Chapter 7
Factorial ANOVA: Two-way ANOVA

Factorial ANOVA
Two-factor ANOVA: Equal *n*

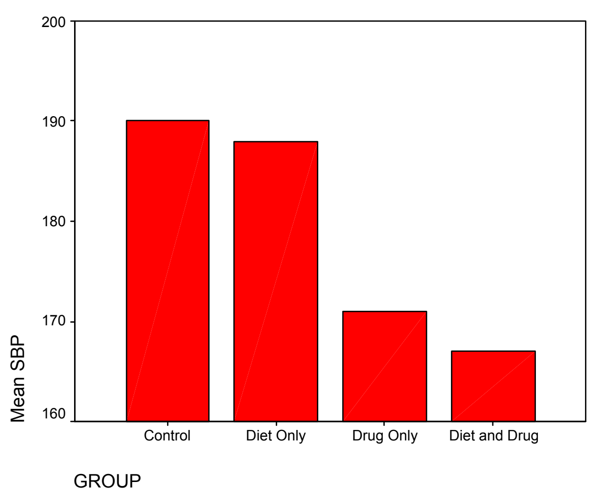
1. Examples of two-factor ANOVA designs

- Example #1: The effect of drugs and diet on systolic blood pressure
  20 individuals with high blood pressure were randomly assigned to one of four treatment conditions
  - Control group (Neither drug nor diet modification)
  - Diet modification only
  - Drug only
  - Both drug and diet modification

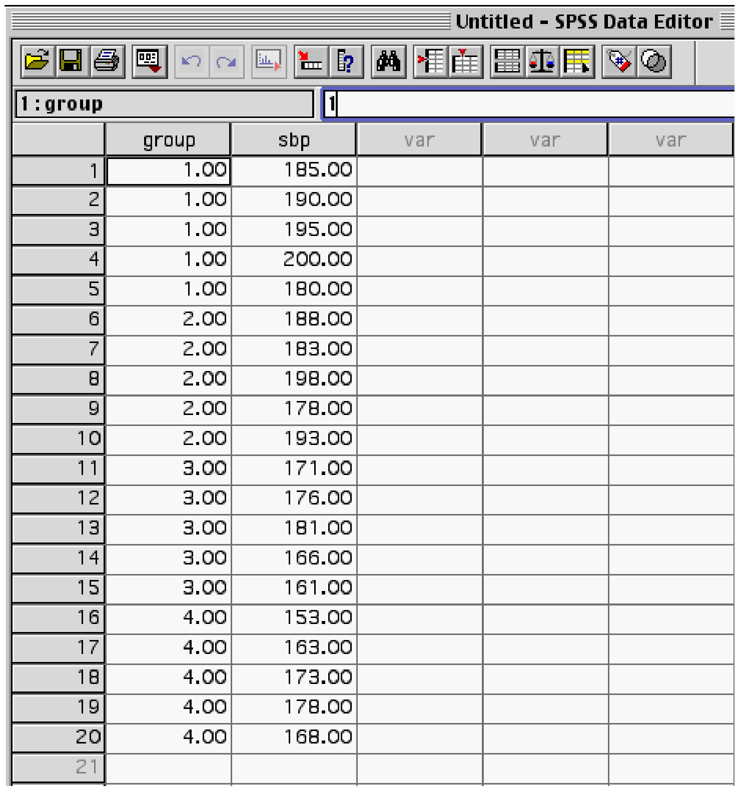
  At the end of the treatment period, SBP was assessed:

	Group			
	Control	Diet Only	Drug Only	Diet and Drug
	185	188	171	153
	190	183	176	163
	195	198	181	173
	200	178	166	178
	180	193	161	168
Mean	190	188	171	167

  - In the past, we would have analyzed these data as a one-way design



- In SPSS, our data file would have one IV with four levels:



Untitled - SPSS Data Editor

1 : group | 1

	group	sbp	var	var	var
1	1.00	185.00			
2	1.00	190.00			
3	1.00	195.00			
4	1.00	200.00			
5	1.00	180.00			
6	2.00	188.00			
7	2.00	183.00			
8	2.00	198.00			
9	2.00	178.00			
10	2.00	193.00			
11	3.00	171.00			
12	3.00	176.00			
13	3.00	181.00			
14	3.00	166.00			
15	3.00	161.00			
16	4.00	153.00			
17	4.00	163.00			
18	4.00	173.00			
19	4.00	178.00			
20	4.00	168.00			
21					

ONEWAY iv BY group.

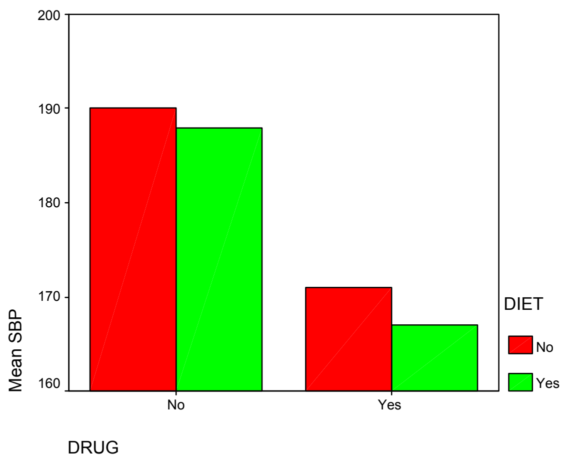
ANOVA

SBP

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2050.000	3	683.333	9.762	.001
Within Groups	1120.000	16	70.000		
Total	3170.000	19			

- Alternatively, we could also set up our data as a two-factor ANOVA

		Diet Modification: No	Diet Modification: Yes	
Drug Therapy	No	$\overline{X}_{\cdot 11} = 190$	$\overline{X}_{\cdot 21} = 188$	$\overline{X}_{\cdot\cdot 1} = 189$
	Yes	$\overline{X}_{\cdot 12} = 171$	$\overline{X}_{\cdot 22} = 167$	$\overline{X}_{\cdot\cdot 2} = 169$
		$\overline{X}_{\cdot 1 \cdot} = 180.5$	$\overline{X}_{\cdot 2 \cdot} = 177.5$	$\overline{X}_{\cdots} = 179$



- In SPSS, our data file would have two IVs each with two levels:

Untitled - SPSS Data Editor

5 :

	iv1	iv2	sbp	var	var
1	.00	.00	185.00		
2	.00	.00	190.00		
3	.00	.00	195.00		
4	.00	.00	200.00		
5	.00	.00	180.00		
6	.00	1.00	188.00		
7	.00	1.00	183.00		
8	.00	1.00	198.00		
9	.00	1.00	178.00		
10	.00	1.00	193.00		
11	1.00	.00	171.00		
12	1.00	.00	176.00		
13	1.00	.00	181.00		
14	1.00	.00	166.00		
15	1.00	.00	161.00		
16	1.00	1.00	153.00		
17	1.00	1.00	163.00		
18	1.00	1.00	173.00		
19	1.00	1.00	178.00		
20	1.00	1.00	168.00		
21					

UNIANOVA sbp BY IV1 IV2.

Tests of Between-Subjects Effects

Dependent Variable: SBP

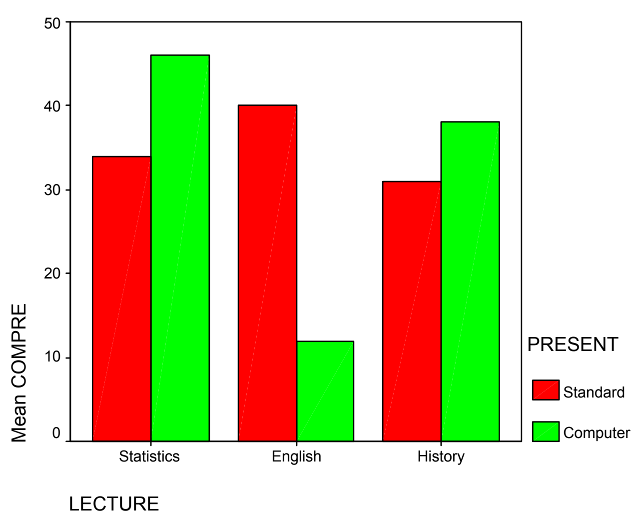
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2050.000[a]	3	683.333	9.762	.001
Intercept	640820.000	1	640820.000	9154.571	.000
DRUG	2000.000	1	2000.000	28.571	.000
DIET	45.000	1	45.000	.643	.434
DRUG * DIET	5.000	1	5.000	.071	.793
Error	1120.000	16	70.000		
Total	643990.000	20			
Corrected Total	3170.000	19			

a. R Squared = .647 (Adjusted R Squared = .580)

- Example #2: The relationship between type of lecture and method of presentation to lecture comprehension

  30 people were randomly assigned to one of six experimental conditions. At the end of the lecture, a measure of comprehension was obtained.

	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	44 18 48 32 35 27	47 37 42 42 39 33	46 21 40 30 29 20
Computer	53 42 49 51 47 34	13 10 16 11 16 6	45 36 41 35 38 33



2. Terminology and notation for a two-factor ANOVA

- Level = the different aspects/amounts of an independent variable
- Factor = an independent variable
  - A one factor ANOVA has one independent variable
  - A two factor ANOVA has two independent variables
  - An m factor ANOVA has m independent variables
- A factorial design = a design where all possible combinations of each independent variable are completely crossed
  - A factorial design with two factors is designated as a $a \text{X} b$ design

    a = the number of levels of the first factor

    b = the number of levels of the second factor

    The blood pressure example is a $2 \text{X} 2$ design

    Factor A (diet modification) has two levels

    Factor B (drug therapy) has two levels

    The lecture comprehension example is a $3 \text{X} 2$ design

    Factor A (type of lecture) has three levels

    Factor B (method of presentation) has two levels

  - This notation can be extended to denote multi-factor designs

    A factorial design with three factors is designated $\text{a X } b \text{ X } c$

    a = the number of levels of the first factor

    b = the number of levels of the second factor

    c = the number of levels of the third factor

    In this class we will not consider non-factorial (or partial factorial) designs

  - Consider an example where participants are randomly assigned to a type of lecture (history, statistics, psychology, or English), to be presented in either a large or small classroom, using different methods of presentation (blackboard, overhead projector, or computer), and given by a graduate student, an assistant professor or a full professor.

    How would you describe this design?

- An example:

Method of Presentation (Factor B)	Type of Lecture (Factor A) Statistics a_1	English a_2	History a_3
Standard b_1	x_{111}	x_{121}	x_{131}
	x_{211}	x_{221}	x_{231}
	x_{311}	x_{321}	x_{331}
	x_{411}	x_{421}	x_{431}
	x_{511}	x_{521}	x_{531}
Computer b_2	x_{112}	x_{122}	x_{132}
	x_{212}	x_{222}	x_{232}
	x_{312}	x_{322}	x_{332}
	x_{412}	x_{422}	x_{432}
	x_{512}	x_{522}	x_{532}

x_{ijk} i = indicator for subject within level jk
j = indicator for level of factor A
k = indicator for level of factor B

Method of Presentation (Factor B)	Type of Lecture (Factor A) Statistics a_1	English a_2	History a_3	
Standard b_1	$\bar{X}_{\cdot 11}$	$\bar{X}_{\cdot 21}$	$\bar{X}_{\cdot 31}$	$\bar{X}_{\cdot\cdot 1}$
Computer b_2	$\bar{X}_{\cdot 12}$	$\bar{X}_{\cdot 22}$	$\bar{X}_{\cdot 32}$	$\bar{X}_{\cdot\cdot 2}$
	$\bar{X}_{\cdot 1\cdot}$	$\bar{X}_{\cdot 2\cdot}$	$\bar{X}_{\cdot 3\cdot}$	$\bar{X}_{\cdot\cdot\cdot}$

- Kinds of effects in a two-factor design
  - Main effects
  - Interaction effects

- A main effect of a factor is the effect of that factor *averaging across all the levels of all the other factors*
  - The main effect of factor A examines if there are any differences in the DV as a function of the levels of factor A, *averaging across the levels of all other IVs*. These means are called the marginal means for factor A

$$H_0 : \mu_{\cdot 1 \cdot} = \mu_{\cdot 2 \cdot} = \ldots = \mu_{\cdot a \cdot}$$
$$H_1 : Not\ all\ \mu_{\cdot j \cdot}'s\ are\ equal$$

  - The main effect of factor B examines if there are any differences in the DV as a function of the levels of factor B, *averaging across the levels of all other IVs*. These means are called the marginal means for factor B

$$H_0 : \mu_{\cdot\cdot 1} = \mu_{\cdot\cdot 2} = \ldots = \mu_{\cdot\cdot b}$$
$$H_1 : Not\ all\ \mu_{\cdot\cdot k}'s\ are\ equal$$

  - Note that the main effect of a factor is not (necessarily) equal to the effect of that factor in the absence of all other factors

  - When a factor has more than two levels, then the test for a main effect is an omnibus test, and follow-up tests are required to identify the effect

  - For the SBP example

Marginal Means for Drug Therapy

		Diet Modification: No	Diet Modification: Yes	
Drug Therapy	No	$\bar{X}_{\cdot 11} = 190$	$\bar{X}_{\cdot 21} = 188$	$\bar{X}_{\cdot\cdot 1} = 189$
	Yes	$\bar{X}_{\cdot 12} = 171$	$\bar{X}_{\cdot 22} = 167$	$\bar{X}_{\cdot\cdot 2} = 169$
		$\bar{X}_{\cdot 1 \cdot} = 180.5$	$\bar{X}_{\cdot 2 \cdot} = 177.5$	$\bar{X}_{\cdots} = 179$

Marginal Means for Diet Modification

- To test the main effect of diet modification, we examine

$$\hat{\mu}_{\cdot 1 \cdot} = \bar{X}_{\cdot 1 \cdot} = 180.5 \quad \text{and} \quad \hat{\mu}_{\cdot 2 \cdot} = \bar{X}_{\cdot 2 \cdot} = 177.5$$

- To test the main effect of drug therapy, we examine

$$\hat{\mu}_{\cdot\cdot 1} = \bar{X}_{\cdot\cdot 1} = 189 \quad \text{and} \quad \hat{\mu}_{\cdot\cdot 2} = \bar{X}_{\cdot\cdot 2} = 169$$

- An interaction of two factors
  - The interaction of A and B examines:
    - if the effect of one variable depends on the level of the other variable
    - if the main effect of factor A is the same for all levels of factor B
    - if the main effect of factor B is the same for all levels of factor A
  - Indicates non-additivity of effects

  - To investigate the interaction of A and B, we examine the cell means
  - For the SBP example

		Diet Modification		
		No	Yes	
Drug Therapy	No	$\bar{X}_{\cdot 11} = 190$	$\bar{X}_{\cdot 21} = 188$	$\bar{X}_{\cdot\cdot 1} = 189$
	Yes	$\bar{X}_{\cdot 12} = 171$	$\bar{X}_{\cdot 22} = 167$	$\bar{X}_{\cdot\cdot 2} = 169$
		$\bar{X}_{\cdot 1 \cdot} = 180.5$	$\bar{X}_{\cdot 2 \cdot} = 177.5$	$\bar{X}_{\cdots} = 179$

- The effect of diet modification (Factor A) among those in the no drug therapy condition (level 1 of Factor B):

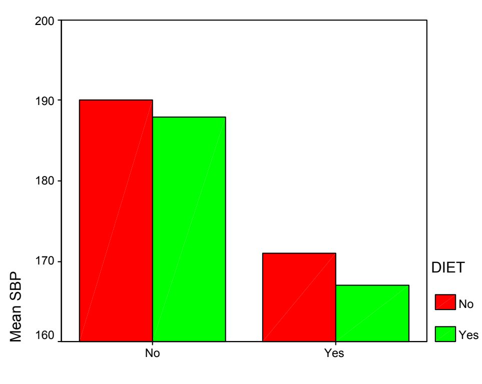
$$\hat{\mu}_{\cdot 11} = \bar{X}_{\cdot 11} = 190 \quad \text{and} \quad \hat{\mu}_{\cdot 21} = \bar{X}_{\cdot 21} = 188$$
$$\hat{\mu}_{\cdot 11} - \hat{\mu}_{\cdot 21} = \bar{X}_{\cdot 11} - \bar{X}_{\cdot 21} = 2$$

- The effect of diet modification (Factor A) among those in the drug therapy condition (level 2 of Factor B):

$$\hat{\mu}_{\cdot 12} = \bar{X}_{\cdot 12} = 171 \quad \text{and} \quad \hat{\mu}_{\cdot 22} = \bar{X}_{\cdot 22} = 167$$
$$\hat{\mu}_{\cdot 12} - \hat{\mu}_{\cdot 22} = \bar{X}_{\cdot 12} - \bar{X}_{\cdot 22} = 4$$

- If there is no interaction, then the effect of diet modification will be the same at each level of drug therapy

  (The ‘difference of differences’ will be zero)



- An exactly equivalent test is to look at the effect of drug therapy (Factor B) within each level of factor A

  - The effect of drug therapy (Factor B) among those in the no diet modification condition (level 1 of Factor A):

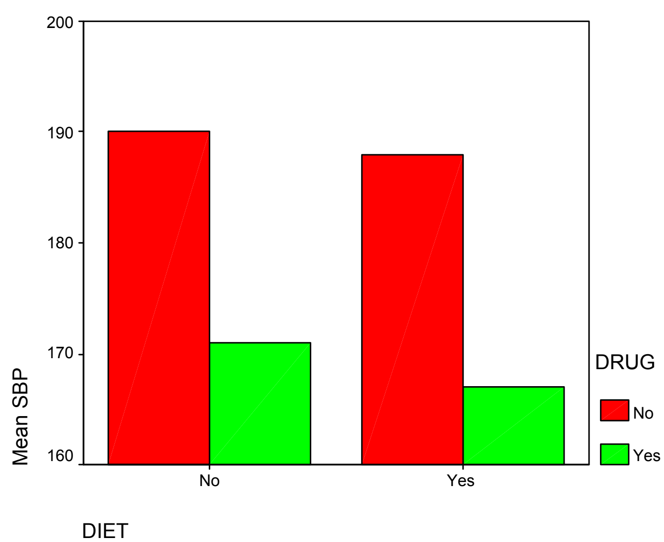
    $\hat{\mu}_{\cdot 11} = \overline{X}_{\cdot 11} = 190$ and $\hat{\mu}_{\cdot 12} = \overline{X}_{\cdot 12} = 171$

    $$\hat{\mu}_{\cdot 11} - \hat{\mu}_{\cdot 12} = \overline{X}_{\cdot 11} - \overline{X}_{\cdot 12} = 19$$

  - The effect of drug therapy (Factor B) among those in the diet modification condition (level 2 of Factor B):

    $\hat{\mu}_{\cdot 21} = \overline{X}_{\cdot 21} = 188$ and $\hat{\mu}_{\cdot 22} = \overline{X}_{\cdot 22} = 167$

    $$\hat{\mu}_{\cdot 21} - \hat{\mu}_{\cdot 22} = \overline{X}_{\cdot 21} - \overline{X}_{\cdot 22} = 21$$



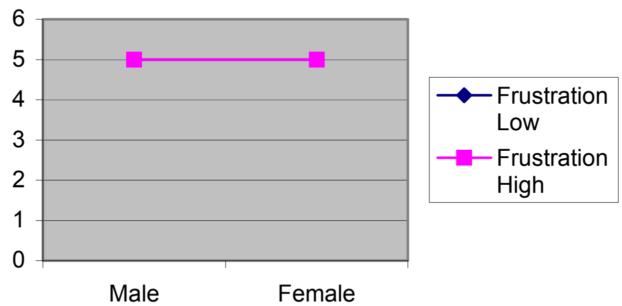
- The main advantage of conducting multi-factor ANOVA designs is the ability to detect and test interactions.
- It may also give you greater generalizability of your results
- Including additional factors may reduce the error term (MSW) which will lead to increased power

3. Understanding main effects and interactions
   - The easiest way to understand main effects and interactions is by graphing cell means.
   - Non-parallel lines indicate the presence of an interaction
     (Non-additivity of effects)

   - Let's consider a 2 * 2 design where male and female participants experience either low or high levels of frustration

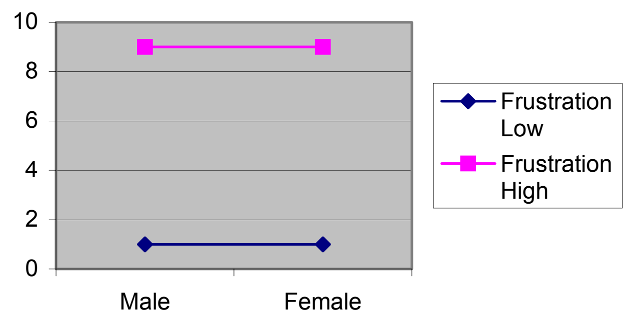
     - Case 1: No main effects and no interactions

	Frustration Low	Frustration High	
Male	5	5	5
Female	5	5	5
	5	5	



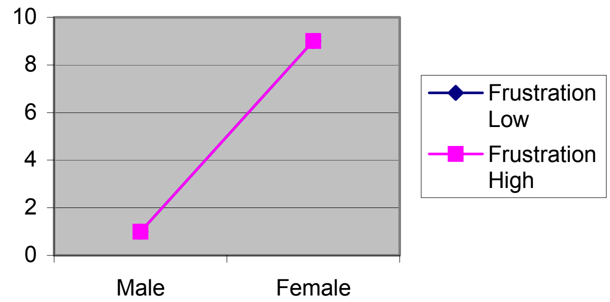
     - Case 2: Main effect for frustration, no main effect for gender, no interaction

	Frustration Low	Frustration High	
Male	1	9	5
Female	1	9	5
	1	9	



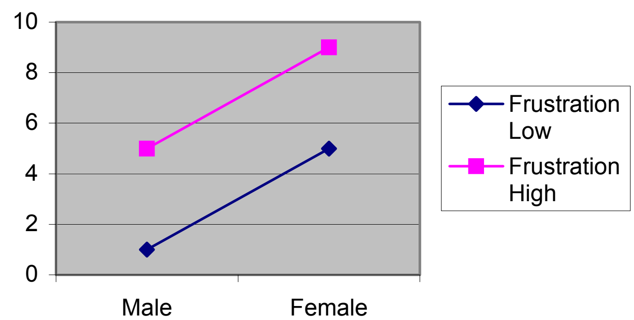
- Case 3: No main effect for frustration, main effect for gender, no interaction

	Frustration		
	Low	High	
Male	1	1	1
Female	9	9	9
	5	5	



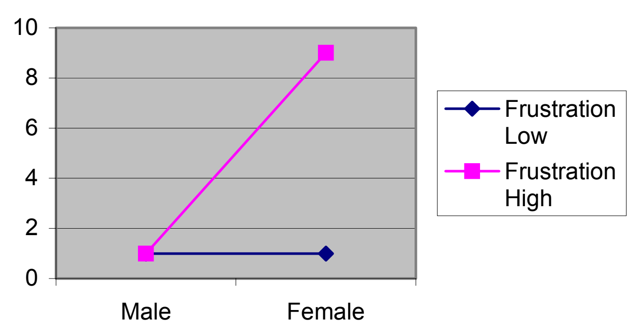
- Case 4: Main effect for frustration, main effect for gender, no interaction

	Frustration		
	Low	High	
Male	1	5	3
Female	5	9	7
	3	7	



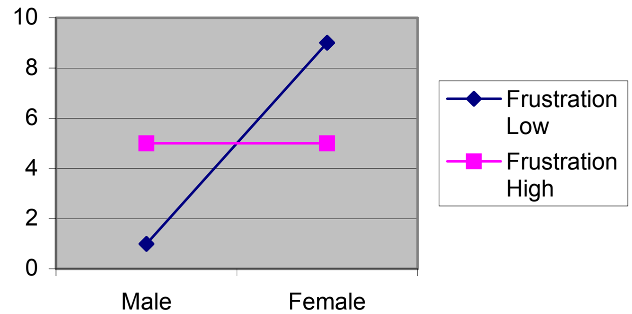
- Case 5: Main effect for frustration, main effect for gender, frustration by gender interaction

	Frustration		
	Low	High	
Male	1	1	1
Female	1	9	5
	1	5	



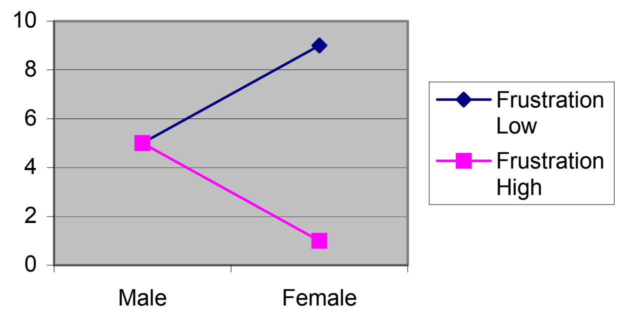
- Case 6: No main effect for frustration, main effect for gender, frustration by gender interaction

	Frustration		
	Low	High	
Male	1	5	3
Female	9	5	7
	5	5	



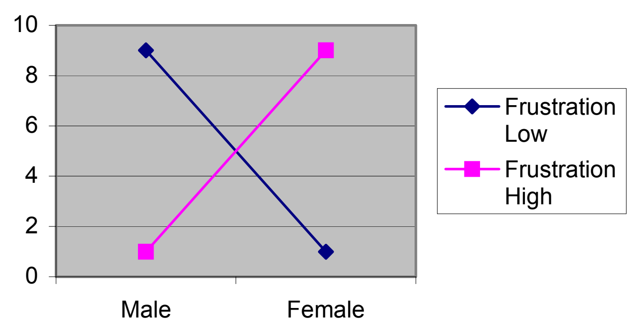
- Case 7: Main effect for frustration, no main effect for gender, frustration by gender interaction

	Frustration		
	Low	High	
Male	5	5	5
Female	9	1	5
	7	3	



- Case 8: No main effect for frustration, no main effect for gender, frustration by gender interaction

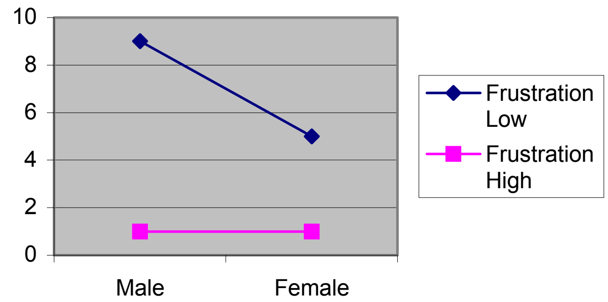
	Frustration		
	Low	High	
Male	9	1	5
Female	1	9	5
	5	5	



- Note when an interaction is present, it can be misleading and erroneous to interpret a main effect (see Case 7)
- If an interaction is present, only true main effects should be interpreted

  - Case 9: A true main effect for frustration and a frustration by gender interaction

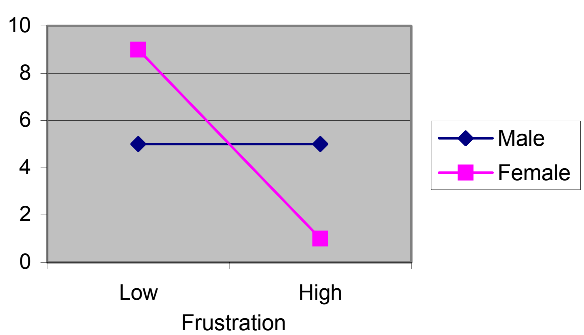
	Frustration		
	Low	High	
Male	9	1	5
Female	5	1	3
	7	1	

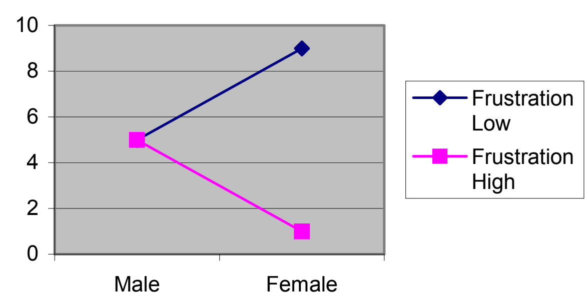


- There are two ways to display/interpret any interactions

  - Case 7 (revisited): Main effect for frustration, no main effect for gender, frustration by gender interaction

	Frustration		
	Low	High	
Male	5	5	5
Female	9	1	5
	7	3	



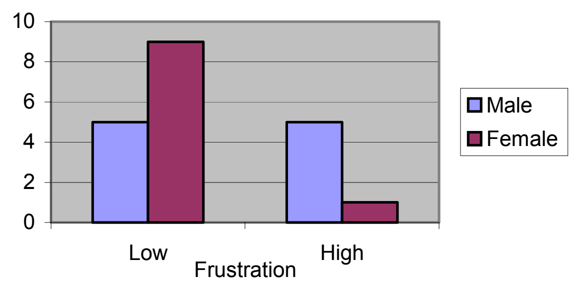


- Graph A: Tend to interpret/read as gender across levels of frustration
- Graph B: Tend to interpret/read as level of frustration across genders

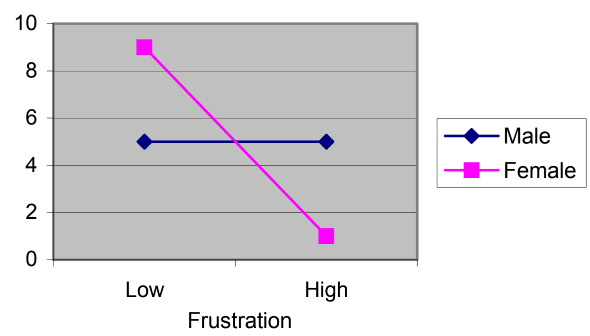
- An aside on graphing interactions
  - For between-subjects factors, it is best to use bar graphs
    (to indicate that each bar is a separate group of people)

  - For within-subjects or repeated measures factors, use line graphs to connect the data points at each level of measurement
    (line graphs have been presented for pedagogical purposes only)

Good



Less Good



4. The structural model for two-way ANOVA

- The purpose of the structural model is to decompose each score into a part we can explain (MODEL) and a part we can not explain (ERROR)

- For a one-way ANOVA design, the model had only two components:

$$Y_{ij} = MODEL + ERROR$$
$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

μ	The overall mean of the scores
α_j	The effect of being in level j
ε_{ij}	The unexplained part of the score

- In a two-way ANOVA design, our model will be more refined, and we will have additional components to the model:

$$
\begin{aligned}
Y_{ijk} &= MODEL + ERROR \\
Y_{ijk} &= \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk}
\end{aligned}
$$

μ The overall mean of the scores

α_j The effect of being in level *j* of Factor A

β_k The effect of being in level *k* of Factor B

$(\alpha\beta)_{jk}$ The effect of being in level *j* of Factor A and level k of Factor B (the interaction of level *j* of Factor A and level *k* of Factor B)

ε_{ijk} The unexplained part of the score

α_j The effect of being in level *j* of Factor A

$$\alpha_j = \mu._j. - \mu...$$

$$\sum_{j=1}^{a} \alpha_j = 0$$

β_k The effect of being in level *k* of Factor B

$$\beta_k = \mu.._k - \mu...$$

$$\sum_{k=1}^{b} \beta_k = 0$$

$(\alpha\beta)_{jk}$ The effect of being in level *j* of Factor A and level *k* of Factor B (the interaction of level *j* of Factor A and level *k* of Factor B)

$$(\alpha\beta)_{jk} = \mu._{jk} - \mu._j. - \mu.._k + \mu...$$

$$\sum_{j=1}^{a} (\alpha\beta)_{jk} = 0 \quad \text{for each level of } j$$

$$\sum_{k=1}^{b} (\alpha\beta)_{jk} = 0 \quad \text{for each level of } k$$

ε_{ijk} The unexplained part of the score

$$
\begin{aligned}
\varepsilon_{ijk} &= Y_{ijk} - MODEL \\
&= Y_{ijk} - \left(\mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}\right)
\end{aligned}
$$

- Blood Pressure Example:

Entries indicate cell means based on *n*=5

		Diet Modification: No	Diet Modification: Yes	
Drug Therapy	No	$\overline{X}_{\cdot 11} = 190$	$\overline{X}_{\cdot 21} = 188$	$\overline{X}_{\cdot\cdot 1} = 189$
	Yes	$\overline{X}_{\cdot 12} = 171$	$\overline{X}_{\cdot 22} = 167$	$\overline{X}_{\cdot\cdot 2} = 169$
		$\overline{X}_{\cdot 1\cdot} = 180.5$	$\overline{X}_{\cdot 2\cdot} = 177.5$	$\overline{X}_{\cdots} = 179$

μ The overall mean of the scores

$$\hat{\mu} = 179$$

α_j The effect of being in level *j* of Diet Modification

$$\alpha_j = \mu_{\cdot j \cdot} - \mu_{\cdots}$$

α_1 is the effect of being in the No Diet Modification condition

$$\hat{\alpha}_1 = 180.5 - 179 = 1.5$$

α_2 is the effect of being in the Diet Modification condition

$$\hat{\alpha}_2 = 177.5 - 179 = -1.5$$

Note that $1.5 + (-1.5) = 0$

The test for the main effect of Diet Modification:

$H_0 : \mu_{\cdot 1 \cdot} = \mu_{\cdot 2 \cdot}$ or $H_0 : \alpha_1 = \alpha_2 = 0$

β_k The effect of being in level *k* of Drug Therapy

$$\beta_k = \mu_{\cdot\cdot k} - \mu_{\cdots}$$

β_1 is the effect of being in the No Drug Therapy condition

$$\hat{\beta}_1 = 189 - 179 = 10$$

β_2 is the effect of being in the Drug Therapy condition

$$\hat{\beta}_2 = 169 - 179 = -10$$

Note that $10 + (-10) = 0$

The test for the main effect of Drug Therapy:

$H_0 : \mu_{\cdot\cdot 1} = \mu_{\cdot\cdot 2}$ or $H_0 : \beta_1 = \beta_2 = 0$

$(\alpha\beta)_{jk}$ The effect of being in level *j* of Factor A and level *k* of Factor B (the interaction of level *j* of Factor A and level *k* of Factor B)

$$(\alpha\beta)_{jk} = \mu._{jk} - \mu._{j}. - \mu.._{k} + \mu...$$

$(\alpha\beta)_{11}$ is the effect of being in the No Diet Modification and in the No Drug Therapy conditions

$$(\hat{\alpha}\beta)_{11} = 190 - 180.5 - 189 + 179 = -0.5$$

$(\alpha\beta)_{12}$ is the effect of being in the No Diet Modification and in the Drug Therapy conditions

$$(\hat{\alpha}\beta)_{12} = 171 - 180.5 - 169 + 179 = 0.5$$

$(\alpha\beta)_{21}$ is the effect of being in the Diet Modification and in the No Drug Therapy conditions

$$(\hat{\alpha}\beta)_{21} = 188 - 177.5 - 189 + 179 = 0.5$$

$(\alpha\beta)_{22}$ is the effect of being in the Diet Modification and in the Drug Therapy conditions

$$(\hat{\alpha}\beta)_{22} = 167 - 177.5 - 169 + 179 = -0.5$$

Note that adding across the Diet Modification factor:

For No Drug Therapy: $-0.5 + 0.5 = 0$

For Drug Therapy: $0.5 + (-0.5) = 0$

Note that adding across the Drug Therapy factor:

For No Diet Modification: $-0.5 + 0.5 = 0$

For Diet Modification: $0.5 - 0.5 = 0$

The test for the interaction of Diet Modification and Drug Therapy:

$$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = (\alpha\beta)_{21} = (\alpha\beta)_{22} = 0$$

- Lecture Comprehension:

Entries indicate cell means based on n=6

	Type of Lecture			
Method of Presentation	Statistics	English	History	
Standard	$\bar{X}_{\cdot 11} = 34.0$	$\bar{X}_{\cdot 21} = 40.0$	$\bar{X}_{\cdot 31} = 31.0$	$\bar{X}_{\cdot\cdot 1} = 35.0$
Computer	$\bar{X}_{\cdot 12} = 46.0$	$\bar{X}_{\cdot 22} = 12.0$	$\bar{X}_{\cdot 32} = 38.0$	$\bar{X}_{\cdot\cdot 2} = 32.0$
	$\bar{X}_{\cdot 1\cdot} = 40.0$	$\bar{X}_{\cdot 2\cdot} = 26.0$	$\bar{X}_{\cdot 3\cdot} = 34.5$	$\bar{X}_{\cdots} = 33.5$

μ The overall mean of the scores

$$\hat{\mu} = 33.5$$

α_j The effect of being in level j of Type of Lecture

$$\alpha_j = \mu_{\cdot j \cdot} - \mu_{\cdots}$$

α_1 is the effect of being in the Statistics Lecture

$$\hat{\alpha}_1 = 40 - 33.5 = 6.5$$

α_2 is the effect of being in the English Lecture

$$\hat{\alpha}_2 = 26 - 33.5 = -7.5$$

α_3 is the effect of being in the History Lecture

$$\hat{\alpha}_3 = 34.5 - 33.5 = 1.0$$

Note that $6.5 - 7.5 + 1.0 = 0$

β_k The effect of being in level k of Method of Presentation

$$\beta_k = \mu_{\cdot\cdot k} - \mu_{\cdots}$$

β_1 is the effect of being in the Standard Presentation condition

$$\hat{\beta}_1 = 35 - 33.5 = 1.5$$

β_2 is the effect of being in the Computer Presentation condition

$$\hat{\beta}_2 = 32 - 33.5 = -1.5$$

Note that $1.5 - 1.5 = 0$

$(\alpha\beta)_{jk}$ The effect of being in level *j* of Factor A and level *k* of Factor B
(the interaction of level *j* of Factor A and level *k* of Factor B)

$$(\alpha\beta)_{jk} = \mu._{jk} - \mu._{j}. - \mu.._{k} + \mu...$$

$(\alpha\beta)_{11}$ is the effect of being in the Statistics lecture and in the Standard Presentation conditions

$(\hat{\alpha\beta})_{11} = 34 - 40 - 35 + 33.5 = -7.5$

$(\alpha\beta)_{12}$ is the effect of being in the Statistics lecture and in the Computer Presentation conditions

$(\hat{\alpha\beta})_{12} = 46 - 40 - 32 + 33.5 = 7.5$

$(\alpha\beta)_{21}$ is the effect of being in the English lecture and in the Standard Presentation conditions

$(\hat{\alpha\beta})_{21} = 40 - 26 - 35 + 33.5 = 12.5$

$(\alpha\beta)_{22}$ is the effect of being in the English lecture and in the Computer Presentation conditions

$(\hat{\alpha\beta})_{22} = 12 - 26 - 32 + 33.5 = -12.5$

$(\alpha\beta)_{31}$ is the effect of being in the History lecture and in the Standard Presentation conditions

$(\hat{\alpha\beta})_{21} = 31 - 34.5 - 35 + 33.5 = -5.0$

$(\alpha\beta)_{32}$ is the effect of being in the History lecture and in the Computer Presentation conditions

$(\hat{\alpha\beta})_{22} = 38 - 34.5 - 32 + 33.5 = 5.0$

Note that adding across the Type of Lecture:

Standard Presentation: $-7.5 + 12.5 - 5.0 = 0$
Computer Presentation: $7.5 - 12.5 + 5.0 = 0$

Note that adding across the Method of Presentation:

Statistics Lecture: $-7.5 + 7.5 = 0$
English Lecture: $12.5 - 12.5 = 0$
History Lecture: $-5.0 + 5.0 = 0$

5. Variance partitioning for two-way ANOVA

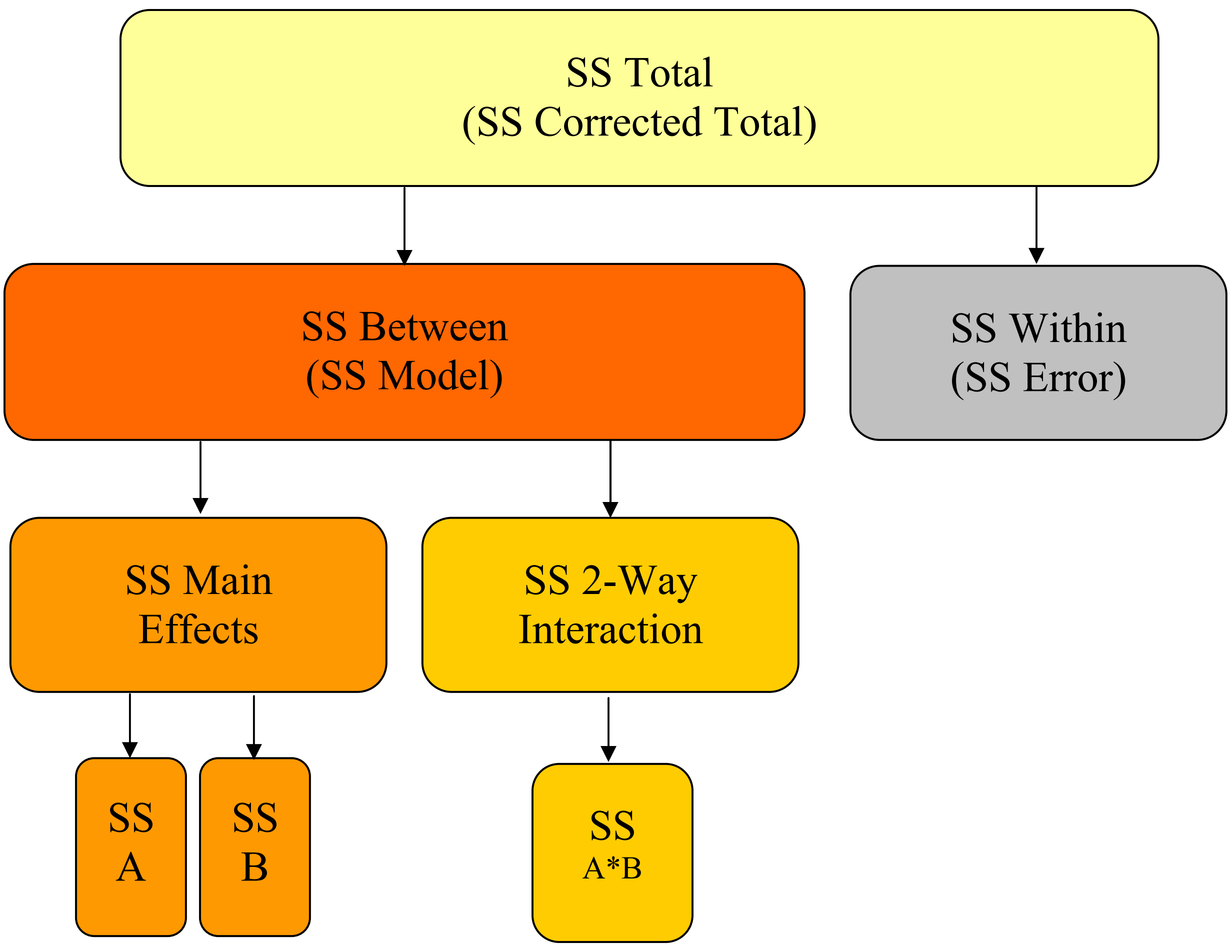
- Recall that for a one-way ANOVA we partitioned the sums of squares total into sum of squares between and sum of squares within

$$\sum_{j}^{a}\sum_{i}^{n}(y_{ij}-\bar{y}_{..})^2 = n\sum_{j}^{a}(\bar{y}_{.j}-\bar{y}_{..})^2+\sum_{j}^{a}\sum_{i}^{n}(y_{ij}-\bar{y}_{.j})^2$$

$$SST = SSBet + SSW$$

Where $SSBet$ is the SS of the model
SSW is the SS that we cannot explain (error)

- For a two-way ANOVA, our model has additional components, so we will be able to partition the SSB into several components



$$Y_{ijk} = MODEL + ERROR$$

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk}$$

$$Y_{ijk} = \left(\bar{Y}_{\cdots}\right) + \left(\bar{Y}_{\cdot j \cdot} - \bar{Y}_{\cdots}\right) + \left(\bar{Y}_{\cdot\cdot k} - \bar{Y}_{\cdots}\right) + \left(\bar{Y}_{\cdot jk} - \bar{Y}_{\cdot j \cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdots}\right) + \left(\bar{Y}_{ijk} - \bar{Y}_{\cdot jk}\right)$$

$$\left(Y_{ijk} - \bar{Y}_{\cdots}\right) = \left(\bar{Y}_{\cdot j \cdot} - \bar{Y}_{\cdots}\right) + \left(\bar{Y}_{\cdot\cdot k} - \bar{Y}_{\cdots}\right) + \left(\bar{Y}_{\cdot jk} - \bar{Y}_{\cdot j \cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdots}\right) + \left(\bar{Y}_{ijk} - \bar{Y}_{\cdot jk}\right)$$

Now if we square both sides of the equation, sum over all the observations, and simplify:

$$\begin{array}{ll}
\displaystyle\sum_{i=i}^{n}\sum_{j=1}^{a}\sum_{k=1}^{b}\left(Y_{ijk} - \bar{Y}_{\cdots}\right)^2 & \text{SS Total} \\
\left.\begin{array}{ll}
\displaystyle = nb\sum_{j=1}^{a}\left(\bar{Y}_{\cdot j \cdot} - \bar{Y}_{\cdots}\right)^2 & \text{SS Factor A} \\
\displaystyle + na\sum_{k=1}^{b}\left(\bar{Y}_{\cdot\cdot k} - \bar{Y}_{\cdots}\right)^2 & \text{SS Factor B} \\
\displaystyle + n\sum_{j=1}^{a}\sum_{k=1}^{b}\left(\bar{Y}_{\cdot jk} - \bar{Y}_{\cdot j \cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdots}\right)^2 & \text{SS AB Interaction}
\end{array}\right\} & \text{SS Between} \\
\displaystyle + \sum_{i=1}^{n}\sum_{j=1}^{a}\sum_{k=1}^{b}\left(\bar{Y}_{ijk} - \bar{Y}_{\cdot jk}\right)^2 & \text{SS Within cell (SS Error)}
\end{array}$$

- A simple computational example:

Data	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	44 18 48 32 35 27	47 37 42 42 39 33	46 21 40 30 29 20
Computer	53 42 49 51 47 34	13 10 16 11 16 6	45 36 41 35 38 33

Means (n=6)	Type of Lecture			
Method of Presentation	Statistics	English	History	
Standard	$\bar{X}_{\cdot 11} = 34$	$\bar{X}_{\cdot 21} = 40$	$\bar{X}_{\cdot 31} = 31$	$\bar{X}_{\cdot\cdot 1} = 35$
Computer	$\bar{X}_{\cdot 12} = 46$	$\bar{X}_{\cdot 22} = 12$	$\bar{X}_{\cdot 32} = 38$	$\bar{X}_{\cdot\cdot 2} = 32$
	$\bar{X}_{\cdot 1\cdot} = 40$	$\bar{X}_{\cdot 2\cdot} = 26$	$\bar{X}_{\cdot 3\cdot} = 34.5$	$\bar{X}_{\cdots} = 33.5$

SS Total

$$\sum_{i=i}^{n}\sum_{j=1}^{a}\sum_{k=1}^{b}\left(Y_{ijk} - \bar{Y}_{\dots}\right)^2$$

$$= \begin{array}{l} (44-33.5)^2 + (18-33.5)^2 + (48-33.5)^2 + (32-33.5)^2 + (35-33.5)^2 + (27-33.5)^2 + \\ (47-33.5)^2 + (37-33.5)^2 + (42-33.5)^2 + (42-33.5)^2 + (39-33.5)^2 + (33-33.5)^2 + \\ (46-33.5)^2 + (21-33.5)^2 + (40-33.5)^2 + (30-33.5)^2 + (29-33.5)^2 + (20-33.5)^2 + \\ (53-33.5)^2 + (42-33.5)^2 + (49-33.5)^2 + (51-33.5)^2 + (47-33.5)^2 + (34-33.5)^2 + \\ (13-33.5)^2 + (10-33.5)^2 + (16-33.5)^2 + (11-33.5)^2 + (16-33.5)^2 + (6-33.5)^2 + \\ (45-33.5)^2 + (36-33.5)^2 + (41-33.5)^2 + (35-33.5)^2 + (38-33.5)^2 + (33-33.5)^2 \end{array}$$

$= 5793$

$\hat{\alpha}_1 = 6.5 \qquad \hat{\alpha}_2 = -7.5 \qquad \hat{\alpha}_3 = 1.0$

SS Factor A

$$nb\sum_{j=1}^{a}\left(\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\dots}\right)^2 = 6*2\left[(40-33.5)^2 + (26-33.5)^2 + (34.5-33.5)^2\right]$$

$$= 12\left[6.5^2 + (-7.5)^2 + (1.0)^2\right]$$

$$= 12\left[42.25 + 56.25 + 1^2\right]$$

$$= 12[99.5]$$

$$= 1194$$

$$\hat{\beta}_1 = 1.5 \qquad \hat{\beta}_2 = -1.5$$

SS Factor B

$$na\sum_{k=1}^{b}\left(\overline{Y}_{\cdot\cdot k} - \overline{Y}_{\cdots}\right)^2 = 6*3\left[(35-33.5)^2 + (32-33.5)^2\right]$$
$$= 18\left[1.5^2 + (-1.5)^2\right]$$
$$= 18\left[2.25 + 2.25\right]$$
$$= 18\left[4.5\right]$$
$$= 81$$

$$(\hat{\alpha\beta})_{11} = -7.5 \qquad (\hat{\alpha\beta})_{12} = 7.5 \qquad (\hat{\alpha\beta})_{21} = 12.5$$
$$(\hat{\alpha\beta})_{22} = -12.5 \qquad (\hat{\alpha\beta})_{21} = -5.0 \qquad (\hat{\alpha\beta})_{22} = 5.0$$

SS AB Interaction

$$n\sum_{j=1}^{a}\sum_{k=1}^{b}\left(\overline{Y}_{\cdot jk} - \overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdot\cdot k} + \overline{Y}_{\cdots}\right)^2$$
$$= 6\left[(34-35-40+33.5)^2 + \ldots + (38-32-34.5+33.5)^2\right]$$
$$= 6\left[(-7.5)^2 + (7.5)^2 + (12.5)^2 + (-12.5)^2 + (-5.0)^2 + (5.0)^2\right]$$
$$= 6\left[56.25 + 56.25 + 156.25 \quad + 156.25 + 25 + 25\right]$$
$$= 6\left[56.25 + 56.25 + 156.25 + 156.25 + 25 + 25\right]$$
$$= 6\left[475\right]$$
$$= 2850$$

SS Within

$$\sum_{i=1}^{n}\sum_{j=1}^{a}\sum_{k=1}^{b}\left(\overline{Y}_{ijk} - \overline{Y}_{\cdot jk}\right)^2$$

$$= \begin{array}{l}(44-34)^2 + (18-34)^2 + (48-34)^2 + (32-34)^2 + (35-34)^2 + (27-34)^2 + \\ (47-40)^2 + (37-40)^2 + (42-40)^2 + (42-40)^2 + (39-40)^2 + (33-40)^2 + \\ (46-31)^2 + (21-31)^2 + (40-31)^2 + (30-31)^2 + (29-31)^2 + (20-31)^2 + \\ (53-46)^2 + (42-46)^2 + (49-46)^2 + (51-46)^2 + (47-46)^2 + (34-46)^2 + \\ (13-12)^2 + (10-12)^2 + (16-12)^2 + (11-12)^2 + (16-12)^2 + (6-12)^2 + \\ (45-38)^2 + (36-38)^2 + (41-38)^2 + (35-38)^2 + (38-38)^2 + (33-38)^2\end{array}$$
$$= 1668$$

$$\begin{array}{ll} SST & = SSA + SSB + SSAB + SSW \\ 5793 & = 1194 + 81 + 2850 + 1668 \\ & = 5793 \end{array}$$

- This partition works because the tests for Factor A, Factor B, and the AB interaction are orthogonal

6. Tests of main effects and interactions for two-way ANOVA

- For a one-way ANOVA, we constructed an F-test for the factor of interest:

$$F(a-1, N-a) = \frac{MSBet}{MSW}$$

- Why does this test work?

$$E(MSW) = \sigma_\varepsilon^2 \qquad\qquad E(MSBet) = \sigma_\varepsilon^2 + \frac{n\sum \alpha_j^2}{a-1}$$

Under the null hypothesis $\alpha_j = 0$ $\qquad \dfrac{MSBet}{MSW} = \dfrac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2} = 1$

Under the alternative hypothesis $\alpha_j \neq 0$ $\qquad \dfrac{MSBet}{MSW} = \dfrac{\sigma_\varepsilon^2 + \dfrac{n\sum \alpha_j^2}{a-1}}{\sigma_\varepsilon^2} > 1$

- For a two-way ANOVA, we may construct F-tests for the main effect of factor A, the main effect of factor B, and the A*B interaction. For each of these tests, we need to make sure that we can interpret the F-test as a measure of the effect of interest.
- We'll skip the math and jump to the main results

- For a two-factor ANOVA:

  - $E(MSW) = \sigma_\varepsilon^2$
  - $E(MSA) = \sigma_\varepsilon^2 + \dfrac{nb\sum \alpha_j^2}{a-1}$

    To test the effect of Factor A

    $$H_0: \mu_{\cdot 1 \cdot} = \mu_{\cdot 2 \cdot} = \ldots = \mu_{\cdot a \cdot}$$
    $$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$$

    $$F(a-1, N-ab) = \frac{MSA}{MSW} = \frac{\sigma_\varepsilon^2 + \dfrac{nb\sum \alpha_j^2}{a-1}}{\sigma_\varepsilon^2}$$

  - $E(MSB) = \sigma_\varepsilon^2 + \dfrac{na\sum \beta_k^2}{b-1}$

    To test the effect of Factor B

    $$H_0: \mu_{\cdot\cdot 1} = \mu_{\cdot\cdot 2} = \ldots = \mu_{\cdot\cdot b}$$
    $$H_0: \beta_1 = \beta_2 = \ldots = \beta_b = 0$$

    $$F(b-1, N-ab) = \frac{MSB}{MSW} = \frac{\sigma_\varepsilon^2 + \dfrac{na\sum \beta_k^2}{b-1}}{\sigma_\varepsilon^2}$$

  - $E(MSAB) = \sigma_\varepsilon^2 + \dfrac{n\sum\sum (\alpha\beta)_{jk}^2}{(a-1)(b-1)}$

    To test the AB interaction

    $$H_0: (\alpha\beta)_{11} = (\alpha\beta)_{12} = \ldots = (\alpha\beta)_{ab} = 0$$

    $$F[(a-1)(b-1), N-ab] = \frac{MSAB}{MSW} = \frac{\sigma_\varepsilon^2 + \dfrac{n\sum\sum (\alpha\beta)_{jk}^2}{(a-1)(b-1)}}{\sigma_\varepsilon^2}$$

- Using this information, we can construct an ANOVA table

ANOVA

Source of Variation	SS	df	MS	F	P-value
Factor A	SSA	(a-1)	SSA/df_a	MSA/MSW	
Factor B	SSB	(b-1)	SSB/df_b	MSB/MSW	
A * B interaction	SSAB	(a-1)(b-1)	SSAB/df_{ab}	MSAB/MSW	
Within	SSW	N-ab	SSW/df_w		
Total	SST	N-1			

Note that $dfw = (N - ab)$

Why?

$$
\begin{aligned}
dfw \quad &= N - dfA - dfB - dfAB - 1 \text{ (for grand mean)} \\
&= N - (a-1) - (b-1) - (a-1)(b-1) - 1 \\
&= N - a + 1 - b + 1 - ab + a + b - 1 - 1 \\
&= N - ab
\end{aligned}
$$

- For our comprehension example:

ANOVA

Source of Variation	SS	df	MS	F	P-value
Factor A (Lecture)	1194	2	597	10.74	0.001
Factor B (Presentation)	81	1	81	1.46	0.237
A * B interaction (Lecture by Presentation)	2850	2	1425	25.63	0.001
Within	1668	30	55.6		
Total	5793	35			

- In SPSS:

```
UNIANOVA dv  BY iv1 iv2
  /PRINT = DESCRIPTIVE.

UNIANOVA compre  BY lecture present
  /PRINT = DESCRIPTIVE.
```

Between-Subjects Factors

		Value Label	N
LECTURE	1.00	Statistics	12
	2.00	English	12
	3.00	History	12
PRESENT	1.00	Standard	18
	2.00	Computer	18

Descriptive Statistics

Dependent Variable: COMPRE

LECTURE	PRESENT	Mean	Std. Deviation	N
Statistics	Standard	34.0000	11.00909	6
	Computer	46.0000	6.98570	6
	Total	40.0000	10.79562	12
English	Standard	40.0000	4.81664	6
	Computer	12.0000	3.84708	6
	Total	26.0000	15.20167	12
History	Standard	31.0000	10.31504	6
	Computer	38.0000	4.38178	6
	Total	34.5000	8.39372	12
Total	Standard	35.0000	9.41213	18
	Computer	32.0000	15.72933	18
	Total	33.5000	12.86524	36

Tests of Between-Subjects Effects

Dependent Variable: COMPRE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4125.000[a]	5	825.000	14.838	.000
Intercept	40401.000	1	40401.000	726.637	.000
LECTURE	1194.000	2	597.000	10.737	.000
PRESENT	81.000	1	81.000	1.457	.237
LECTURE * PRESENT	2850.000	2	1425.000	25.629	.000
Error	1668.000	30	55.600		
Total	46194.000	36			
Corrected Total	5793.000	35			

a. R Squared = .712 (Adjusted R Squared = .664)

7. Testing assumptions for two-way ANOVA and alternatives to ANOVA

i. All samples are drawn from *normally distributed* populations
ii. All populations have a *common variance*
iii. All *samples were drawn independently* from each other
iv. Within each sample, the *observations were sampled randomly and independently* of each other

- For a two-way ANOVA, we can use the same techniques for testing assumptions that we used for a one-way ANOVA.

- We need to check these assumptions on a cell-by-cell basis
  (NOT on a factor-by-factor basis)

  - Example of a 4*3 design

	Factor A				
Factor B	a_1	a_2	a_3	a_4	
b_1	$N(\mu_{11},\sigma)$	$N(\mu_{21},\sigma)$	$N(\mu_{31},\sigma)$	$N(\mu_{41},\sigma)$	$\mu_{\cdot 1}$
b_2	$N(\mu_{12},\sigma)$	$N(\mu_{22},\sigma)$	$N(\mu_{32},\sigma)$	$N(\mu_{42},\sigma)$	$\mu_{\cdot 2}$
b_3	$N(\mu_{13},\sigma)$	$N(\mu_{23},\sigma)$	$N(\mu_{33},\sigma)$	$N(\mu_{43},\sigma)$	$\mu_{\cdot 3}$
	$\mu_{1\cdot}$	$\mu_{2\cdot}$	$\mu_{3\cdot}$	$\mu_{4\cdot}$	$\mu_{\cdot\cdot}$

- SPSS conducts tests on a factor-by-factor basis

  For the lecture comprehension example:

```
EXAMINE compre BY lecture present
  /PLOT BOXPLOT STEMLEAF NPPLOT SPREADLEVEL.
```

  This syntax will give us:
  - Plots and tests on the type of lecture factor
  - Plots and tests on the type of presentation factor

  But what we need are tests on each cell of the design!

- We can con SPSS into giving us the tests we need by making SPSS think that we have a one-factor design with 6 levels instead of a 2X3 design.

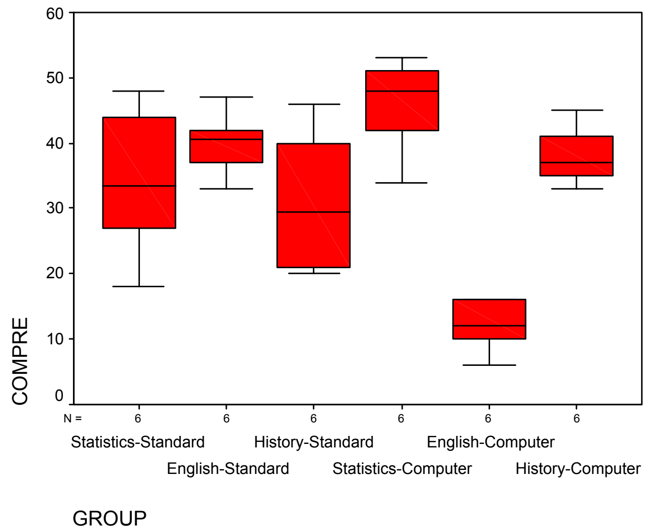
	Type of Lecture (Factor A)		
Method of Presentation (Factor B)	Statistics a_1	English a_2	History a_3
Standard b_1	1	2	3
Computer b_2	4	5	6

Factor					
	Standard			Computer	
Stat	English	History	Stat	English	History
$a_1\ b_1$	$a_2\ b_1$	$a_3\ b_1$	$a_1\ b_2$	$a_2\ b_2$	$a_3\ b_2$
1	2	3	4	5	6

```
if (present=1 and lecture=1) group = 1.
if (present=1 and lecture=2) group = 2.
if (present=1 and lecture=3) group = 3.
if (present=2 and lecture=1) group = 4.
if (present=2 and lecture=2) group = 5.
if (present=2 and lecture=3) group = 6.
```

Now the following command will provide us with all the tests and graphs we need on a cell-by-cell basis.

```
EXAMINE compre BY group
  /PLOT BOXPLOT STEMLEAF NPPLOT SPREADLEVEL.
```



Descriptives

	GROUP		Statistic	Std. Error
COMPRE	Statistics-Standard	Mean	34.0000	4.49444
		Median	33.5000	
		Variance	121.200	
		Std. Deviation	11.00909	
		Interquartile Range	20.2500	
		Skewness	-.158	.845
		Kurtosis	-.705	1.741
	English-Standard	Mean	40.0000	1.96638
		Median	40.5000	
		Variance	23.200	
		Std. Deviation	4.81664	
		Interquartile Range	7.2500	
		Skewness	-.032	.845
		Kurtosis	.143	1.741
	History-Standard	Mean	31.0000	4.21110
		Median	29.5000	
		Variance	106.400	
		Std. Deviation	10.31504	
		Interquartile Range	20.7500	
		Skewness	.482	.845
		Kurtosis	-1.189	1.741
	Statistics-Computer	Mean	46.0000	2.85190
		Median	48.0000	
		Variance	48.800	
		Std. Deviation	6.98570	
		Interquartile Range	11.5000	
		Skewness	-1.141	.845
		Kurtosis	.834	1.741
	English-Computer	Mean	12.0000	1.57056
		Median	12.0000	
		Variance	14.800	
		Std. Deviation	3.84708	
		Interquartile Range	7.0000	
		Skewness	-.506	.845
		Kurtosis	-.415	1.741
	History-Computer	Mean	38.0000	1.78885
		Median	37.0000	
		Variance	19.200	
		Std. Deviation	4.38178	
		Interquartile Range	7.5000	
		Skewness	.749	.845
		Kurtosis	-.166	1.741

Tests of Normality

	GROUP	Shapiro-Wilk Statistic	df	Sig.
COMPRE	Statistics-Standard	.976	6	.933
	English-Standard	.981	6	.958
	History-Standard	.921	6	.515
	Statistics-Computer	.912	6	.451
	English-Computer	.929	6	.571
	History-Computer	.955	6	.783

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
COMPRE	Based on Mean	2.058	5	30	.099
	Based on Median	1.613	5	30	.187
	Based on Median and with adjusted df	1.613	5	19.476	.203
	Based on trimmed mean	1.973	5	30	.112

- We can also examine cell-by-cell histograms and Q-Q plots (But with n=6, these will be difficult to interpret)

- What can we do if the assumptions are violated?

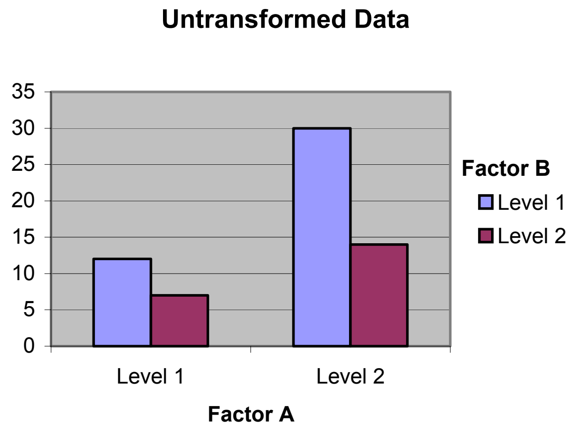
  - Transformations tend to be dangerous with a higher-order ANOVA
    - One application of transformations is to eliminate or reduce an interaction

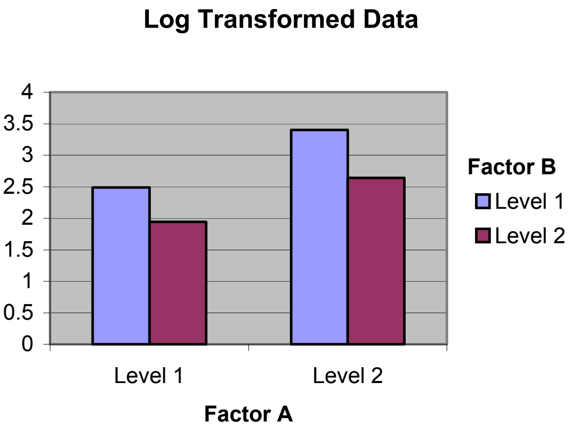
$$y = ab$$

This equation specifies a model with an AxB interaction (and no main effects)

$$\ln(y) = \ln(a) + \ln(b)$$

After a log transformation, we have a main effect of ln(a) and a main effect of ln(b), but no interaction





Here we see an A*B interaction

Now the interaction has disappeared

- In other words, transformations applied to fix heterogeneity of variances and/or non-normality may eliminate or produce interactions!

- Which method/analysis is "right"?
  - In psychology, we typically do not know what the true model is (nor do we have a clue what the real model would look like)
  - Looking at residuals can help determine if you have a good model for your data

  - The main point is that what appears to be an interaction may be a question of having the right scale
  - And remember that when you transform your data, the conclusions you draw are always on the transformed scale!

- Non-parametric/rank based methods for higher-order ANOVA are not very straightforward either
  - Different tests are needed to examine the main effects and the interactions
  - The statistical properties of these tests have not been fully ironed out.

- For <u>equal *n*</u> two-factor designs, a relatively simple extension of the Brown-Forsythe F^* test is available (but not included in SPSS).
    - Recall that for equal n designs, $F^* = F$
    - Also, the numerator dfs remain the same for both F and F^*
    - We just need to calculate the adjusted denominator dfs, f

      (This adjusted df is used for all three F^* tests: the main effect of A, the main effect of B, and the A*B interaction)

$$g = \frac{1}{\sum_{k=1}^{b}\sum_{j=1}^{a} s_{jk}^2}$$

$$f = \frac{n-1}{\sum_{k=1}^{b}\sum_{j=1}^{a}\left(s_{jk}^2 g\right)^2}$$

$$F^*_{obs}\,(ndf, f) = F$$

    - For unequal *n* two-factor designs, the process gets considerately more complicated (for details, see Brown & Forsythe, 1974)

    - Although multi-factor ANOVA offers some nice advantages, one disadvantage is that we do not have many options when the statistical assumptions are not met.

- If we can live without the omnibus tests then we can ignore the fact that we have a two-way design, and treat the design as a one factor ANOVA. We can run contrasts to test our specific hypotheses AND we can use the Welch's unequal variances correction for contrasts.

8. Follow-up tests and contrasts in two-way ANOVA

i. Contrasts

- In general, a contrast is a set of weights that defines a specific comparison over the cell means.
- For a one-way ANOVA, we had:

$$\psi = \sum_{j=1}^{a} c_i \mu_i = c_1\mu_1 + c_2\mu_2 + c_3\mu_3 + ... + c_a\mu_a$$

$$\hat{\psi} = \sum_{j=1}^{a} c_i \bar{X}_i = c_1\bar{X}_1 + c_2\bar{X}_2 + c_3\bar{X}_3 + ... + c_a\bar{X}_a$$

- For a multi-factor ANOVA, we have many more means:
  - Main effect means (marginal means)
  - Cell means

	IV 1			
IV 2	Level 1	Level 2	Level 3	
Level 1	$\bar{X}_{\bullet 11}$	$\bar{X}_{\bullet 21}$	$\bar{X}_{\bullet 31}$	$\bar{X}_{\bullet\bullet 1}$
Level 2	$\bar{X}_{\bullet 12}$	$\bar{X}_{\bullet 22}$	$\bar{X}_{\bullet 32}$	$\bar{X}_{\bullet\bullet 2}$
Level 3	$\bar{X}_{\bullet 13}$	$\bar{X}_{\bullet 23}$	$\bar{X}_{\bullet 33}$	$\bar{X}_{\bullet\bullet 3}$
	$\bar{X}_{\bullet 1\bullet}$	$\bar{X}_{\bullet 2\bullet}$	$\bar{X}_{\bullet 3\bullet}$	

  - Contrasts on IV1 means involve the marginal means for IV1: $\bar{X}_{\bullet 1\bullet}$, $\bar{X}_{\bullet 2\bullet}$, $\bar{X}_{\bullet 3\bullet}$

$$\hat{\psi}_{IV1} = \sum_{j=1}^{r} c_j \bar{X}_{\bullet j\bullet} = c_1\bar{X}_{\bullet 1\bullet} + c_2\bar{X}_{\bullet 2\bullet} + c_3\bar{X}_{\bullet 3\bullet}$$

  - Contrasts on IV2 means involve the marginal means for IV2: $\bar{X}_{\bullet\bullet 1}$, $\bar{X}_{\bullet 2\bullet}$, $\bar{X}_{\bullet\bullet 3}$

$$\hat{\psi}_{IV2} = \sum_{k=1}^{q} c_k \bar{X}_{\bullet\bullet k} = c_1\bar{X}_{\bullet\bullet 1} + c_2\bar{X}_{\bullet\bullet 2} + c_3\bar{X}_{\bullet\bullet 3}$$

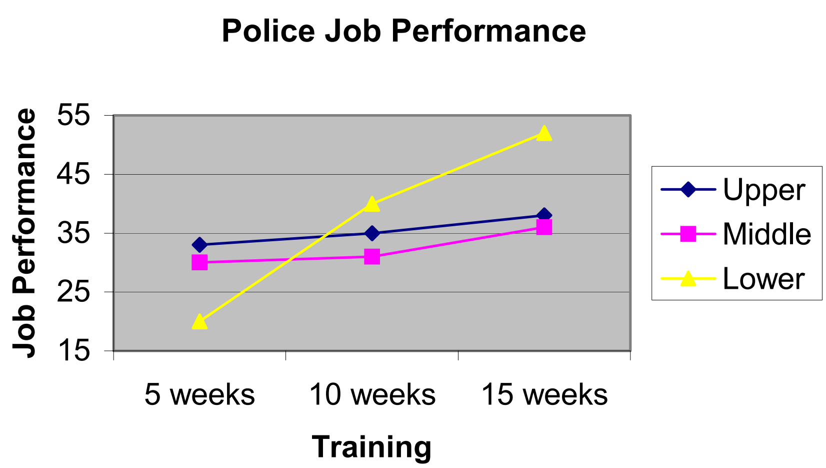
  - Interaction contrasts and more specific contrasts can be performed on the cell means

$$\hat{\psi} = \sum_{k=1}^{b}\sum_{j=1}^{a} c_{jk} \bar{X}_{\bullet jk}$$

- As for a oneway ANOVA, t-tests or F-tests can be used to determine significance

An Example: Police job performance

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks	
Level 1: Upper Class	24 33 37 29 42 $\bar{X}_{\cdot 11} = 33$	44 36 25 27 43 $\bar{X}_{\cdot 21} = 35$	38 29 28 47 48 $\bar{X}_{\cdot 31} = 38$	$\bar{X}_{\cdot\cdot 1} = 35.33$
Level 2: Middle Class	30 21 39 26 34 $\bar{X}_{\cdot 12} = 30$	35 40 27 31 22 $\bar{X}_{\cdot 22} = 31$	26 27 36 46 45 $\bar{X}_{\cdot 32} = 36$	$\bar{X}_{\cdot\cdot 2} = 32.33$
Level 3: Lower Class	21 18 10 31 20 $\bar{X}_{\cdot 13} = 20$	41 39 50 36 34 $\bar{X}_{\cdot 23} = 40$	42 52 53 49 64 $\bar{X}_{\cdot 33} = 52$	$\bar{X}_{\cdot\cdot 3} = 37.33$
	$\bar{X}_{\cdot 1\cdot} = 27.67$	$\bar{X}_{\cdot 2\cdot} = 35.33$	$\bar{X}_{\cdot 3\cdot} = 42$	




```
UNIANOVA perform  BY duration location
  /PRINT = DESCRIPTIVE.
```

Tests of Between-Subjects Effects

Dependent Variable: PERFORM

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2970.000[a]	8	371.250	5.940	.000
Intercept	55125.000	1	55125.000	882.000	.000
DURATION	1543.333	2	771.667	12.347	.000
LOCATION	190.000	2	95.000	1.520	.232
DURATION * LOCATION	1236.667	4	309.167	4.947	.003
Error	2250.000	36	62.500		
Total	60345.000	45			
Corrected Total	5220.000	44			

a. R Squared = .569 (Adjusted R Squared = .473)

ii. Follow-up tests on main effects: Main Effect Contrasts

- When an independent variable has more than 2 levels, the test for the main effect of that variable is an omnibus test. When you reject the null hypothesis, you can only say that not all the marginal means are equal for the IV. We would like to be able to specify where the significant differences are.

- Contrasts on the marginal means of an independent variable are called Main Effect Contrasts

  - To conduct Main Effect Contrasts on the duration of training:

IV 1: Training Duration			
Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks	
$\bar{X}_{\cdot 1 \cdot} = 27.67$	$\bar{X}_{\cdot 2 \cdot} = 35.33$	$\bar{X}_{\cdot 3 \cdot} = 42$	$n_{\cdot j \cdot} = 15$

  - To conduct Main Effect Contrasts on the office location:

IV 2: Location of Office	
Level 1: Upper Class	$\bar{X}_{\cdot\cdot 1} = 35.33$
Level 2: Middle Class	$\bar{X}_{\cdot\cdot 2} = 32.33$
Level 3: Lower Class	$\bar{X}_{\cdot\cdot 3} = 37.33$
	$n_{\cdot\cdot k} = 15$

- Computing and testing a Main Effect Contrast

$$\hat{\psi} = \sum_{j=1}^{a} c_j \overline{X}_{\cdot j \cdot} = c_1 \overline{X}_{\cdot 1 \cdot} + \ldots + c_r \overline{X}_{\cdot a \cdot}$$

$$\text{Std error } (\hat{\psi}) = \sqrt{MSW \sum_{j=1}^{a} \frac{c_j^2}{n_j}}$$

Where c_j^2 is the squared weight for each marginal mean
n_j is the sample size for each marginal mean
MSW is MSW from the omnibus ANOVA

$$t \sim \frac{\hat{\psi}}{\text{standard error}(\hat{\psi})} \qquad t_{observed} = \frac{\sum c_j \overline{X}_{\cdot j \cdot}}{\sqrt{MSW \sum \frac{c_j^2}{n_j}}}$$

$$SS(\hat{\psi}) = \frac{\hat{\psi}^2}{\sum \frac{c_j^2}{n_j}}$$

$$F(1, dfw) = \frac{SSC/dfc}{SSW/dfw} = \frac{SSC}{MSW}$$

- For example, let's test for linear and quadratic trends in amount of training on job performance

$\psi_{lin} : (-1,0,1)$ $\qquad$ $\psi_{quad} : (1,-2,1)$

$$\begin{aligned}\hat{\psi}_{linear} &= -\overline{X}_1 + 0\overline{X}_2 + \overline{X}_3 \\ &= -(27.67) - (0) + (42.0) \\ &= 14.33\end{aligned} \qquad \begin{aligned}\hat{\psi}_{quadratic} &= \overline{X}_1 - 2\overline{X}_2 + \overline{X}_3 \\ &= (27.67) - 2(35.33) + (42.0) \\ &= -1\end{aligned}$$

$$SS(\hat{\psi}_{linear}) = \frac{(14.33)^2}{\frac{(-1)^2}{15} + \frac{(0)^2}{15} + \frac{(1)^2}{15}} = \frac{205.44}{.133} = 1540.83$$

$$F(1,36) = \frac{1540.83}{62.5} = 24.65, p < .01$$

$$SS(\hat{\psi}_{quadratic}) \frac{(-1)^2}{\frac{(1)^2}{15} + \frac{(-2)^2}{15} + \frac{(1)^2}{15}} = \frac{1}{.4} = 2.5$$

$$F(1,36) = \frac{2.5}{62.5} = .04, p = .84$$

- Note that this process is identical to the oneway contrasts we previously developed. The only difference is that we now average across the levels of another IV
  - You need all the assumptions to be satisfied for the marginal means of interest
  - If the assumptions are not satisfied, you can rely on the fixes we developed for oneway ANOVA
- Main effect contrasts are usually post-hoc tests and require adjustment of the *p*-value. However, there is no reason why you cannot hypothesize about main effect contrasts, making these tests planned contrasts. (More to follow regarding planned and post-hoc tests for multi-factor ANOVA)

- Main effect contrasts in SPSS GLM/UNIANOVA using the CONTRAST subcommand
  - The CONTRAST subcommand can be used to test main effect contrasts if you wish to conduct the built-in, brand-name contrasts (polynomial, Helmert, etc.)

```
UNIANOVA perform  BY duration location
  /CONTRAST (duration)=Polynomial
  /PRINT = DESCRIPTIVE.
```

  - Note: This syntax will provide polynomial main effect contrasts on the *duration* marginal means.

Contrast Results (K Matrix)

DURATION Polynomial Contrast[a]			Dependent Variable PERFORM
Linear	Contrast Estimate		10.135
	Hypothesized Value		0
	Difference (Estimate - Hypothesized)		10.135
	Std. Error		2.041
	Sig.		.000
	95% Confidence Interval for Difference	Lower Bound	5.995
		Upper Bound	14.275
Quadratic	Contrast Estimate		-.408
	Hypothesized Value		0
	Difference (Estimate - Hypothesized)		-.408
	Std. Error		2.041
	Sig.		.843
	95% Confidence Interval for Difference	Lower Bound	-4.548
		Upper Bound	3.732

a. Metric = 1.000, 2.000, 3.000

Linear trend for duration: $t(36) = 4.97, p < .001$
Quadratic trend for duration: $t(36) = -.20, p = .84$

  - These results match our hand calculations on the previous page

  - If you cannot test your main effect contracts using SPSS's brand-name contrasts, then you must resort to hand calculations.

- Main effect contrasts in SPSS GLM/UNIANOVA using the EMMEANS subcommand
  - The EMMEANS subcommand can be used to test all possible pairwise contrasts on the marginal main effect means.

```
UNIANOVA perform  BY duration location
  /EMMEANS = TABLES(duration) COMPARE
  /EMMEANS = TABLES(location) COMPARE
  /PRINT = DESCRIPTIVE.
```

  - The first EMMEANS comment asks for pairwise contrasts on the marginal duration means

Estimates

Dependent Variable: perform

duration	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
5 Weeks	27.667	2.041	23.527	31.806
10 Weeks	35.333	2.041	31.194	39.473
15 Weeks	42.000	2.041	37.860	46.140

Pairwise Comparisons

Dependent Variable: perform

(I) duration	(J) duration	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
5 Weeks	10 Weeks	-7.667*	2.887	.012	-13.521	-1.812
	15 Weeks	-14.333*	2.887	.000	-20.188	-8.479
10 Weeks	5 Weeks	7.667*	2.887	.012	1.812	13.521
	15 Weeks	-6.667*	2.887	.027	-12.521	-.812
15 Weeks	5 Weeks	14.333*	2.887	.000	8.479	20.188
	10 Weeks	6.667*	2.887	.027	.812	12.521

Based on estimated marginal means

*. The mean difference is significant at the .050 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

5 Weeks v. 10 Weeks: $t(36) = -2.64, p = .01$

5 Weeks v. 15 Weeks: $t(36) = -4.96, p < .01$

10 Weeks v. 15 Weeks: $t(36) = -2.31, p = .03$

- The second EMMEANS comment asks for pairwise contrasts on the marginal location means

Estimates

Dependent Variable: perform

location	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Upper Class	35.333	2.041	31.194	39.473
Middle Class	32.333	2.041	28.194	36.473
Lower Class	37.333	2.041	33.194	41.473

Pairwise Comparisons

Dependent Variable: perform

(I) location	(J) location	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
Upper Class	Middle Class	3.000	2.887	.306	-2.855	8.855
	Lower Class	-2.000	2.887	.493	-7.855	3.855
Middle Class	Upper Class	-3.000	2.887	.306	-8.855	2.855
	Lower Class	-5.000	2.887	.092	-10.855	.855
Lower Class	Upper Class	2.000	2.887	.493	-3.855	7.855
	Middle Class	5.000	2.887	.092	-.855	10.855

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Upper v. Middle Class: $t(36) = 1.04, p = .31$

Upper v. Lower Class: $t(36) = -0.69, p = .49$

Middle v. Lower Class: $t(36) = -1.73, p = .09$

- To confirm these tests, let's compute Upper v. Middle Class by hand:

$$t(36) = \frac{\sum c_j \overline{X}_{\cdot j \cdot}}{\sqrt{MSW \sum \frac{c_j^2}{n_j}}} = \frac{1*35.333 + (-1)*32.333 + 0*37.333}{\sqrt{62.5\left[\frac{1}{15} + \frac{1}{15} + 0\right]}} = \frac{3.00}{2.887} = 1.04$$

iii. Follow-up tests on interactions: Simple (Main) Effects

- When you have an interaction with more than 1 degree of freedom (either a>2 or b>2), the test for the interaction between those variables is an omnibus test. When you reject the null hypothesis, you can only say that the main effect of one IV is not equal across all levels of the second IV. We would like to be able to specify where the significant differences are.

- Contrasts on the cell means of one IV *within one level* of another IV are called Simple Effect Contrasts

  - Is there an effect of training duration on job performance among police officers who work in upper class neighborhoods? … in middle class neighborhoods? … in lower class neighborhoods?

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	$\boxed{\bar{X}_{\cdot 11} = 33}$	$\boxed{\bar{X}_{\cdot 21} = 35}$	$\boxed{\bar{X}_{\cdot 31} = 38}$
Level 2: Middle Class	$\bar{X}_{\cdot 12} = 30$	$\bar{X}_{\cdot 22} = 31$	$\bar{X}_{\cdot 32} = 36$
Level 3: Lower Class	$\bar{X}_{\cdot 13} = 20$	$\bar{X}_{\cdot 23} = 40$	$\bar{X}_{\cdot 33} = 52$

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	$\bar{X}_{\cdot 11} = 33$	$\bar{X}_{\cdot 21} = 35$	$\bar{X}_{\cdot 31} = 38$
Level 2: Middle Class	$\boxed{\bar{X}_{\cdot 12} = 30}$	$\boxed{\bar{X}_{\cdot 22} = 31}$	$\boxed{\bar{X}_{\cdot 32} = 36}$
Level 3: Lower Class	$\bar{X}_{\cdot 13} = 20$	$\bar{X}_{\cdot 23} = 40$	$\bar{X}_{\cdot 33} = 52$

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	$\bar{X}_{\cdot 11} = 33$	$\bar{X}_{\cdot 21} = 35$	$\bar{X}_{\cdot 31} = 38$
Level 2: Middle Class	$\bar{X}_{\cdot 12} = 30$	$\bar{X}_{\cdot 22} = 31$	$\bar{X}_{\cdot 32} = 36$
Level 3: Lower Class	$\boxed{\bar{X}_{\cdot 13} = 20}$	$\boxed{\bar{X}_{\cdot 23} = 40}$	$\boxed{\bar{X}_{\cdot 33} = 52}$

- Let's return to the lecture comprehension example

  - We found that there is a main effect for type of lecture and a lecture by presentation interaction
  - The presence of the interaction indicates that the main effect for type of lecture is not equal across all methods of presentation (or equivalently, that the main effect of method of presentation is not equal across all types of lectures)

$n_{jk} = 6$	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	$\bar{X}_{\cdot 11} = 34$	$\bar{X}_{\cdot 21} = 40$	$\bar{X}_{\cdot 31} = 31$
Computer	$\bar{X}_{\cdot 12} = 46$	$\bar{X}_{\cdot 22} = 12$	$\bar{X}_{\cdot 32} = 38$

- Computing and testing simple effects contrasts using SPSS

  - Is there an effect of method of presentation for statistics lectures?

$n_{jk} = 6$	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	-1	0	0
Computer	1	0	0

```
ONEWAY compre by group
 /CONT = -1 0 0 1 0 0.
```

Contrast Tests

Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
COMPRE	12.0000	4.30504	2.787	30	.009

The test of the simple effect of method of presentation within <u>statistics lectures</u> reveals that computer presentations were understood better than standard presentations, $t(30) = 2.79, p = .009$.

- Is there an effect of method of presentation for English lectures?

$n_{jk}=6$	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	0	-1	0
Computer	0	1	0

```
ONEWAY compre by group
 /CONT = 0 -10 0 1 0.
```

Contrast Tests

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
COMPRE	1	-28.0000	4.30504	-6.504	30	.000

The test of the simple effect of method of presentation within English lectures reveals that standard presentations were understood better than computer presentations, $t(30) = -6.50, p < .001$.

- Is there an effect of method of presentation for history lectures?

$n_{jk}=6$	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	0	0	-1
Computer	0	0	1

```
ONEWAY compre by group
 /CONT = 0 0 -1 0 0 1.
```

Contrast Tests

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
COMPRE	1	7.0000	4.30504	1.626	30	.114

The test of the simple effect of method of presentation within history lectures reveals no significant differences in comprehension between standard presentations and computer presentations, $t(30) = 1.63, p = .114$.

- Alternatively, the simple effects of presentation within each type of lecture can be obtained by using the EMMEANS subcommand of GLM/UNIANOVA:

```
UNIANOVA compre  BY lecture present
/EMMEANS = TABLES(lecture*present) COMPARE (present).
```

  - The EMMEANS command asks for cell means (lecture*present) and for comparisons of the variable present within each level of lecture.

Estimates

Dependent Variable: compre

lecture	present	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	Standard	34.000	3.044	27.783	40.217
	Computer	46.000	3.044	39.783	52.217
English	Standard	40.000	3.044	33.783	46.217
	Computer	12.000	3.044	5.783	18.217
History	Standard	31.000	3.044	24.783	37.217
	Computer	38.000	3.044	31.783	44.217

Pairwise Comparisons

Dependent Variable: compre

lecture	(I) present	(J) present	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
Statistics	Standard	Computer	-12.000*	4.305	.009	-20.792	-3.208
	Computer	Standard	12.000*	4.305	.009	3.208	20.792
English	Standard	Computer	28.000*	4.305	.000	19.208	36.792
	Computer	Standard	-28.000*	4.305	.000	-36.792	-19.208
History	Standard	Computer	-7.000	4.305	.114	-15.792	1.792
	Computer	Standard	7.000	4.305	.114	-1.792	15.792

Based on estimated marginal means

*. The mean difference is significant at the .050 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

  - Simple effect of presentation within statistics lectures:
    $t(30) = 2.79, p = .009$
  - Simple effect of presentation within English lectures:
    $t(30) = 6.50, p < .001$
  - Simple effect of presentation within history lectures:
    $t(30) = 1.62, p = .114$

- Computing and testing simple effects contrasts by hand

$$\hat{\psi} = \sum_{k=1}^{b}\sum_{j=1}^{a} c_{jk}\overline{X}_{\cdot jk} = c_{11}\overline{X}_{\cdot 11} + ... + c_{ab}\overline{X}_{\cdot ab}$$

$$\text{Std error } (\hat{\psi}) = \sqrt{MSW\sum_{k=1}^{b}\sum_{j=1}^{a}\frac{c_{jk}^2}{n_{jk}}}$$

Where c_{jk}^2 is the squared weight for each cell
n_{jk} is the sample size for each cell
MSW is MSW from the omnibus ANOVA

$$t \sim \frac{\hat{\psi}}{\text{standard error}(\hat{\psi})} \qquad t_{observed} = \frac{\sum\sum c_{jk}\overline{X}_{\cdot jk}}{\sqrt{MSW\sum\sum\frac{c_{jk}^2}{n_{jk}}}}$$

$$\text{SS}\,\hat{\psi} = \frac{\hat{\psi}^2}{\sum\sum\frac{c_{jk}^2}{n_{jk}}} \qquad F(1,dfw) = \frac{SSC/dfc}{SSW/dfw} = \frac{SSC}{MSW}$$

- For example, let's test if there is an effect of method of presentation for history lectures.

$$\hat{\psi} = \sum_{k=1}^{q}\sum_{j=1}^{r} c_{jk}\overline{X}_{\cdot jk} = 0\overline{X}_{\cdot 11} + 0\overline{X}_{\cdot 21} - \overline{X}_{\cdot 31} + 0\overline{X}_{\cdot 12} + 0\overline{X}_{\cdot 22} + \overline{X}_{\cdot 32}$$

$$\hat{\psi} = -31 + 38 = 7$$

$$t_{observed} = \frac{7}{\sqrt{55.6\left(0+0+\frac{1}{6}+0+0+\frac{1}{6}\right)}} = \frac{7}{4.305} = 1.62$$

$$t(30) = 1.62, p = .114$$

- Note that if we had decided to investigate the effect of type of lecture within each method of presentation, our lives would have been more complicated!

$n_{jk} = 6$	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	$\overline{X}_{\cdot 11} = 34$	$\overline{X}_{\cdot 21} = 40$	$\overline{X}_{\cdot 31} = 31$
Computer	$\overline{X}_{\cdot 12} = 46$	$\overline{X}_{\cdot 22} = 12$	$\overline{X}_{\cdot 32} = 38$

- Each simple effect would have 2 degrees of freedom (They would be omnibus tests)

- In this case where the simple effect has more than 1 degree of freedom, a significant simple effect test will have to be followed by additional tests to identify where the differences are.

- To test an omnibus simple effect
  - ⇒ Construct a-1 orthogonal contrasts (in this case 2)
  - ⇒ Compute the sums of squares of each contrast
  - ⇒ Test the contrasts simultaneously with an $(a\text{-}1)$ df omnibus test

$$F(a-1, dfw) = \frac{\left(\dfrac{SS\hat{\psi}_1 + ... + SS\hat{\psi}_{(a-1)}}{a-1}\right)}{MSW}$$

- Alternatively (and more simply), we can use the EMMEANS subcommand of GLM/UNIANOVA to compute the omnibus simple effects.

```
UNIANOVA compre  BY lecture present
  /EMMEANS = TABLES(lecture*present) COMPARE (lecture)
```

$n_{jk} = 6$	Type of Lecture		
Method of Presentation	Statistics	English	History
Standard	$\overline{X}_{\cdot 11} = 34$	$\overline{X}_{\cdot 21} = 40$	$\overline{X}_{\cdot 31} = 31$
Computer	$\overline{X}_{\cdot 12} = 46$	$\overline{X}_{\cdot 22} = 12$	$\overline{X}_{\cdot 32} = 38$

Univariate Tests

Dependent Variable: compre

present		Sum of Squares	df	Mean Square	F	Sig.
Standard	Contrast	252.000	2	126.000	2.266	.121
	Error	1668.000	30	55.600		
Computer	Contrast	3792.000	2	1896.000	34.101	.000
	Error	1668.000	30	55.600		

Each F tests the simple effects of lecture within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

- Simple effect of type of lecture within standard presentations:
  $F(2,30) = 2.27, p = .121$
- Simple effect of type of lecture within computer presentations:
  $F(2,30) = 34.10, p < .001$

- This syntax also gives us all pairwise contrasts with each level of presentation (but that output is not displayed here).

iv. Tests of more specific hypotheses involving cell means

- Thus far, we have developed procedures for understanding:
  - Main effects (Main effect contrasts)
  - Interactions (Simple effects)

- But you may have developed a specific hypothesis that does not fall into one of these categories.
  - Suppose you want to compare officers who work in upper class neighborhoods and receive up to 10 weeks of training, to officers who work in lower class neighborhoods and receive up to 10 weeks of training

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	$\overline{X}_{\cdot 11} = 33$	$\overline{X}_{\cdot 21} = 35$	$\overline{X}_{\cdot 31} = 38$
Level 2: Middle Class	$\overline{X}_{\cdot 12} = 40$	$\overline{X}_{\cdot 22} = 31$	$\overline{X}_{\cdot 32} = 36$
Level 3: Lower Class	$\overline{X}_{\cdot 11} = 20$	$\overline{X}_{\cdot 21} = 40$	$\overline{X}_{\cdot 31} = 52$

- We need to convert the hypothesis to a set of contrast coefficients

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	-1	-1	0
Level 2: Middle Class	0	0	0
Level 3: Lower Class	1	1	0

- Now, we can use the same formulas we developed for simple effect/interaction contrasts to test this specific contrast

$$\hat{\psi} = \sum_{k=1}^{b}\sum_{j=1}^{a} c_{jk}\overline{X}_{\cdot jk} = c_{11}\overline{X}_{\cdot 11} + \ldots + c_{ab}\overline{X}_{\cdot ab}$$

$$SS\hat{\psi} = \frac{\hat{\psi}^2}{\sum\sum \frac{c_{jk}^2}{n_{jk}}} \qquad F(1, dfw) = \frac{SSC/dfc}{SSW/dfw} = \frac{SSC}{MSW}$$

$$\hat{\psi} = -33 - 35 + 20 + 40 = -8$$

$$SS\hat{\psi} = \frac{8^2}{\sum \frac{-(1)^2}{5} + \frac{(-1)^2}{5} + \frac{(1)^2}{5} + \frac{(1)^2}{5}} = 80$$

$$F(1,36) = \frac{MSC}{MSW} = \frac{80}{62.5} = 1.28$$

$$F(1,36) = 1.28, p = .27$$

- Or we can have SPSS ONEWAY compute these contrasts

	IV 1: Training Duration		
IV 2: Location of Office	Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	1	2	3
Level 2: Middle Class	4	5	6
Level 3: Lower Class	7	8	9

```
if (duration=1 and location=1) group = 1.
if (duration=2 and location=1) group = 2.
if (duration=3 and location=1) group = 3.
if (duration=1 and location=2) group = 4.
if (duration=2 and location=2) group = 5.
if (duration=3 and location=2) group = 6.
if (duration=1 and location=3) group = 7.
if (duration=2 and location=3) group = 8.
if (duration=3 and location=3) group = 9.
```

```
ONEWAY perform by group
  /CONT = -1 -1 0 0 0 0 1 1 0.
```

Contrast Tests

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
PERFORM	1	-8.0000	7.07107	-1.131	36	.265

$$t(36) = -1.13, p = .27$$

- Now you can conduct any contrasts and omnibus tests for a two-way ANOVA designs

- But a caveat! Consider the following contrast:

	IV 1: Training Duration		
IV 2: Location of Office	Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	0	1	0
Level 2: Middle Class	-1	0	0
Level 3: Lower Class	0	0	0

  - This contrast confounds two variables
    10 weeks training AND Upper class neighborhood
    vs. 5 weeks training AND Middle class neighborhood

  - If you find a difference, you will not know if it is due to the difference in training, or due to the difference in location.

- Be careful of conducting contrasts that are statistically valid, but that are ambiguous in interpretation!

- To cement your understanding of main effects and contrasts, it is very illuminating to see how omnibus main effect tests can be conducted by combining contrasts. For this information, see Appendix A.

9. Planned tests and post-hoc tests

- The same logic we outlined for the oneway design applies to a two-way design

  - Planned tests: If you plan to conduct tests before looking at the data, then you need to worry about the problem of multiple tests inflating the type one error rate

  - Post-hoc tests: If you decide to conduct tests after looking at the data, then you need to worry about the problem of multiple tests, but you also need to worry that your tests may be capitalizing on random differences

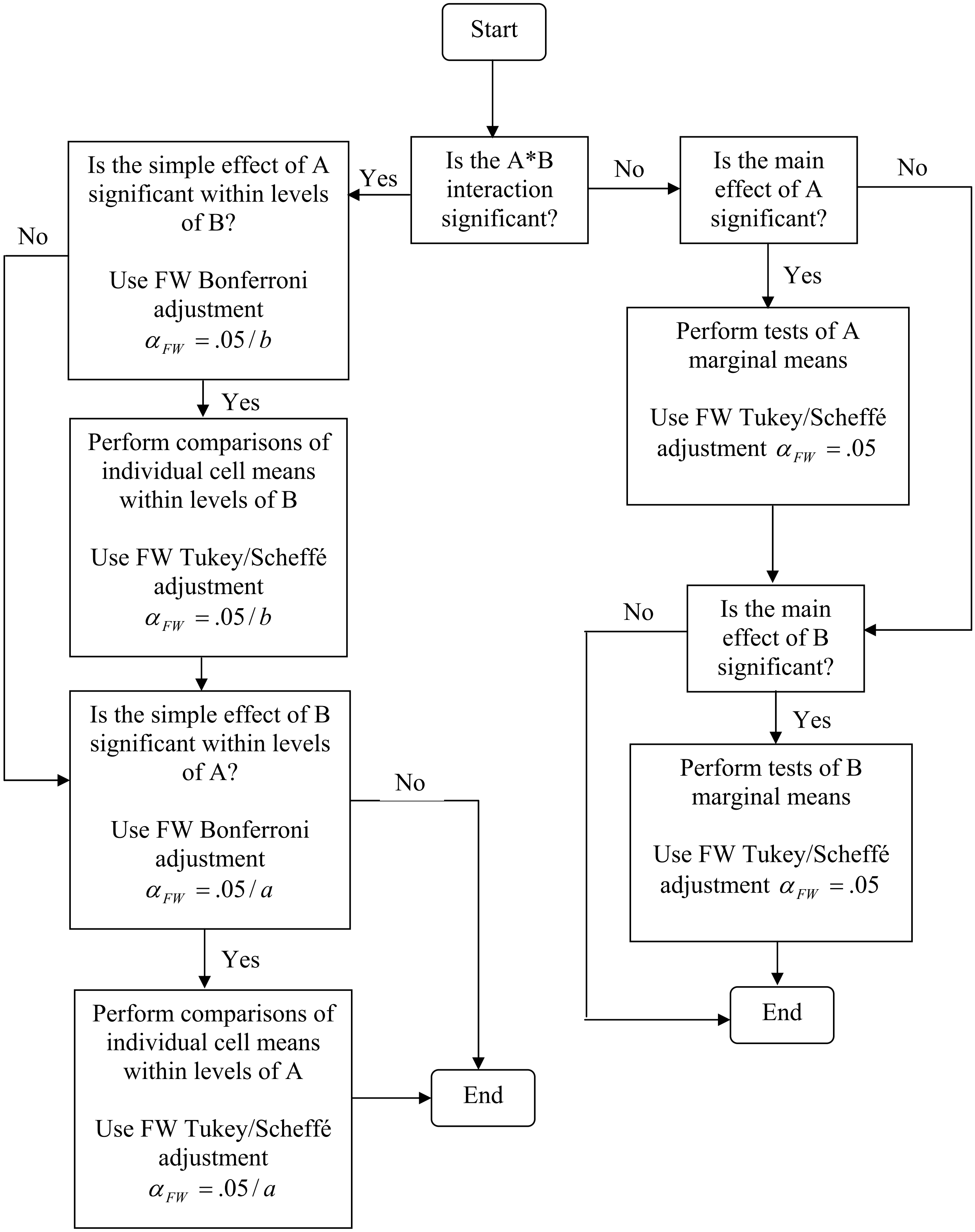
- Which error rate to control – Experiment-wise or Family-wise?
  - To control the Experiment-wise error rate (α_{EW}) we would like to keep the probability of committing a Type 1 error across the entire experiment at $\alpha_{EW} = .05$.
  - Because we use $\alpha = .05$ for testing the main effects and the interaction, we have an inflated Type 1 error rate if we only conduct the omnibus tests!

$$\alpha_{EW} = 1 - (1 - .05)^3 = .14$$

  - Thus, common convention to control the Family-wise error rate at $\alpha_{FW} = .05$ instead

$$\alpha_{FactorA} = .05 \qquad \alpha_{FactorB} = .05 \qquad \alpha_{A*B} = .05$$

Maxwell and Delaney's (1990)
Guidelines for Analyzing Effects in a Two-factor Design



- Advantages of the Maxwell and Delaney Model
  - DO NOT interpret main effects in the presence of an interaction!
  - $\alpha_{FW} = .05$

- Disadvantages of the Maxwell and Delaney Model
  - You may never test your research hypotheses!
  - Can be cumbersome to conduct post-hoc tests with $\alpha = .05 / a$

- A contrast-based method of analysis
  - For an $a*b$ two factor design, you would use ab-1 degrees of freedom if you conduct the omnibus tests:
    - a-1 df for the main effect of Factor 1
    - b-1 df for the main effect of Factor 2
    - $(a\text{-}1)(b\text{-}1)$ df for the interaction of Factor 1 and Factor 2

  - Thus, according to the logic I outlined for a one-factor design, you should be entitled to ab-1 uncorrected planned contrasts

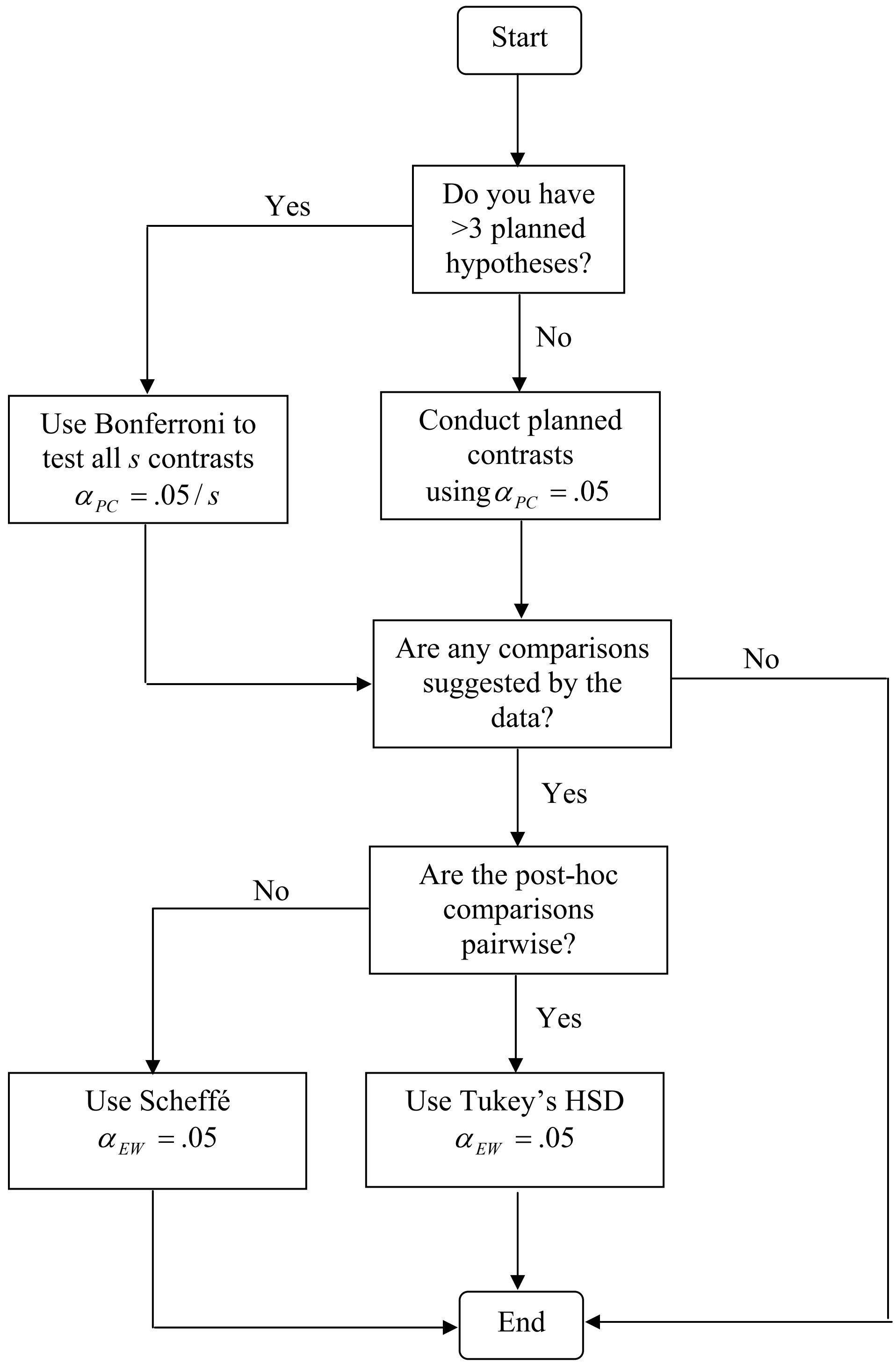
2*2 design	3 uncorrected contrasts
2*3 design	5 uncorrected contrasts
4*5 design	19 uncorrected contrasts

  - But this logic can lead to a large number of uncorrected contrasts. For example in a 4*5 design with $\alpha_{PC} = .05$, the actual probability of making a type 1 error across the entire experiment is:

$$\alpha_{EW} = 1 - (1 - .05)^{19} = .62$$

  - To be on the safe side, we should probably only conduct at most three uncorrected tests in a two-way design – the same number of uncorrected omnibus tests others may have conducted. And remember, you are conducting these contrasts in place of (not in addition to) the omnibus main effect and interaction tests!

A Contrast-Based Approach
for Analyzing Effects in a Two-Factor Design



- The method for conducting post-hoc adjustments is same as for one-way design
  - Obtain observed t- or F-statistic by hand (or using SPSS, but discard printed p-value)
  - Look up critical value and compare to observed value

    - For Tukey's HSD using marginal means: $q(1-\alpha, a, \nu)$

      Where α = Familywise error rate
      a = Number of groups in the factor
      $\nu = DFw = N\text{-}ab$

    - For Tukey's HSD using cell means: $q(1-\alpha, ab, \nu)$

      Where α = Familywise error rate
      ab = Number of cells in the design
      $\nu = DFw = N\text{-}ab$

Compare $t_{observed}$ to $\frac{q_{crit}}{\sqrt{2}}$ or $F_{observed}$ to $\frac{(q_{crit})^2}{2}$

- For Scheffé using marginal means: $F_{Crit} = (a-1)F_{\alpha=.05;a-1,N-ab}$

- For Scheffé using cell means: $F_{Crit} = (a-1)(b-1)F_{\alpha=.05;(a-1)(b-1),N-ab}$

Compare $F_{observed}$ to F_{crit}

10. Effect Sizes

- Omega Squared(ω^2)
  - Omega squared is a measure of the proportion of the variance of the dependent variable that is explained by the factor/contrast of interest. ω^2 generalizes to the population

  - Previously we used the following formulas

$$\hat{\omega}^2 = \frac{SSBetween - (a-1)MSWithin}{SSTotal + MSWithin} \qquad \text{or} \qquad \hat{\omega}^2 = \frac{SS\hat{\psi} - MSW}{SST + MSW}$$

  - Now, we can adjust these for a two-factor ANOVA, and use <u>partial omega squared</u>

    - The proportion of the variance of the dependent variable that is explained by Factor A:

$$\hat{\omega}_A^2 = \frac{SSA - (dfA)MSWithin}{SSA + (N - dfA)MSWithin} = \frac{dfA(F_A - 1)}{dfA(F_A - 1) + N}$$

    - The proportion of the variance of the dependent variable that is explained by Factor B:

$$\hat{\omega}_B^2 = \frac{SSB - (dfB)MSWithin}{SSB + (N - dfB)MSWithin} = \frac{dfB(F_B - 1)}{dfB(F_B - 1) + N}$$

    - The proportion of the variance of the dependent variable that is explained by Factor A by Factor B interaction:

$$\hat{\omega}_{AB}^2 = \frac{SSAB - (dfAB)MSWithin}{SSAB + (N - dfAB)MSWithin} = \frac{dfAB(F_{AB} - 1)}{dfAB(F_{AB} - 1) + N}$$

    - The proportion of the variance of the dependent variable that is explained by a contrast:

$$\hat{\omega}_\psi^2 = \frac{SS\psi - MSWithin}{SS\psi + (N-1)MSWithin} = \frac{(F_\psi - 1)}{(F_\psi - 1) + N}$$

$\omega^2 = .01$	small effect size
$\omega^2 = .06$	medium effect size
$\omega^2 = .15$	large effect size

- It is possible to calculate an overall omega squared – interpreted as the proportion of the variance of the dependent variable that is explained by all the factors and interactions in the study

$$\hat{\omega}^2 = \frac{SSModel - (dfModel)MSWithin}{SSModel + (N - dfModel)MSWithin}$$

- Remember, if the partial omega squared is calculated to be less than zero, we report partial omega squared to be very small

$$\omega^2 < .01$$

- An example using the lecture comprehension data

Tests of Between-Subjects Effects

Dependent Variable: COMPRE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4125.000[a]	5	825.000	14.838	.000
Intercept	40401.000	1	40401.000	726.637	.000
LECTURE	1194.000	2	597.000	10.737	.000
PRESENT	81.000	1	81.000	1.457	.237
LECTURE * PRESENT	2850.000	2	1425.000	25.629	.000
Error	1668.000	30	55.600		
Total	46194.000	36			
Corrected Total	5793.000	35			

a. R Squared = .712 (Adjusted R Squared = .664)

$$\hat{\omega}^2_{Lecture} = \frac{SSA - (dfA)MSWithin}{SSA + (N - dfA)MSWithin} = \frac{1194 - 2(55.6)}{1194 + (36 - 2)(55.6)} = \frac{1182.8}{3084.4} = .351$$

$$\hat{\omega}^2_{Presentation} = \frac{81 - (1)55.6}{81 + (36 - 1)55.6} = \frac{25.4}{2027} = .013$$

$$\hat{\omega}^2_{Lecture*Presentation} = \frac{2850 - (2)55.6}{2850 + (36 - 2)55.6} = \frac{2738.8}{4740.4} = .578$$

$$\hat{\omega}^2_{Model} = \frac{4125 - (5)55.6}{4125 + (36 - 5)55.6} = \frac{4097}{5848.6} = .658$$

- *f*
  - *f* is a measure of the average standardized difference between the means and the grand mean
  - It can be difficult to interpret and should not be used when more than 2 means are involved

$$f = \sqrt{\frac{\omega^2}{1-\omega^2}}$$

  - If you substitute the appropriate partial omega squared into the formula, you can obtain *f* for Factor A, Factor B and the AB interaction.

- When conducting contrasts, it is much more informative to report Hedges's *g*, or *r*.

$$g = \frac{\hat{\psi}}{\sqrt{MSW}} \qquad \text{where} \qquad \sum|a_i| = 2$$

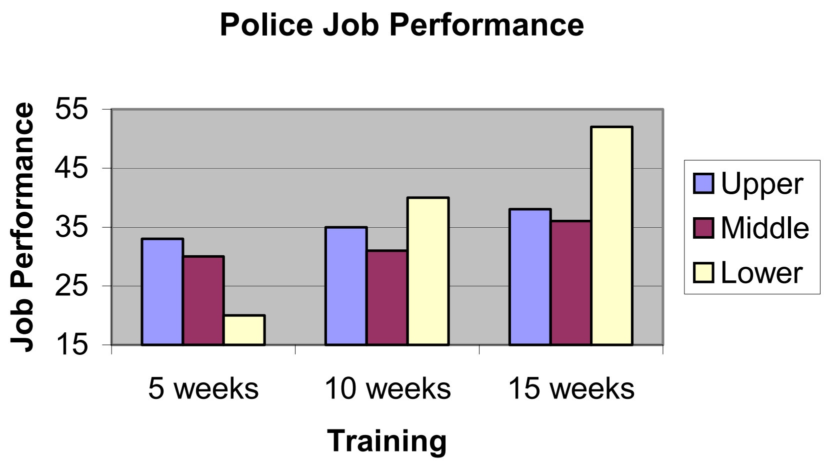
$$r = \sqrt{\frac{F_{contrast}}{F_{contrast} + df_{within}}} = \frac{t_{contrast}}{\sqrt{t^2_{contrast} + df_{within}}}$$

  - For the presentation example, the lecture main effect and the lecture*presentation interaction are omnibus tests. Thus, if we choose to report these tests, we are stuck reporting ω^2.

11. Examples

- Example #1: Let's return to the job performance example and imagine that we had no hypotheses.

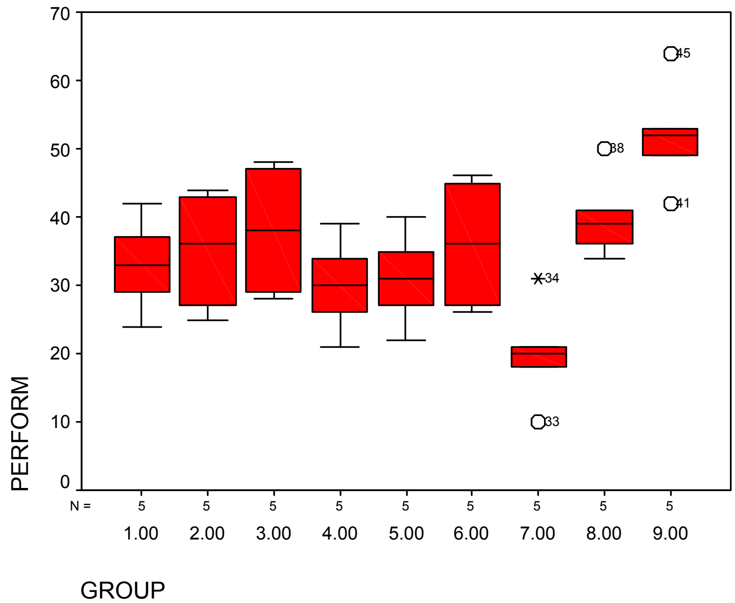
- The only approach to analysis is to use the traditional main effects and interaction approach (see Maxwell and Delaney's flowchart).



  - From the graph, we can see that there appears to be
    - A location by training interaction such that amount of training makes little difference in performance for upper and middle class police, but training does affect performance for lower class police
    - A main effect for training such that as training increases, performance increases (but we should not interpret this!)

- First, let's do a quick check of assumptions (with n = 5, we will not be able to tell much!)

```
EXAMINE VARIABLES=perform BY group
  /PLOT BOXPLOT STEMLEAF NPPLOT SPREADLEVEL.
```



Tests of Normality

	GROUP	Shapiro-Wilk Statistic	df	Sig.
PERFORM	1.00	.995	5	.994
	2.00	.877	5	.297
	3.00	.859	5	.226
	4.00	.995	5	.994
	5.00	.995	5	.994
	6.00	.859	5	.226
	7.00	.954	5	.764
	8.00	.910	5	.470
	9.00	.957	5	.784

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
PERFORM	Based on Mean	.437	8	36	.891

- Again, it is difficult to make judgments based on 5/cell, but nothing looks too out of the ordinary.

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks	
Level 1: Upper Class	$\overline{X}_{\cdot 11} = 33$	$\overline{X}_{\cdot 21} = 35$	$\overline{X}_{\cdot 31} = 38$	$\overline{X}_{\cdot\cdot 1} = 35.33$
Level 2: Middle Class	$\overline{X}_{\cdot 12} = 30$	$\overline{X}_{\cdot 22} = 31$	$\overline{X}_{\cdot 32} = 36$	$\overline{X}_{\cdot\cdot 2} = 32.33$
Level 3: Lower Class	$\overline{X}_{\cdot 13} = 20$	$\overline{X}_{\cdot 23} = 40$	$\overline{X}_{\cdot 33} = 52$	$\overline{X}_{\cdot\cdot 3} = 37.33$
$n_{jk} = 5$	$\overline{X}_{\cdot 1\cdot} = 27.67$	$\overline{X}_{\cdot 2\cdot} = 35.33$	$\overline{X}_{\cdot 3\cdot} = 42$	

$\hat{\mu} = 35$

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	$\hat{\alpha}_1 = -7.33$	$\hat{\alpha}_2 = 0.33$	$\hat{\alpha}_3 = 7.00$
	$\hat{\beta}_1 = 0.33$	$\hat{\beta}_1 = 0.33$	$\hat{\beta}_1 = 0.33$
	$(\hat{\alpha\beta})_{11} = 5$	$(\hat{\alpha\beta})_{21} = -0.67$	$(\hat{\alpha\beta})_{31} = -4.33$
Level 2: Middle Class	$\hat{\alpha}_1 = -7.33$	$\hat{\alpha}_2 = 0.33$	$\hat{\alpha}_3 = 7.00$
	$\hat{\beta}_2 = -2.67$	$\hat{\beta}_2 = -2.67$	$\hat{\beta}_2 = -2.67$
	$(\hat{\alpha\beta})_{12} = 5$	$(\hat{\alpha\beta})_{22} = -1.67$	$(\hat{\alpha\beta})_{32} = -3.33$
Level 3: Lower Class	$\hat{\alpha}_1 = -7.33$	$\hat{\alpha}_2 = 0.33$	$\hat{\alpha}_3 = 7.00$
	$\hat{\beta}_3 = 2.33$	$\hat{\beta}_3 = 2.33$	$\hat{\beta}_3 = 2.33$
	$(\hat{\alpha\beta})_{13} = -10$	$(\hat{\alpha\beta})_{23} = 2.34$	$(\hat{\alpha\beta})_{33} = 7.67$

- Next, we run the tests for main effects and interactions

```
UNIANOVA perform  BY duration location.
```

Tests of Between-Subjects Effects

Dependent Variable: PERFORM

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2970.000[a]	8	371.250	5.940	.000
Intercept	55125.000	1	55125.000	882.000	.000
DURATION	1543.333	2	771.667	12.347	.000
LOCATION	190.000	2	95.000	1.520	.232
DURATION * LOCATION	1236.667	4	309.167	4.947	.003
Error	2250.000	36	62.500		
Total	60345.000	45			
Corrected Total	5220.000	44			

a. R Squared = .569 (Adjusted R Squared = .473)

- The duration by location interaction is significant:

$$F(4,36) = 4.95, p = .003, \omega^2 = .26$$

But this is an omnibus test; we need to do follow-up tests to identify the effect. (But the presence of a significant interaction indicates that we should refrain from interpreting the significant main effect for duration, and instead should proceed to simple effects.)

- Let's examine the simple effect of duration within levels of location (using Bonferroni $\alpha_{FW} = .05/b = .05/3 = .0167$)

    - $b=3$ indicating that the simple effects tests will each be a $3-1=2$ df test.
    - We need to compute two orthogonal contrasts for each simple effect and conduct a simultaneous test of those contrasts.

    - The simple effect of duration for police officers in upper class / middle class / lower class neighborhoods:

IV 2: Location of Office	IV 1: Training Duration Level 1: 5 weeks	Level 2: 10 Weeks	Level 3: 15 Weeks
Level 1: Upper Class	$\overline{X}_{\cdot 11} = 33$	$\overline{X}_{\cdot 21} = 35$	$\overline{X}_{\cdot 31} = 38$
Level 2: Middle Class	$\overline{X}_{\cdot 12} = 30$	$\overline{X}_{\cdot 22} = 31$	$\overline{X}_{\cdot 32} = 36$
Level 3: Lower Class	$\overline{X}_{\cdot 13} = 20$	$\overline{X}_{\cdot 23} = 40$	$\overline{X}_{\cdot 33} = 52$

```
UNIANOVA perform  BY duration location
  /EMMEANS = TABLES(duration*location) COMPARE (duration)
  /PRINT = DESCRIPTIVE .
```

Univariate Tests

Dependent Variable: perform

location		Sum of Squares	df	Mean Square	F	Sig.
Upper Class	Contrast	63.333	2	31.667	.507	.607
	Error	2250.000	36	62.500		
Middle Class	Contrast	103.333	2	51.667	.827	.446
	Error	2250.000	36	62.500		
Lower Class	Contrast	2613.333	2	1306.667	20.907	.000
	Error	2250.000	36	62.500		

Each F tests the simple effects of duration within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

- (Unadjusted) Simple effect of duration within upper class offices:
  $F(2,36) = 0.51, p = .607$
- (Unadjusted) Simple effect of duration within middle class offices:
  $F(2,36) = 0.83, p = .446$
- (Unadjusted) Simple effect of duration within lower class offices:
  $F(2,36) = 20.91, p < .001$

o Now apply Bonferroni correction:

$$p_{crit} = \frac{.05}{3} = .0167$$

- Simple effect of duration within upper class offices:
  $F(2,36) = 0.51, ns$
- Simple effect of duration within middle class offices:
  $F(2,36) = 0.83, ns$
- Simple effect of duration within lower class offices:
  $F(2,36) = 20.91, p < .05$

- We need to perform comparisons of individual cell means to identify the effects (using Tukey $\alpha_{FW} = .05/3 = .0167$ within each simple effect).

$$q_{crit}\left(\alpha = \frac{.05}{3}, 3, 36\right) = 4.11$$

$$t_{crit} = \frac{q_{crit}}{\sqrt{2}} = 2.91$$

- First, let's conduct these contrasts using the ONEWAY command:

```
ONEWAY perform BY group
 /CONT = 1 -1 0 0 0 0 0 0 0
 /CONT = 1 0 –1 0 0 0 0 0 0
 /CONT = 0 1 –1 0 0 0 0 0 0
 /CONT = 0 0 0 1 -1 0 0 0 0
 /CONT = 0 0 0 1 0 –1 0 0 0
 /CONT = 0 0 0 0 1 –1 0 0 0
 /CONT = 0 0 0 0 0 0 1 -1 0
 /CONT = 0 0 0 0 0 0 1 0 -1
 /CONT = 0 0 0 0 0 0 0 1 -1.
```

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
PERFORM	Upper Class	5 vs. 10	-2.0000	5.00000	-.400	36	.692
		5 vs. 15	-5.0000	5.00000	-1.000	36	.324
		10 vs. 15	-3.0000	5.00000	-.600	36	.552

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
PERFORM	Middle Class	5 vs. 10	-1.0000	5.00000	-.200	36	.843
		5 vs. 15	-6.0000	5.00000	-1.200	36	.238
		10 vs. 15	-5.0000	5.00000	-1.000	36	.324

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
PERFORM	Lower Class	5 vs. 10	-20.0000	5.00000	-4.000	36	.000
		5 vs. 15	-32.0000	5.00000	-6.400	36	.000
		10 vs. 15	-12.0000	5.00000	-2.400	36	.022

- Using Tukey's HSD correction, we find no significant pairwise differences in upper and middle class neighborhoods.
- In lower class neighborhoods, we find:
  - Better job performance for those with 10 weeks of training vs 5 weeks of training, $t(36) = 4.00, p < .05, \omega^2 = .25$
  - Better job performance for those with 15 weeks of training vs 5 weeks of training, $t(36) = 6.40, p < .05, \omega^2 = .47$
  - No significant difference in job performance for 10 weeks of training vs 15 weeks of training, $t(36) = 2.40, ns, \omega^2 = .10$

$$\hat{\omega}^2_{\psi 1} = \frac{(F_\psi - 1)}{(F_\psi - 1) + N} = \frac{(16-1)}{(16-1)+45} = .25 \qquad r_{\psi 1} = \sqrt{\frac{F_{contrast}}{F_{contrast} + df_{within}}} = \sqrt{\frac{16}{16+36}} = .56$$

$$\hat{\omega}^2_{\psi 2} = \frac{(40.96-1)}{(40.96-1)+45} = .47 \qquad r_{\psi 2} = \sqrt{\frac{40.96}{40.96+36}} = .73$$

$$\hat{\omega}^2_{\psi 3} = \frac{(5.76-1)}{(5.76-1)+45} = .10 \qquad r_{\psi 3} = \sqrt{\frac{5.76}{5.76+36}} = .37$$

- We could have also conducted these tests with GLM/UNIANOVA

```
UNIANOVA perform  BY duration location
  /EMMEANS = TABLES(duration*location) COMPARE (duration)
```

Pairwise Comparisons

Dependent Variable: perform

location	(I) duration	(J) duration	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
Upper Class	5 Weeks	10 Weeks	-2.000	5.000	.692	-12.140	8.140
		15 Weeks	-5.000	5.000	.324	-15.140	5.140
	10 Weeks	5 Weeks	2.000	5.000	.692	-8.140	12.140
		15 Weeks	-3.000	5.000	.552	-13.140	7.140
	15 Weeks	5 Weeks	5.000	5.000	.324	-5.140	15.140
		10 Weeks	3.000	5.000	.552	-7.140	13.140
Middle Class	5 Weeks	10 Weeks	-1.000	5.000	.843	-11.140	9.140
		15 Weeks	-6.000	5.000	.238	-16.140	4.140
	10 Weeks	5 Weeks	1.000	5.000	.843	-9.140	11.140
		15 Weeks	-5.000	5.000	.324	-15.140	5.140
	15 Weeks	5 Weeks	6.000	5.000	.238	-4.140	16.140
		10 Weeks	5.000	5.000	.324	-5.140	15.140
Lower Class	5 Weeks	10 Weeks	-20.000*	5.000	.000	-30.140	-9.860
		15 Weeks	-32.000*	5.000	.000	-42.140	-21.860
	10 Weeks	5 Weeks	20.000*	5.000	.000	9.860	30.140
		15 Weeks	-12.000*	5.000	.022	-22.140	-1.860
	15 Weeks	5 Weeks	32.000*	5.000	.000	21.860	42.140
		10 Weeks	12.000*	5.000	.022	1.860	22.140

Based on estimated marginal means

*. The mean difference is significant at the .050 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

  - These results exactly match the contrasts we obtained from ONEWAY

- According to Maxwell & Delaney, we should redo these analyses to examine the simple effects of location on duration of training. This is left as an exercise to the reader. These analyses will give you a second way of looking at the same effects.

```
UNIANOVA perform  BY duration location
  /EMMEANS = TABLES(duration*location) COMPARE (location)
```

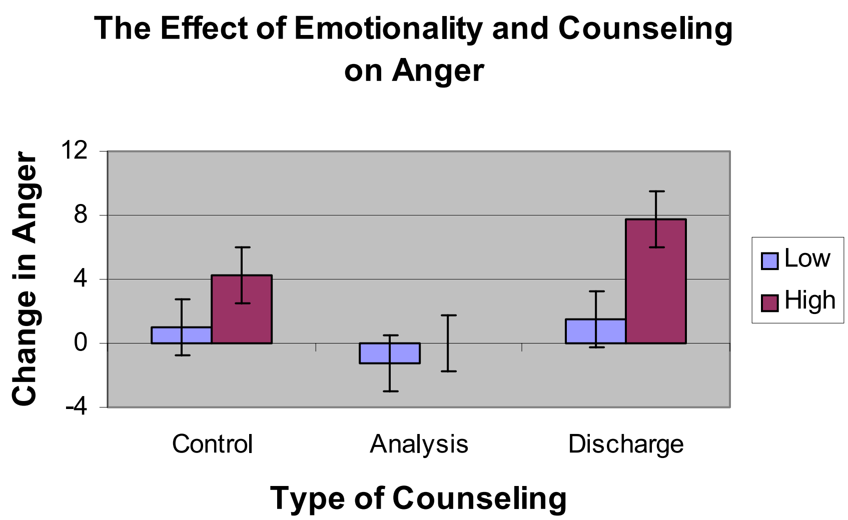

- Finally, always remember to graph your data with error bars/confidence intervals.

Note that the standard error bars = $\sqrt{\frac{MSW}{n_{jk}}}$

- Example #2: The effect of counseling and emotionality on anger

DV = Change in anger scores

	Type of Counseling		
Emotionality	Control	Analysis	Discharge
Low	-1 2 -1 4	-4 -1 2 -2	3 2 1 0
High	5 6 4 2	9 -3 2 8	9 9 7 6



- In this case, we have a prediction:

  Compared to the control group, the discharge group will have higher anger scores, and compared to the control group, the analysis group will have lower anger scores only for those low in emotionality

- We can test the two parts of this prediction separately

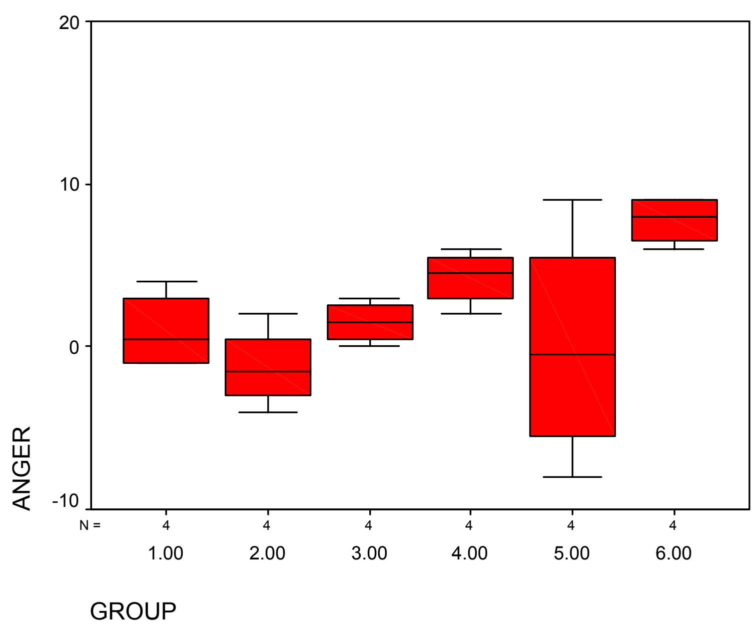
	Type of Counseling		
Emotionality	Control	Analysis	Discharge
Low	-1	0	1
High	-1	0	1

	Type of Counseling		
Emotionality	Control	Analysis	Discharge
Low	1	-3	0
High	1	1	0

- First, let's do a quick check of the assumptions:

```
if (counsel=1 and emotion=1) group = 1.
if (counsel=2 and emotion=1) group = 2.
if (counsel=3 and emotion=1) group = 3.
if (counsel=1 and emotion=2) group = 4.
if (counsel=2 and emotion=2) group = 5.
if (counsel=3 and emotion=2) group = 6.

EXAMINE VARIABLES=anger BY group
  /PLOT BOXPLOT STEMLEAF NPPLOT SPREADLEVEL.
```



Tests of Normality

	GROUP	Shapiro-Wilk Statistic	df	Sig.
ANGER	1.00	.860	4	.262
	2.00	.982	4	.911
	3.00	.993	4	.972
	4.00	.971	4	.850
	5.00	.991	4	.964
	6.00	.849	4	.224

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
ANGER	Based on Mean	4.173	5	18	.011

- We do not have equality of variances. We will have to analyze the data in a manner that does not assume homogeneity of variances

- Because we have specific hypotheses, let's use the contrast method of analyzing the data to directly test those hypotheses

```
ONEWAY anger by group
  /CONT = -1 0 1 -1 0 1
  /CONT =  1 -3 0 1 1 0.
```

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
ANGER	Does not assume equal variances	1	4.0000	1.7912	2.233	9.439	.051
		2	9.0000	5.4276	1.658	6.963	.141

- We find moderate support that the discharge group has higher anger scores than the control group, $t(9.44) = 2.23, p = .051, r \approx .47$

- There is no evidence that the analysis group had significantly lower anger scores than the control group only for those low in emotionality, $t(6.96) = 1.66, p = .14, r = .36$

- After looking at the data, we decide to compare each cell mean to its control mean.

Emotionality	Type of Counseling: Control	Analysis	Discharge
Low	$\overline{X}_{\cdot 11} = 1.00$	$\overline{X}_{\cdot 21} = -1.25$	$\overline{X}_{\cdot 31} = 1.50$
High	$\overline{X}_{\cdot 12} = 4.25$	$\overline{X}_{\cdot 22} = 0.00$	$\overline{X}_{\cdot 32} = 7.75$

Because these comparisons were made after looking at the data, we must use the Tukey correction (technically, the Dunnett T3 correction because the variances are not equal)

```
ONEWAY anger by group
 /CONT = -1 1 0 0 0 0
 /CONT = -1 0 1 0 0 0
 /CONT = 0 0 0 -1 1 0
 /CONT = 0 0 0 -1 0 1.
```

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
ANGER	Does not assume equal variances	1	-2.2500	1.75000	-1.286	5.998	.246
		2	.5000	1.38444	.361	4.547	.734
		3	-4.2500	3.72771	-1.140	3.331	.330
		4	3.5000	1.13652	3.080	5.902	.022

- Low Emotionality: Control vs. Analysis

  $t(6.00) = -1.29$ $\qquad q(1-\alpha, ab, \nu) = q(.95,6,6) \approx 5.63$

  $$t_{crit} \approx \frac{5.63}{\sqrt{2}} = 3.98$$

  Not significant

- Low Emotionality: Control vs. Discharge

  $t(4.55) = 0.36$ $\qquad q(1-\alpha, ab, \nu) = q(.95,6,4.55) \approx 6.71$

  $$t_{crit} \approx \frac{6.71}{\sqrt{2}} = 4.74$$

  Not significant

- High Emotionality: Control vs. Analysis

  $t(3.33) = -1.14$ $\qquad$ $q(1-\alpha, ab, \nu) = q(.95,6,3.33) \approx 8.04$

  $$t_{crit} \approx \frac{8.04}{\sqrt{2}} = 5.68$$

  Not significant

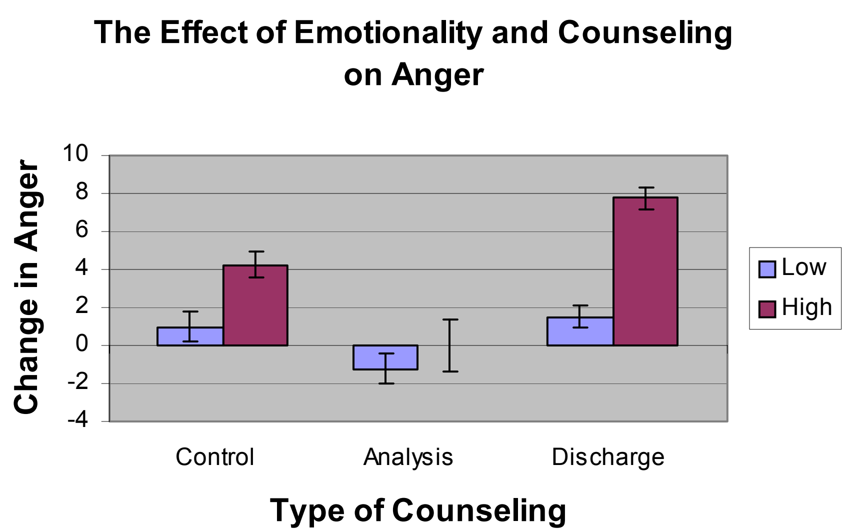
- High Emotionality: Control vs. Discharge

  $t(5.90) = 3.08$ $\qquad$ $q(1-\alpha, ab, \nu) = q(.95,6,5.90) \approx 6.03$

  $$t_{crit} \approx \frac{6.03}{\sqrt{2}} = 4.27$$

  Not significant

- We find no evidence that any of the cell means differ significantly from their control means.



Appendix

A. Conducting main effect and interaction tests using contrasts

- As a way of solidifying what we have learned regarding contrasts, let's apply this knowledge to testing the omnibus main effect and interaction tests using contrasts.

- As we know, for a oneway ANOVA with a levels, we can test the omnibus hypothesis by conducting a simultaneous test of a-1 orthogonal contrasts.

$$F(a-1, dfw) = \frac{\left(\dfrac{SS\hat{\psi}_1 + \ldots + SS\hat{\psi}_{(a-1)}}{a-1}\right)}{MSW}$$

- For a two-way ANOVA, we can follow a similar logic:
    - Test for the main effect of IV1 (a levels): simultaneous test of a-1 orthogonal contrasts <u>on the marginal means for IV1</u>

    - Test for the main effect of IV2 (b levels): simultaneous test of b-1 orthogonal contrasts <u>on the marginal means for IV2</u>

    - Test for IV 1 by IV 2 interaction: simultaneous test of $(a\text{-}1)(b\text{-}1)$ orthogonal contrasts <u>on the cell means</u>

- A 2x2 example: SBP example

		Diet Modification		
		No	Yes	
Drug Therapy	No	$\bar{X}_{\cdot 11} = 190$	$\bar{X}_{\cdot 21} = 188$	$\bar{X}_{\cdot\cdot 1} = 189$
	Yes	$\bar{X}_{\cdot 12} = 171$	$\bar{X}_{\cdot 22} = 167$	$\bar{X}_{\cdot\cdot 2} = 169$
		$\bar{X}_{\cdot 1 \cdot} = 180.5$	$\bar{X}_{\cdot 2 \cdot} = 177.5$	$\bar{X}_{\cdots} = 179$

$n_{jk} = 5$

- First, let's let SPSS do all the work for us:

  UNIANOVA sbp BY drug diet.

Tests of Between-Subjects Effects

Dependent Variable: SBP

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2050.000[a]	3	683.333	9.762	.001
Intercept	640820.000	1	640820.000	9154.571	.000
DRUG	2000.000	1	2000.000	28.571	.000
DIET	45.000	1	45.000	.643	.434
DRUG * DIET	5.000	1	5.000	.071	.793
Error	1120.000	16	70.000		
Total	643990.000	20			
Corrected Total	3170.000	19			

a. R Squared = .647 (Adjusted R Squared = .580)

Main effect for diet: $F(1,16) = 0.64, p = .43$
Main effect for drug: $F(1,16) = 28.57, p < .01$
Diet by drug interaction: $F(1,16) = 0.07, p = .79$

- Now, let's replicate these results using contrasts

- Test for Main Effect of Diet modification (a=2):
  Only 1 contrast is required (Main effect for diet modification has 1 df)

		Diet Modification	
		No	Yes
Drug Therapy	No		
	Yes		
		-1	1

Contrast on Marginal Means for Diet Modification

When we have equal *n*, we can conduct a test of marginal means on the cell means

		Diet Modification	
		No	Yes
Drug Therapy	No	-1	1
	Yes	-1	1

Contrast performed on Cell Means

These contrast coefficients will only work if we have equal *n* (Why?)

To test this contrast, we can:

- Compute the contrast by hand
  (using either the marginal means or the cell means)
- Trick SPSS into thinking this is a oneway design and use the ONEWAY command

		Diet Modification	
		No	Yes
Drug Therapy	No	Cell 1	Cell 2
	Yes	Cell 3	Cell 4

```
if (diet=1 and drug=1) group = 1.
if (diet=2 and drug=1) group = 2.
if (diet=1 and drug=2) group = 3.
if (diet=2 and drug=2) group = 4.
```

```
ONEWAY sbp BY group.
 /cont = -1 1 -1 1
```

Contrast Tests

Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
SBP	-6.0000	7.48331	-.802	16	.434

Main effect for Diet: $t(16) = -0.80, p = .43$

$F(1,16) = 0.64, p = .43$

- Test for Main Effect of Drug Therapy modification (b=2):
  Only 1 contrast is required (Main effect for drug therapy has 1 df)

		Diet Modification		
		No	Yes	
Drug Therapy	No			-1
	Yes			1

Contrast on Marginal Means for Drug Therapy

		Diet Modification	
		No	Yes
Drug Therapy	No	-1	-1
	Yes	1	1

Contrast performed on Cell Means

```
ONEWAY sbp BY group
 /cont = -1 -1 1 1.
```

Contrast Tests

Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
SBP	-40.0000	7.48331	-5.345	16	.000

Main effect for Diet: $t(16) = 5.35, p < .01$

$F(1,16) = 28.57, p < .01$

- Test for Main Effect of Diet by Drug interaction:
  Only 1 contrast is required (Diet by drug interaction has 1 df)

  What contrast coefficients should we use?
  Orthogonal interaction contrasts can be obtained by multiplying the marginal main effect contrasts

		Diet Modification: No	Diet Modification: Yes	
Drug Therapy	No			**-1**
	Yes			**1**
		-1	**1**	

$n_{jk} = 5$

		Diet Modification: No	Diet Modification: Yes
Drug Therapy	No	1	-1
	Yes	-1	1

$n_{jk} = 5$

```
ONEWAY sbp BY group
 /cont = 1 -1 -1 1.
```

Contrast Tests

Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
SBP	-2.0000	7.48331	-.267	16	.793

Diet by drug interaction: $t(16) = 0.27, p = .79$
$F(1,16) = 0.07, p = .79$

- Let's look at the set of contrasts we have used:

$$c_A = (-1,1,-1,1) \qquad c_A \perp c_B$$
$$c_B = (-1,-1,1,1) \qquad c_B \perp c_{A*B}$$
$$c_{A*B} = (1,-1,-1,1) \qquad c_A \perp c_{A*B}$$

  This is an orthogonal set of contrasts, and because these contrasts are orthogonal, the ANOVA *SS* partition works!

- Things get more complicated when the omnibus tests have more than 1 df, but the same logic applies

- Finally, it is important to remember what we learned about omnibus tests – they rarely address your research hypothesis. It is almost always preferable to skip the omnibus tests and use contrasts to directly examine your hypotheses.

