The Overall Two-Factor Analysis

ERSH 8310

Lecture 10
October 28, 2009

Today's Class

- Chapter 11 (sections 5-7)
  - Blocking factors
  - Measuring effect size
  - Determining sample size with power analyses

- Chapter 12
  - Interpreting a Two-Way Design
  - Comparing the Marginal Means
  - Interpreting the Interaction
  - Testing the Simple Effects
  - Simple Comparisons
  - Effect Sizes and Power for Simple Effects
  - Controlling Familywise Type-I Error

- Chapter 13 (not discussed this week – see next week)

DESIGNS WITH A BLOCKING FACTOR

Recall From Last Week…

- Last week we discovered the two-way ANOVA model
  - Two IVs
  - Main effects
  - Interactions
- This week, we begin by introducing the concept of using one of the IVs to help control for experimental errors
  - Called a blocking factor
- Blocking factors are statistical methods that control for factors that may adversely impact the results of an experiment

The Randomized-Blocks Design

- You can use a blocking factor to capture variability that is irrelevant to the effect of interest and thereby reduce the size of error term
  - Makes groups of subjects more homogeneous within a block

- In a randomized-blocks design, the blocking is part of the original study and controls the assignment of subjects to groups

RBD Example

- Suppose that a researcher is investigating the effects of four sets of instructional material on how well college students learn a body of quantitative material
  - For example, say statistics
- The simplest procedure is to obtain a sample of 60 students
  - Randomly assign n=15 subjects to each of four groups to create a completely randomized single-factor experiment
- Suppose the researcher realizes there is great variability in the student performance arising from differences in their quantitative skills before the study started

More of the RBD Example

- The variability makes the $MS_{S/A}$ large and limits the power of the design to detect differences among the instruction conditions
  - To increase power one can increase the sample size – but here cannot do that
  - Another way to increase power is to decrease the variability of the scores

- Creating a blocking factor will aid in decreasing the variability of the scores

Table 11.10

Table 11.10: A comparison of a completely randomized single-factor design and a randomized-blocks design.

Completely randomized (unblocked) design

Instructions (Factor A)			
a_1	a_2	a_3	a_4
$n = 15$	$n = 15$	$n = 15$	$n = 15$

Source	df
A	$a - 1 = 3$
S/A	$a(n-1) = 56$
Total	$an - 1 = 59$

Randomized-blocks design

	Instructions (Factor A)			
Blocks	a_1	a_2	a_3	a_4
b_1	$n = 5$	$n = 5$	$n = 5$	$n = 5$
b_2	$n = 5$	$n = 5$	$n = 5$	$n = 5$
b_3	$n = 5$	$n = 5$	$n = 5$	$n = 5$

Source	df
A	$a - 1 = 3$
B	$b - 1 = 2$
$A \times B$	$(a-1)(b-1) = 6$
S/AB	$ab(n-1) = 48$
Total	$abn - 1 = 59$

Post-Hoc Blocking

- In a post-hoc design, the second factor is created after the data are collected
  - Likely to be more common in quasi-experimental research
- Use the analysis of covariance (see Chapter 15) when the potential blocking information is available as a numerical quantity
  - A quantitative factor
- Using blocking factors is analogous to ANCOVA with qualitative/categorical IVs

MEASURING EFFECT SIZE

A Numerical Example from Last Class

- A hypothetical investigation of the role of certain drugs [factor A-Control (a1), Drug X (a2), Drug Y (a3)] and drive level [factor B-1 hour of food deprivation (b1), 24 hour of food deprivation (b2)] on learning performance (Y) of monkeys

- The animals are given a series of 20 ``oddity'' problems and the response measure Y is the number of errors in the 20 training trials

- The design is a 3×2 factorial with a cell sample size of n = 4

Statistical Output

Between-Subjects Factors

		N
Drug Condition (Factor A)	1.00	8
	2.00	8
	3.00	8
Deprivation Condition (Factor B)	1.00	12
	2.00	12

Main Effect Test for Factor A

Main Effect Test for Factor B

Interaction Test for AxB

Tests of Between-Subjects Effects

Dependent Variable: Errors (Y)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	280.000[a]	5	56.000	3.055	.036
Intercept	2400.000	1	2400.000	130.909	.000
drug	112.000	2	56.000	3.055	.072
deprive	24.000	1	24.000	1.309	.268
drug * deprive	144.000	2	72.000	3.927	.038
Error	330.000	18	18.333		
Total	3010.000	24			
Corrected Total	610.000	23			

a. R Squared = .459 (Adjusted R Squared = .309)

Reasoning About Effect Size Calculations

- I have had a change of heart - I am now changing the way we discuss effect size estimates

- This is due to several factors
  - Hand calculations can cause problems
  - The problems having to calculate effect sizes by hand

- We will use a slightly different measure of effect size
  - Given in output by SPSS
  - May be an over-estimate but is still approximately correct

Introducing…The Partial Eta Squared

- SPSS provides an easy to calculate measure of effect size called the partial eta squared:

$$\hat{\eta}_{\langle effect \rangle} = \frac{SS_{Effect}}{SS_{Effect} + SS_{Error}}$$

- Again, the measure may result in an over-estimate of your effect
  - However, it is very easy to compute in all situations
  - You should remember that bias can result and that other measures of effect size do exist (omega squared)
- The same metric of “size” applies

Partial Eta Squared – From SPSS

- From SPSS:

Tests of Between-Subjects Effects

Dependent Variable:errors

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	280.000[a]	5	56.000	3.055	.036	.459
Intercept	2400.000	1	2400.000	130.909	.000	.879
drug	112.000	2	56.000	3.055	.072	.253
deprivation	24.000	1	24.000	1.309	.268	.068
drug * deprivation	144.000	2	72.000	3.927	.038	.304
Error	330.000	18	18.333			
Total	3010.000	24				
Corrected Total	610.000	23				

a. R Squared = .459 (Adjusted R Squared = .309)

More Descriptive Measures

- Another descriptive measure is the standard difference between means
  - Here SS is the within-group sums of squares

- For example,

$$d_{a_1,b_1} = \frac{\bar{Y}_{11} - \bar{Y}_{21}}{\left(\dfrac{SS_{11} + SS_{21}}{df_{11} + df_{21}}\right)}$$

Chapter 12

DETAILED ANALYSIS OF MAIN EFFECTS AND SIMPLE EFFECTS

Detailed Analysis of Main Effects and Simple Effects

- The test for interaction comes first
  - Outcome influences all the analyses that follow

- Significant Interactions:
  - Less attention paid to the two main effects
  - Analysis tends to focus on the individual cell means
    - The joint variation of the two independent variables

- Non-Significant Interactions:
  - Attention is directed to the marginal means
  - The variation of each IV is considered in absence of the other

Detailed Analysis of Main Effects and Simple Effects

- The analysis of any study must return eventually to the actual pattern of means
- You cannot just say that one factor is significant and another is not
  - Or that an interaction is or is not present
- A detailed description of the resulting patterns of means is always necessary
  - The focus of Chapter 12 (for simple effects and main effects)

A picture is worth a thousand words.

INTERPRETING A TWO-WAY DESIGN

Interpreting a Two-Way Design

- The first step in examining data from a factorial study is to plot the means

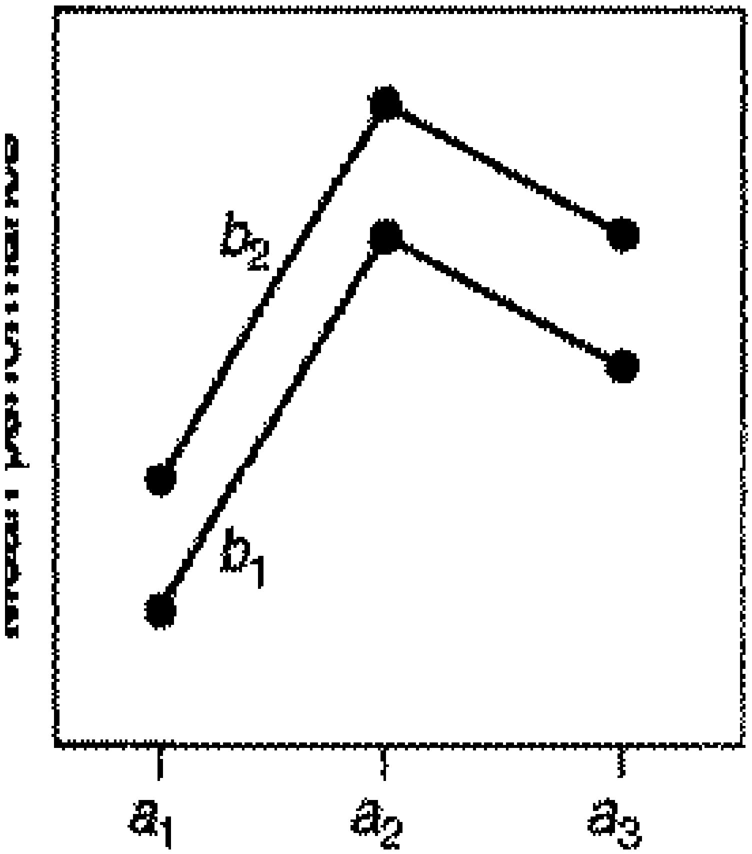
- Line graphs are usually clearer than bargraphs, particularly when exploring the data
  - Try several plots before finding a good representation

- The pattern of means from a factorial design can be expressed as:
  - Main effects
  - Simple effects
  - Interaction components
  - Special patterns implied by theory

Possible Outcomes

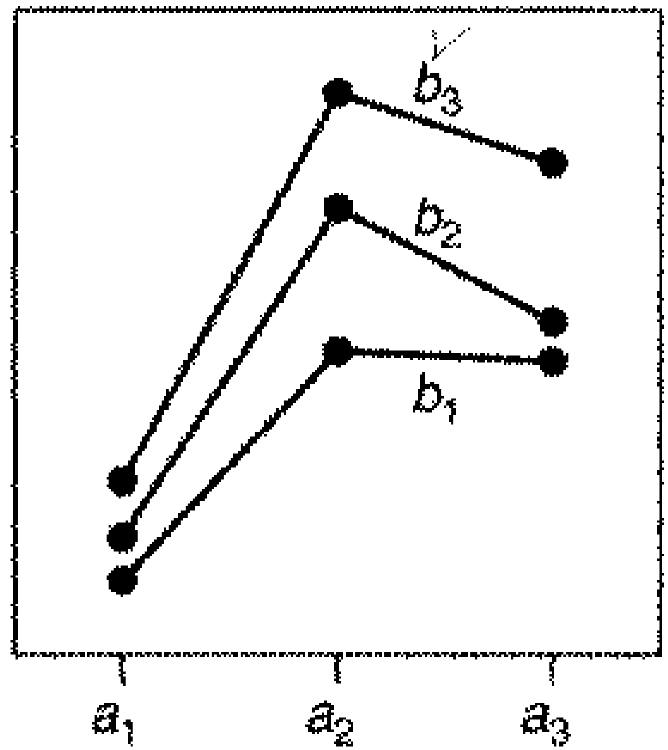
Three possible outcomes that might dictate the subsequent analyses are:

1. No interaction is present
   - The two-way design is reduced to multiple one-way effects
   - Attention is directed at follow-up tests that investigate analytical questions about the marginal means



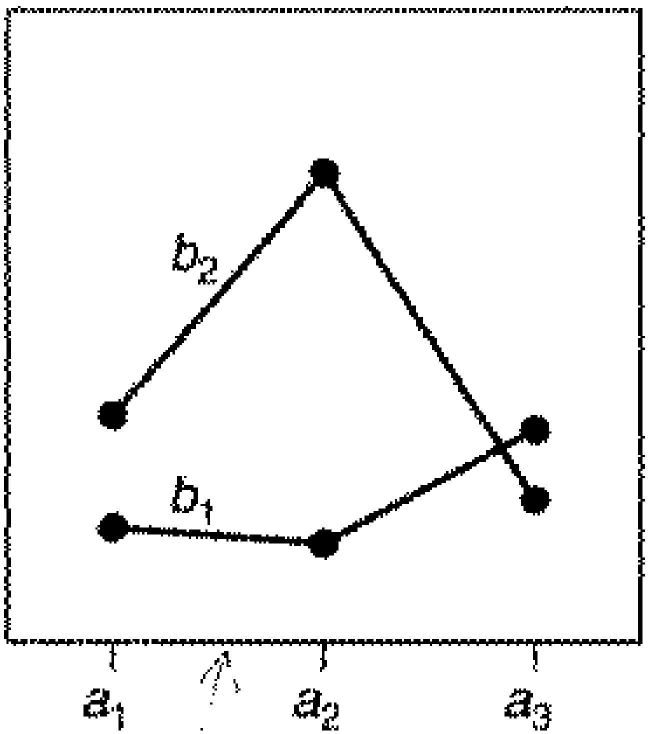
Possible Outcome #2

2. An interaction is present, but it is dominated by the main effects.
   - The effect of either factor changes with the levels of the other
   - The simple picture of two main effects is not appropriate, and the two factors cannot be treated completely separately
   - Must consider how the simple effects of one factor differ with the levels of the other
   - Main effects that dominate the interaction usually represent solid, well-known and often-replicated manipulations



Possible Outcome #3

3. The interaction dominates the main effects
   - It can be deceptive to look at the marginal effects as all
   - We would be justified in ignoring the main effects altogether



COMPARING THE MARGINAL MEANS

Comparing the Marginal Means

- We have two sets of marginal means, one for each of the two factors in the design
  - Factor A
  - Factor B

- The significance of comparisons is evaluated with the error term from the overall analysis, namely, $MS_{S/AB}$

- Consider this section what to do when no interaction is present
  - Interaction is not significant

First…Our Example Data

- The marginal means

errors * drug

errors

drug	Mean	N	Std. Deviation
Control	7.00	8	5.398
Drug X	11.00	8	4.870
Drug Y	12.00	8	4.276
Total	10.00	24	5.150

errors * deprivation

errors

deprivation	Mean	N	Std. Deviation
1-hour deprivation	9.00	12	5.970
24-hour deprivation	11.00	12	4.200
Total	10.00	24	5.150

- The ANOVA table
  - We will use $MS_{S/AB}$

Tests of Between-Subjects Effects

Dependent Variable:errors

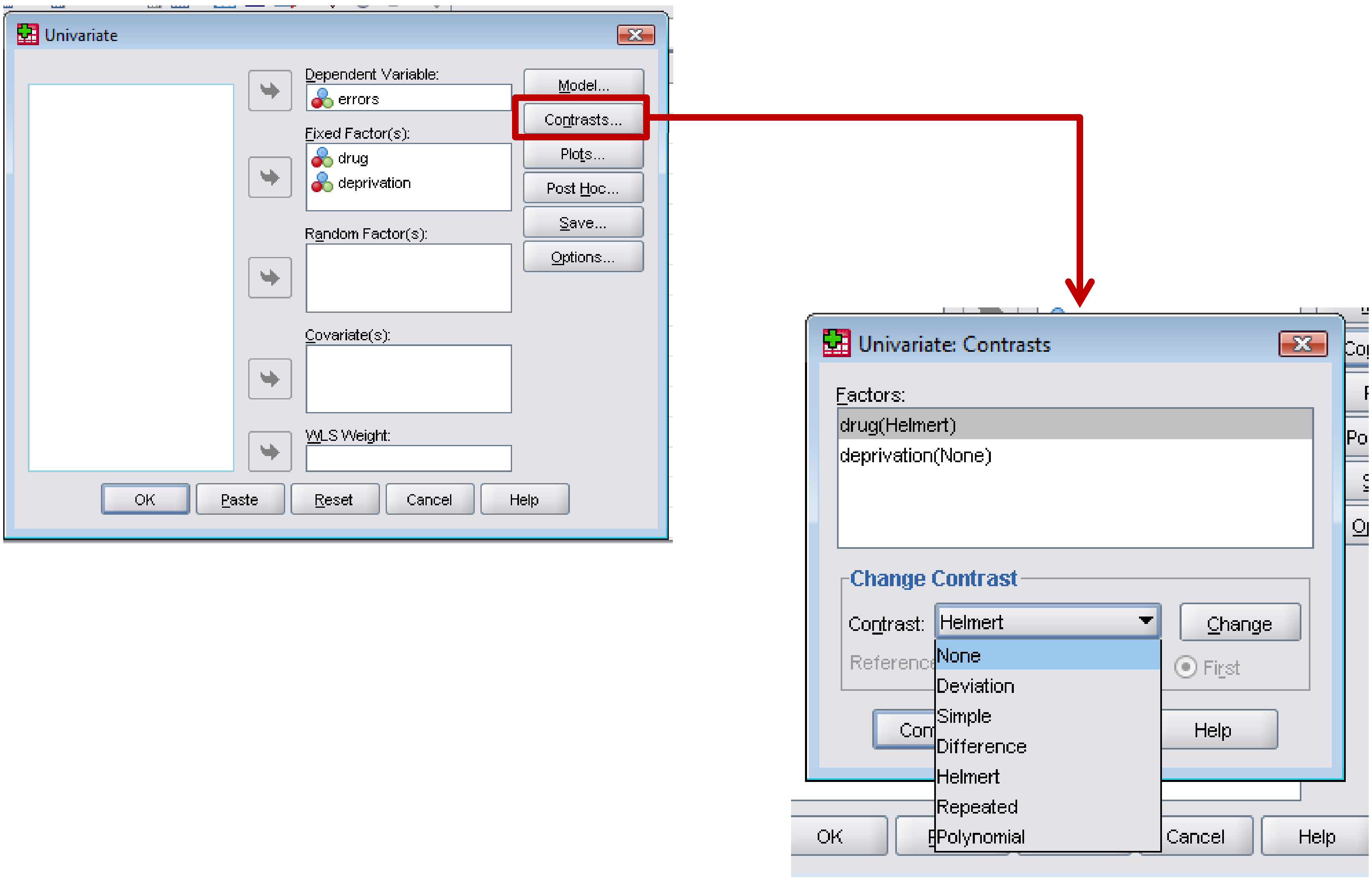
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	280.000[a]	5	56.000	3.055	.036	.459
Intercept	2400.000	1	2400.000	130.909	.000	.879
drug	112.000	2	56.000	3.055	.072	.253
deprivation	24.000	1	24.000	1.309	.268	.068
drug * deprivation	144.000	2	72.000	3.927	.038	.304
Error	330.000	18	18.333			
Total	3010.000	24				
Corrected Total	610.000	23				

a. R Squared = .459 (Adjusted R Squared = .309)

Torturing Marginal Means

- Examinations of marginal means are just like what we carried out on our single factor in one-way ANOVA
  - We can construct ***<u>CONTRASTS</u>***?
- Just as in one-way ANOVA, contrasts are formed by selecting coefficients that multiply the marginal means
- For this example, <u>pretend</u> we do not have a significant interaction present
  - Let’s compare both drugs with the control group, simultaneously
- Everything is the same as in the one-way ANOVA
  - The denominator, however, is $MS_{S/AB}$

Contrasting Main Effects in SPSS



Our Results

Contrast Results (K Matrix)

drug Helmert Contrast			Depende... errors
Level 1 vs. Later	Contrast Estimate		-4.500
	Hypothesized Value		0
	Difference (Estimate - Hypothesized)		-4.500
	Std. Error		1.854
	Sig.		.026
	95% Confidence Interval for Difference	Lower Bound	-8.395
		Upper Bound	-.605
Level 2 vs. Level 3	Contrast Estimate		-1.000
	Hypothesized Value		0
	Difference (Estimate - Hypothesized)		-1.000
	Std. Error		2.141
	Sig.		.646
	95% Confidence Interval for Difference	Lower Bound	-5.498
		Upper Bound	3.498

Contrast of…

Ψ

P-value

So…What does this mean?

INTERPRETING THE INTERACTION

Interpreting the Interaction

- Now…what if we had a significant interaction?
  - We will use the simple effects (slices) to investigate our interaction more thoroughly

- Interactions can be analyzed in two ways:
  1. The analysis of the simple effects
     - Chapter 12
  2. The analysis of interaction comparisons
     - Chapter 13

Selecting a Set of Simple Effects for Analysis

- Analyze the set of simple effects that is the most natural, useful, or potentially revealing
  - The manipulation that will be the easiest to explain
  - The choice really is arbitrary
    - You pick what tells the best story

- So, choose:
  - The factor with the greater number of levels
  - A quantitative factor
  - The factor with the greater main-effect sum of squares
  - A manipulated factor

TESTING THE SIMPLE EFFECTS

Testing the Simple Effects

- Simple effects are based on the differences among the cell means within a particular row or column of the matrix of means
  - Slices of the experiment

- No easy way to get SPSS to cooperate to provide estimates or significance tests
  - Must build your own augmented ANOVA tables
  - Recall the last example from lab last week

Our Example…

- In our example data, a significant interaction was found.
  - Let's examine the simple effects of drug (factor A) at the two different deprivation conditions (factor B)
  - Basically, we must ask whether or not there is a significant difference in mean errors across drug type for 1-hour deprivation **<u>and</u>** for the 24-hour condition, separately

- Steps for analysis:
  1. Run two-way ANOVA
     - Note the S/AB source ($SS_{S/AB}$, $df_{S/AB}$, and $MS_{S/AB}$) - to be used as the denominator in our analyses
  2. Run two one-way ANOVAs
     - Splitting the SPSS file by deprivation condition
     - Gives SS, df, and MS for each level of b (A at b_1 and A at b_2)
  3. Compile all results into a single ANOVA table
     - The S/AB is from the two-way condition
     - The other terms are from the one-way ANOVAs

Step 1...Check

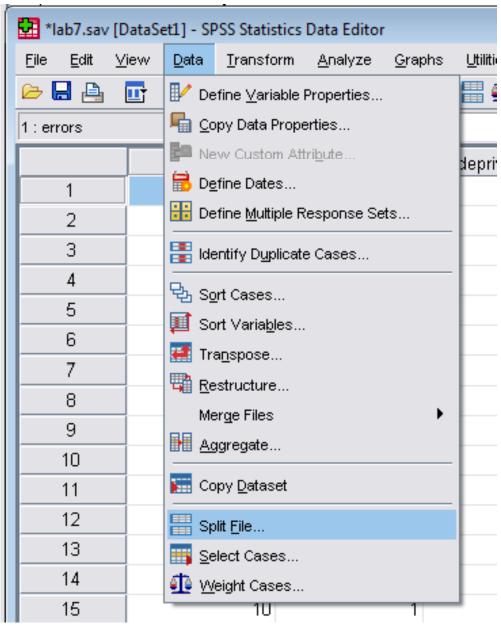
Tests of Between-Subjects Effects

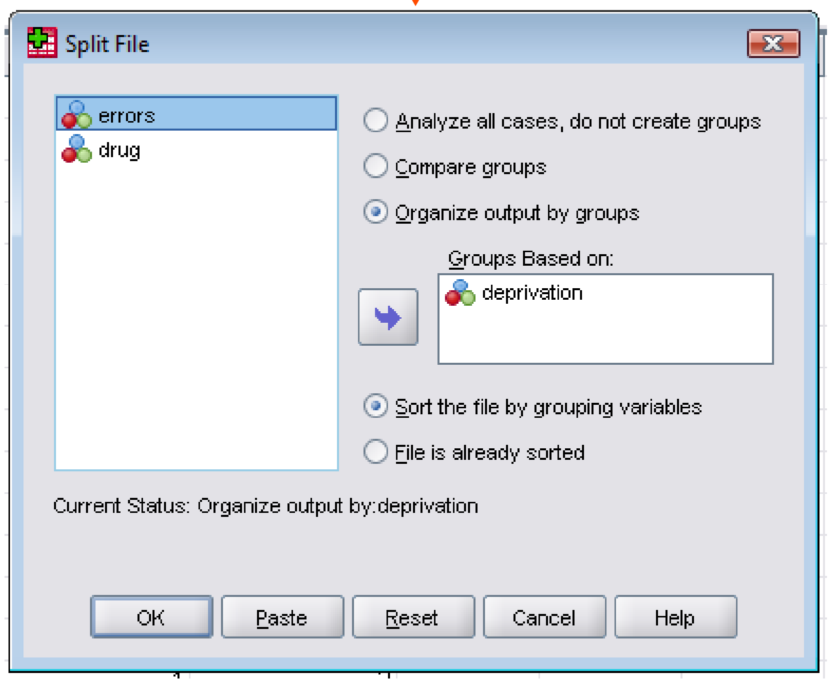
Dependent Variable:errors

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	280.000[a]	5	56.000	3.055	.036	.459
Intercept	2400.000	1	2400.000	130.909	.000	.879
drug	112.000	2	56.000	3.055	.072	.253
deprivation	24.000	1	24.000	1.309	.268	.068
drug * deprivation	144.000	2	72.000	3.927	.038	.304
Error	330.000	18	18.333			
Total	3010.000	24				
Corrected Total	610.000	23				

a. R Squared = .459 (Adjusted R Squared = .309)

Step 2…Separate One-Way ANOVAs





More Step 2…

deprivation = 1-hour deprivation

Tests of Between-Subjects Effects[b]

Dependent Variable:errors

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	248.000[a]	2	124.000	7.750	.011
Intercept	972.000	1	972.000	60.750	.000
drug	248.000	2	124.000	7.750	.011
Error	144.000	9	16.000		
Total	1364.000	12			
Corrected Total	392.000	11			

a. R Squared = .633 (Adjusted R Squared = .551)

b. deprivation = 1-hour deprivation

A at b_1

NOT THE RIGHT P-VALUES! (WRONG ERROR TERMS)

deprivation = 24-hour deprivation

Tests of Between-Subjects Effects[b]

Dependent Variable:errors

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	8.000[a]	2	4.000	.194	.827
Intercept	1452.000	1	1452.000	70.258	.000
drug	8.000	2	4.000	.194	.827
Error	186.000	9	20.667		
Total	1646.000	12			
Corrected Total	194.000	11			

a. R Squared = .041 (Adjusted R Squared = -.172)

b. deprivation = 24-hour deprivation

A at b_2

Step 3…Put Things Together

From Step 3

	Source	SS	df	MS	F	P-value
From Step 1	A	112	2	56	3.06	.072
	B	24	1	24	1.31	.268
	AxB	144	2	72	3.93	.038
From Step 2	A at b_1	248	2	124	6.76	.006
	A at b_2	8	2	4	0.22	.805
From Step 1	S/AB	330	18	18.333		
	Total	610	23			

Interpretation

- We can tell two things from our analysis:

  1. At condition b_1 (1-hour deprivation), there is a significant effect of drug on the number of errors

  2. At condition b_2 (24-hour deprivation), there is no significant difference between drug types on number of errors

- Now, we do have a significant effect…but between what drugs?
  - How did we delve further into effects previously?
  - Contrasts?

Contrasts for Simple Effects

SIMPLE COMPARISONS

Simple Comparisons

- As we did in all other ANOVA contrasts, we can form a contrast using the treatment means at a given level of a factor:

$$\hat{\psi}_{A \text{ at } b_k} = \sum_{j=1}^{a} c_j \bar{Y}_{jk}$$

- To demonstrate, let's consider the contrast where the two drugs were compared with the control, simultaneously, but only for the 1-hour deprivation condition:

	Control	Drug X	Drug Y
c_j	1	-.5	-.5

Building Our Contrast

- Our contrast coefficients:

	Control	Drug X	Drug Y
c_j	1	-.5	-.5

- Our means at b_1 (one hour deprivation)

	Control	Drug X	Drug Y
$\bar{Y}_{j1}$	3	10	14

- Our estimated contrast:

$$\psi = 1*3 + -.5*10 + -.5*14 = 3-5-7 = -9$$

- As in one-way ANOVA, we now need to form the SS for the contrast
  - The error term is $MS_{S/AB}$ from the two-way ANOVA table

Computational Formulas

- In general, the contrast sum of squares is:

$$SS_{\psi_{A \text{ at } b_k}} = \frac{n\hat{\psi}^2_{A \text{ at } b_k}}{\sum_{j=1}^{a} c_j^2}$$

- For our example, this is:

$$SS_{\psi_{A \text{ at } b_k}} = \frac{4 * -9^2}{1^2 + -.5^2 + -.5^2} = \frac{324}{1.5} = 216$$

Building the F-test

- We also need the $MS_{\psi_{A \text{ at } b_k}}$, which is SS divided by df
  - What are the df for a contrast again?

- So, $MS_{\psi_{A \text{ at } b_k}} = 216$

- The denominator of the F-ratio is $MS_{S/AB}$ from the two-way table $MS_{S/AB} = 18.333$

- So, the F-ratio is: $F_{\psi_{A \text{ at } b_k}} = \dfrac{216}{18.333} = 11.782$
- The p-value is: 0.003
  - From Excel (=fdist(11.782,1,18))

Interpretation…

- The interpretation of our contrast is the same as we would interpret a contrast in a one-way ANOVA
  - Just conditional on level b_1 (one-hour deprivation)

- Therefore, we conclude that the drug conditions have significantly more errors than the control condition when monkeys have a one-hour deprivation of food

EFFECT SIZES AND POWER FOR SIMPLE EFFECTS

Effect Sizes and Power for Simple Effects

- In most studies that use a factorial design, effect sizes are reported for the overall main effects and interaction (see Sections 11.6 and 11.7)

- That being said, you can always compute the partial eta-squared:

$$\hat{\eta}_{\langle effect \rangle} = \frac{SS_{Effect}}{SS_{Effect} + SS_{Error}}$$

- Here SS_{Effect} is the SS for the simple effect and SS_{Error} is the $SS_{S/AB}$

Effect Sizes for Simple Effects

Source	SS	df	MS	F	P-value
A	112	2	56	3.06	.072
B	24	1	24	1.31	.268
AxB	144	2	72	3.93	.038
A at b_1	248	2	124	6.76	.006
A at b_2	8	2	4	0.22	.805
S/AB	330	18	18.333		
Total	610	23			

$$\hat{\eta}_{\langle A\,\text{at}\,b_1\rangle} = \frac{248}{248+330} = 0.429$$

$$\hat{\eta}_{\langle A\,\text{at}\,b_2\rangle} = \frac{8}{8+330} = 0.024$$

Effect Sizes and Power

- If you wanted to compute power (or get sample size calculations), everything follows from one-way ANOVA

- Just use the partial eta-squared from the previous slide

CONTROLLING FAMILYWISE TYPE I ERROR

Controlling Familywise Type I Error

- There is a general consensus that the three principal effects (i.e., the two main effects and the interaction) are planned tests and do not require error correction
  - They are evaluated at a conventional significance level, such as $\alpha = .05$

- With a non-significant interaction, attention is usually drawn to one or both of the two main effects
  - Each factor should have a familywise error equal to the level of the original tests (e.g., $\alpha_{FW} = .05$)
    - Need for post-hoc adjustments

Main-Effect Comparisons

- A comparison that is a central planned portion of the study is evaluated without error control
  - For a few contrasts, the familywise error rate can be controlled with the Bonferroni method by taking $\alpha = \alpha_{FW}/c$ (or the Sidák-Bonferroni correction)
  - The set of all pairwise differences between means is most easily tested with the Tukey or Scheffé procedures
    - The Scheffé procedure is most conservative

Simple Effects

- There is no consensual standard for the level of familywise error to use with the simple effects.
  - The Bonferroni (or Sidák-Bonferroni) procedure is the most practical approach here

- The formulas from Chapter 6 apply directly, except that the error term $MS_{S/AB}$ from the factorial design is used instead of $MS_{S/A}$

Wrapping Up

- Today's class delved into what to do following the overall two-way analysis
- Main and simple effects were the focus of the discussion
- Most everything discussed today came in the presence of the possible interaction between independent variables
- The nature of the interaction dictates the level to which you describe the main effects.

Up Next...

- Lab Tonight:
  - How to do everything in SPSS
- Homework
  - Posted tomorrow – due next week before class
- For next week
  - Read chapters 13 and 14