there are three "kinds" of 2-way factorial designs:

- **All 3 of which** can involve any combination of manipulated IVs and/or selected IVs
- **All 3 of which** have 3 F-tests
  - one for each of two Main Effects and one for the Interaction

BETWEEN GROUPS FACTORIAL DESIGN:
- both of the IVs use between groups comparison
- each participant completes only one condition of the design

WITHIN-GROUPS FACTORIAL DESIGN:
- both of the IVs use within-groups comparison
- each participant completes all conditions of the design

MIXED FACTORIAL DESIGN:
- one of the IVs uses a between groups comparison and one of the IVs uses a within-groups comparison.
- each participant completes both conditions of the within-groups IV, but completes only one condition of the between groups IV.
- it is important to specify which IV uses a between groups comparison and which IV uses a within-groups comparison

Attrition – also known as drop-out, data loss, response refusal, & experimental mortality – we're gonna have 2 kinds

Attrition endangers the population portion of externa validity

- After carefully obtaining a representative sample, some people drop out
- that attrition could render the sample non-representative
- E.g., the study is "harder" so participant drop out – producing a sample that doesn't represent the motivation of the pop

Differential Attrition endangers initial equivalence part of internal validity (of subject variables)

• random assignment is intended to produce initial equivalence of subject variables – so that the groups (IV conditions) have equivalent means on all subject variables

• e.g., If one condition is "harder" and so more participants drop out of that condition, producing a "motivation" difference between the two conditions.

So, "attrition" works much like "self assignment" to trash initial equivalence

Both involve a non-random determination of who provides data for what condition of the study!

Imagine a study that involves a "standard treatment" and an "experimental treatment"…

• random assignment would be used to ensure that the participants in the two groups are equivalent

• self-assignment is likely to produce non-equivalence (different "kinds" of folks likely to elect the different treatments)

• attrition (i.e., rejecting the randomly assigned condition) is similarly likely to produce non-equivalence (different "kinds" of folks likely to remain in the different treatments)

"Counterbalancing failure" causes Initial Equivalence problems!

Huh?? "counterbalancing failure" happens "during the procedure" – so why isn't it ongoing equivalence!

Initial equivalence – before manipulation of the IV, participants in the conditions are equivalent (on average) on all subject variables.

Ongoing equivalence – during manipulation of the IV, completion of the task, and measurement of the DV, participants in the conditions are equivalent (on average) on all procedural variables (except for the IV).

Two things to notice:

- Initial equivalence is about subject variables
- Ongoing is not "during the study" but during manip of IV, task, measurement of DV – for a particular condition

Keep going…

In a WG design, we are counting on counterbalancing to ensure that the set of participants are "equivalent on average on all subject variables" just before they begin the manipulation of the Control condition, as they are just before they begin the manipulation of the Treatment condition.

If the participants are different on any subject variables before they begin manip of the Control condition than they are before manipulation of the Treatment condition – then those difference on those subject variables is an Initial Equivalence problem!

So, if we don't counter balance (or if counterbalancing fails) then the differences produced are subject variable differences – a problem of Initial Equivalence

Let's say we are studying weightlifting…

We want to see if, for some particular movement, there is a difference in the number of reps a person can do with a "wide grip" vs. a "narrow grip". We decide to use a WG design.

We have everybody do the wide grip trial first and then immediately do the narrow grip trial. See a problem?

What is the problem? The variable of concern is "fatigue" – a subject variable! Participants are more fatigued just before the second narrow grip condition than just before the first wide grip condition!

Doing the study with this improper counterbalancing leads to a "fatigue difference" between the conditions – a subject variable problem → an initial equivalence problem! ;)

Between groups factorial design --

Each participant is in only one condition, having a particular combination of Initial Diagnosis and Type of Treatment.

	Type of Treatment	
Initial Diagnosis	Individual Therapy	Group Therapy
Depression	Clients diagnosed as depressed who are treated with individual therapy	Clients diagnosed as depressed who are treated with group therapy
Social Anxiety	Clients diagnosed with social anxiety who are treated with individual therapy	Clients diagnosed with social anxiety who are treated with group therapy

Mixed group factorial design

Species was a between groups IV (a turtle can only be a member of one species). Each turtle participated in both the mid-morming & dusk conditions of the Time of Day IV.

Time of Day	Species of Turtle: Snapping Turtle	Species of Turtle: Painted Turtle
Mid-morning	Each snapping turtle completed a trial during mid-morning	Each painted turtle completed a trial during mid-morning
Evening	Each snapping turtle completed a trial during the evening	Each painted turtle completed a trial during the evening

Within-groups factorial design --

Each participant completed four trials, one of each combination of Retention Interval and Word Type.

Word Type	Retention Interval: Immediate Test	Retention Interval: Delayed Test
Familiar	The test was given immediately after the study of a list of 40 familiar words.	The test was given 5 minutes after the study of a list of 40 familiar words.
Unfamiliar	The test was given immediately after the study of a list of 40 unfamiliar words.	The test was given 5 minutes after the study of a list of 40 unfamiliar words.

Practice Identifying Types of Factorial Designs - answers next page

The purpose of the study was to examine the possible influence of two variables upon maze-learning by rats, length of the maze (either 10 feet or 30 feet) and the size of the reward (either 1sugar pellet or 5 sugar pellets).

Here are three “versions” of the study tell which is BG, WG & MG

a. Each rat completed one trial. Each was assigned to either the longer or the shorter maze, and also assigned to receive either 1 or 5 sugar pellets upon completing the maze. BG

b. Each rat completed two trials in either the longer or the shorter maze. Following one trial in the assigned maze, each received 1 pellet reward, after the other trial they received the 5 pellets. MG

c. Each rat completed four trials, two in the shorter maze and two in the longer maze. Each received 1 pellet after one of the short-maze trials and 5 pellets after the other, and also 1 pellet after one of the long-maze trials and 5 pellets after the other. WG

Another Example -- 3 versions of the same study

The researcher wanted to investigate infant's startle responses to loud sounds. The two variables of interest were the Position of the Sound (in front of versus behind the infant) and the Type of Sound (a hand-clap versus deep male voice saying "Hey").

Here are three “versions” of the study tell which is BG, WG & MG

a. Each infant completed trials all involving a hand-clap or all involving the voice saying "Hey". During some of the trials, the appropriate type of sound was made in front of the infant. During other trials, the appropriate type of sound was made behind the infant. MG

b. Each infant had some trials during which the sound was made in front of then and some during which the sound was made behind them. Some of the sounds were the hand-clap and the others were the voice saying "Hey". WG

c. Each infant always heard either the hand-clap or the “Hey”, and whatever sound they heard was always played either in front of them or behind them. BG

Remember about the causal interpretation of effects of a factorial design

Start by assessing the causal interpretability of each main effect

Remember, in order to causally interpret an interaction, you must be able to casually interpret BOTH main effects.

For each of the following: Tell the IVs and tell what effects could be causally interpreted (assuming proper RA, IV manip. and confound control were used):

1. Novice and Expert golfers, who preferred Ping, Callaway or Titleist clubs hit their driver. The DV was distance from the tee until first bounce. zilcho-causo

2. Children played with either a toy gun, a toy car or a puzzle, some while their parents were in the room and some not. The DV was the amount of aggressive behavior they exhibited. All three !

3. Participants played with either a simple puzzle or a complex puzzle in pairs made up of two puzzle lovers, two puzzle haters, or one puzzle lover & one puzzle hater. Puzzle type only.

F-tests of the BG Factorial Designs

$$SS_{total} = SS_A + SS_B + SS_{INT} + SS_{Error}$$

$$df_{total} = df_A + df_B + df_{INT} + df_{Error}$$

$$(N - 1) = (a - 1) + (b-1) + (a-1)(b-1) + ab(n-1)$$

$$F_A = \frac{SS_A / df_A}{SS_E / df_E} \qquad F_B = \frac{SS_B / df_B}{SS_E / df_E} \qquad F_{INT} = \frac{SS_{INT} / df_{INT}}{SS_E / df_E}$$

Things to notice:

- There is a single error term that is used for all the Fs
- All of the effects are “equally powerful” (all use same sample size -- power depends upon sample size)

F-tests of the MG Factorial Designs

$$SS_{total} = SS_A + SS_{S/A} + SS_B + SS_{INT} + SS_{BxS/A}$$

$$df_{total} = df_A + df_{S/A} + df_B + df_{INT} + df_{BxS/A}$$

$$(N - 1) = (a - 1) + a(S-1) + (b-1) + (a-1)(b-1) + b(S-1)$$

$$F_A = \frac{SS_A / df_A}{SS_{S/A} / df_{S/A}} \qquad F_B = \frac{SS_B / df_B}{SS_{BxS/A} / df_{BxS/A}} \qquad F_{INT} = \frac{SS_{INT} / df_{INT}}{SS_{BxS/A} / df_{bxS/A}}$$

Things to notice:

- There are two error terms -- one for the BG main effect and another for the WG main effect and the interaction
- The WG ME test and the interaction are usually more powerful, than the BG ME
- BG main effect is less powerful in this design than for BG design

F-tests of the WG Factorial Designs

$$SS_{total} = SS_{Sub} + SS_A + SS_{SxA} + SS_B + SS_{SxB} + SS_{INT} + SS_{SxI}$$

$$df_{total} = df_{Sub} + df_A + df_{SxA} + df_B + df_{SxB} + df_{INT} + df_{SxI}$$

$$(N - 1) = (S-1) + (a - 1) + (S-1)(a-1) + (b-1) + (S-1)(b-1) + (a-1)(b-1) + (S-1)(a-1)(b-1)$$

$$F_A = \frac{SS_A / df_A}{SS_{SxA} / df_{SxA}} \qquad F_B = \frac{SS_B / df_B}{SS_{SxB} / df_{SxB}} \qquad F_{INT} = \frac{SS_{INT} / df_{INT}}{SS_{SxINT} / df_{SxINT}}$$

Things to notice:

- There is a separate error term that is used for each F
- These tests are about equally powerful as the WG effects of the MG design, and more powerful than the BG effects of the other

Pairwise Comparisons & LSD Follow-ups

As for the BG factorials, any design larger than a 2x2 will require follow-up analyses cell and/or marginal means

- Main Effects
  - any significant ME with more than 2 conditions will require follow-ups to describe the pattern of marginal means
  - follow-up of the corresponding simple effects will be required to determine if that ME is descriptive or potentially misleading
  - LSDs use the error term specific to that ME F-test
- Interaction
  - the LSD minimum mean difference can be used for either 2- or k-condition simple effects, to describe the pattern
  - LSD uses WG error term from the interaction F-test