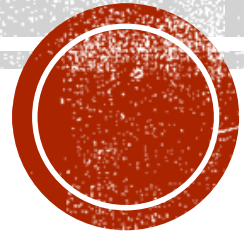
Chapter 14

FACTORIAL ANOVA



'People can be divided into two classes:
Those who go ahead and do something,
and those who sit still and inquire,
'Why wasn't it done the other way?"

– Oliver Wendell Holmes, American Physician, Writer,
Humorist, Harvard Professor, 1809-1894

Dr. Petrov is interested in conducting an experiment where:

- *30 high school students are randomly assigned to a new **computer simulation tool** for learning geometry and*

- *30 other students are randomly assigned to the standard **lecture and paper/pencil problem** solving format.*

*However, Dr. Petrov is **also** interested in the **effect of sex** differences on learning outcomes.*

ANALYSIS OF VARIANCE

- ANOVA types…
  - 1-Way ANOVA = 1 factor
  - 2-Way ANOVA = 2 factors (focus of lecture)
  - 3-Way ANOVA = 3 factors
  - 4-Way ANOVA = 4 factors

- # levels of each factor determines ANOVA design
  - # Levels: Row factor = 2, Column factor = 3
    - 2-way ANOVA, 2X3 factorial design
  - # Levels: Row factor = 4, Column factor = 3
    - 2-way ANOVA, 4X3 factorial design

FACTORIAL 2-WAY ANOVA

- Simultaneously evaluate effect of 2 or more factors on continuous outcome
  - Cross-classification
  - Participants only belong to 1 mutually exclusive 'cell'
    - Within 1 level of row factor and 1 level of columns factor

- Typical 2-way ANOVA
  - 3x2 design
  - Row factor (A): 3 levels
  - Column factor (B): 2 levels

		B	
		B1	**B2**
A	**A1**	11	12
	A2	21	22
	A3	31	32

TEST OF ROW MAIN EFFECT

- Do row marginal means differ?
  - Do population means differ across levels of row factor, averaging across levels of column factor?
  - H_0: $\mu_{j1} = \mu_{j2} = \mu_{jr}$
  - H_1: Not H_0

		B		
		B1	B2	Marginals
A	A1	M_{11}	M_{12}	M_{A1}
	A2	M_{21}	M_{22}	M_{A2}
	A3	M_{31}	M_{32}	M_{A3}
	Marginals	M_{B1}	M_{B2}	

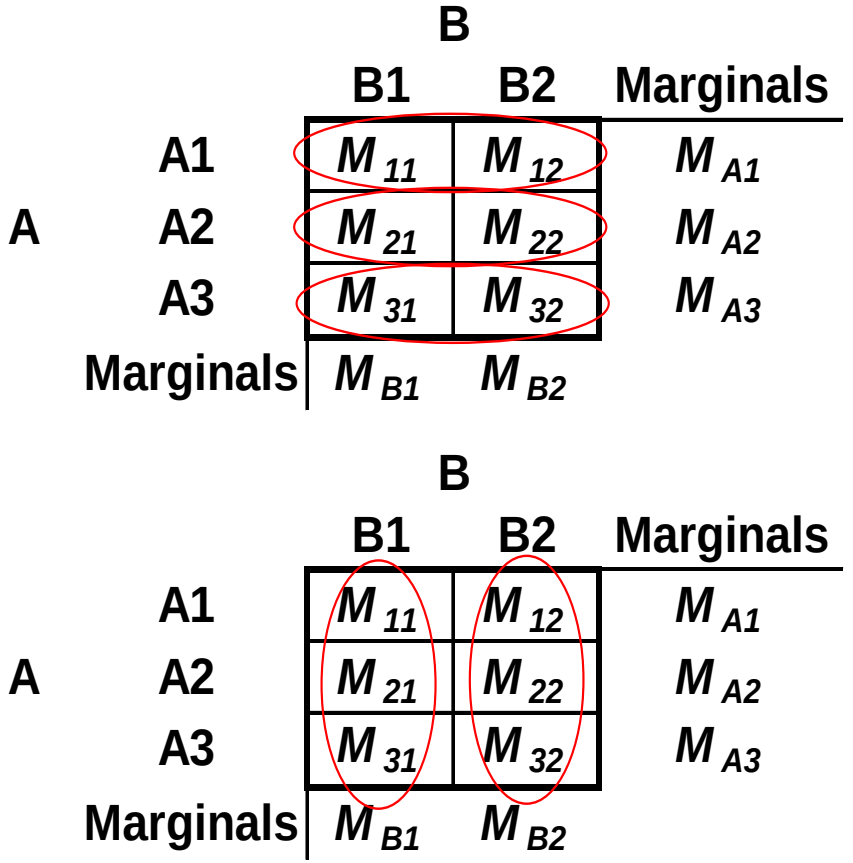
TEST OF COLUMN MAIN EFFECT

- Do column marginal means differ?
  - Do population means differ across levels of column factor, averaging over row factor levels?
  - H_0: $\mu_{k1} = \mu_{k2} = \mu_{kc}$
  - H_1: Not H_0

		B		
		B1	B2	Marginals
A	A1	M_{11}	M_{12}	M_{A1}
	A2	M_{21}	M_{22}	M_{A2}
	A3	M_{31}	M_{32}	M_{A3}
	Marginals	M_{B1}	M_{B2}	

TEST OF INTERACTION EFFECT

- Does pattern of cell means differ?
  - Are differences among population means across row factor similar across all levels of column factor (and vice versa)?
  - H_0: Differences among levels for 1 factor do not vary across levels of other factor
  - H_1: Not H_0



OUTCOMES

- No significant main effects or interaction(s)

- No significant interaction
  - Significant main effect for rows, but not for columns
  - Significant main effect for columns, but not for rows
  - Significant main effects for both rows and columns

- Significant interaction and…
  - Non-significant main effects for rows or columns
  - Significant main effect for rows, but not for columns
  - Significant main effect for columns, but not for rows
  - Significant main effects for both rows and columns

REDUCED ERROR

- Subject-to-subject variability contributes to increased MS_W
  - Result: Less power

- Adding factors to analysis that explain subject-to-subject variability in outcome reduces MS_W and increases power of analysis
  - Variance within (and thus across) individual cells is reduced as cases become more homogeneous in terms of their characteristics
- Factors that do not have this effect may slightly decrease power
  - df_W (which = $N - rc$) decreases as # cells increases, increasing MS_W and decreasing F-ratios

- Alternatives
  - Restriction (subjects from 1-level only – reduced generalizability)
  - Repeated-measures (matched) designs

ASSUMPTIONS

Similar to 1-Way ANOVA

- Independence
- Outcome is normally distributed in population
- Homogeneity of variance
  - Variances within each cell are equal

VARIANCE COMPONENTS

- SS_{Total} partitioned into 4 components

$$SS_{Total} = SS_{(R)ows} + SS_{(C)olumns} + SS_{RC} + SS_{Within}$$

- When balanced, previous SS_B from 1-Way ANOVA partitioned into 3 components: *R, C, RC*
- 1-way ANOVA uses groups and factorial ANOVA uses cells to compute *SS*
- Following equations are for balanced designs

SS_R

- In computing row means all scores in a given row are averaged regardless of column
  - n_{row} = # participants per row

$$SS_R = n_{row}[(\bar{X}_{row1} - \bar{X}_{GM})^2 + (\bar{X}_{row2} - \bar{X}_{GM})^2 + \ldots + (\bar{X}_{row\ r} - \bar{X}_{GM})^2]$$

$$SS_R = \frac{\left(\sum_{i=1}^{n} X_{row1}\right)^2 + \left(\sum_{i=1}^{n} X_{row2}\right)^2 + \ldots + \left(\sum_{i=1}^{n} X_{row\ r}\right)^2}{n_{row}} - \frac{\left(\sum_{i=1}^{n} X\right)^2}{N}$$

SS_C

- In computing column means all scores in a given column are averaged regardless of row
  - n_{col} = # participants per column

$$SS_C = n_{column}[(\bar{X}_{col1} - \bar{X}_{GM})^2 + (\bar{X}_{col2} - \bar{X}_{GM})^2 + \ldots + (\bar{X}_{col\ r} - \bar{X}_{GM})^2]$$

$$SS_C = \frac{\left(\sum_{i=1}^{n} X_{col1}\right)^2 + \left(\sum_{i=1}^{n} X_{col2}\right)^2 + \ldots + \left(\sum_{i=1}^{n} X_{col\ r}\right)^2}{n_{col}} - \frac{\left(\sum_{i=1}^{n} X\right)^2}{N}$$

SS_{RC}

- Variability among cell means when variability due to individual row and column effects have been removed

$$SS_{RC} = n_{cell}[(\bar{X}_{cell11} - \bar{X}_{GM})^2 + (\bar{X}_{cell12} - \bar{X}_{GM})^2 + \ldots + (\bar{X}_{cell\ rc} - \bar{X}_{GM})^2] - SS_R - SS_C$$

$$SS_{RC} = \frac{\left(\sum_{i=1}^{n} X_{cell11}\right)^2 + \left(\sum_{i=1}^{n} X_{cell12}\right)^2 + \ldots + \left(\sum_{i=1}^{n} X_{cell\ rc}\right)^2}{n_{cell}} - \frac{\left(\sum_{i=1}^{n} X\right)^2}{N} - SS_R - SS_C$$

SS_W

- *SS* <u>within</u> each cell added together
  - $SS_W = SS_{11} + SS_{12} + \ldots + SS_{rc}$

For each cell, all scores within that cell are subtracted from cell mean, squared, and summed

$$SS_W = \sum_{j=1,1}^{rc}\left(\sum_{i=1}^{n_{rc}}\left(X_{irc} - \bar{X}_{rc}\right)^2\right)$$

$$SS_W = \sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_{cell11}\right)^2 + \left(\sum_{i=1}^{n} X_{cell12}\right)^2 + \ldots + \left(\sum_{i=1}^{n} X_{cell\ rc}\right)^2}{n_{cell}}$$

$$SS_W = SS_T - SS_R - SS_C - SS_{RC}$$

DEGREES OF FREEDOM

$$df_{Total} = N_T - 1$$

- **Partitioned into 4 parts**
  - $df_{Total} = df_R + df_C + df_{RC} + df_W$
    - $df_R = r - 1$
    - $df_C = c - 1$
    - $df_{RC} = (r - 1)(c - 1)$
    - $df_W = (N - rc)$
      - **Assumes** n are same for all cells
      - Otherwise, $\Sigma(n_{rc} - 1)$: sum of $n - 1$ per cell

VARIANCE ESTIMATES

- Obtain 4 variance estimates when each variance component is divided by its *df*
  - MS_R = *Row* variance estimate
    - Sensitive to effects of factor A
  - MS_C = *Column* variance estimate
    - Sensitive to effects of factor B
  - MS_{RC} = *Row x Column* variance estimate
    - Sensitive to interaction effects of A and B
  - MS_W = Within-cells variance estimate
    - Not sensitive to effects of any factor

VARIANCE ESTIMATES

Alternative calculations

- SS_R = *POPULATION* variance of row means x N_{Total}
  - $MS_R = SS_R / df_R$
- SS_C = *POPULATION* variance of column means x N_{Total}
  - $MS_C = SS_C / df_C$
- SS_{RC} = Subtract SS_R and SS_C from $SS_{Between}$
  - $SS_{Between}$ = *POPULATION* variance of cell means*N_{Total}
  - $MS_{RC} = SS_{RC} / df_{RC}$
- MS_W = Average of all cell *SAMPLE* variances

F-STATISTICS

- Significance testing of 3 variance estimates
  - Distinct F_{stat} for each
    - MS_R / MS_{Within} : Factor A
    - MS_C / MS_{Within} : Factor B
    - MS_{RC} / MS_{Within} : Interaction between factors A and B

- Each F_{stat} compared to distinct F_{crit}
  - Based on df_{Effect} (e.g., df_R) and df_{Within}
  - Reject H_0: $F_{stat} > F_{crit}$

SUMMARY TABLE

Source	*SS*	*df*	*MS*	*F*	*p*
Row					
Column					
R x C					
Within				**X**	**X**
Total			**X**	**X**	**X**

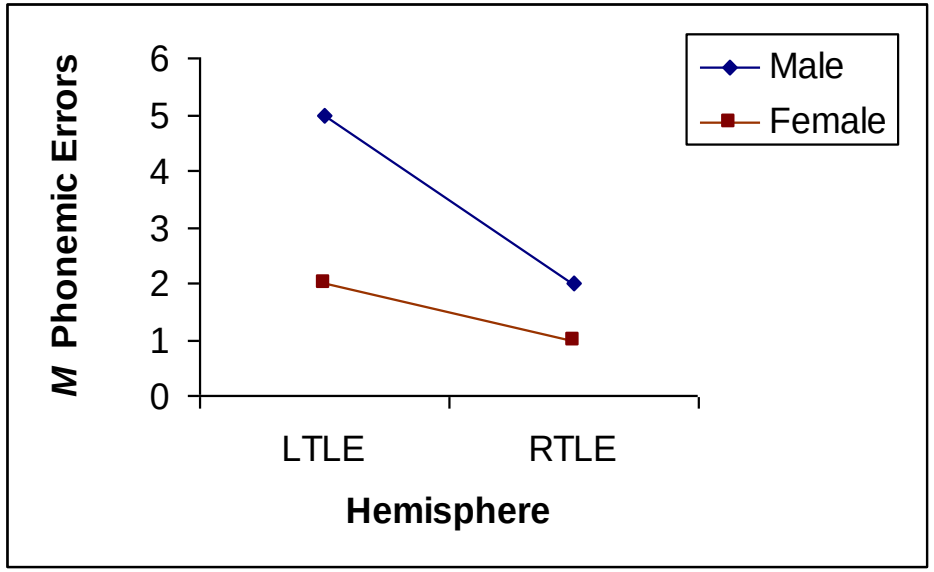
INTERACTIONS

- Interaction between
  - 2 factors: 2-way interaction
  - 3 factors: 3-way interaction
    - Quite rare, be skeptical
- Significance indicates that the effect of 1 factor is not same at all levels of another factor
  - Effect of variables combined is different than would be predicted by either variable alone
- Most interesting results, but more difficult to explain or interpret than main effects

INTERACTIONS

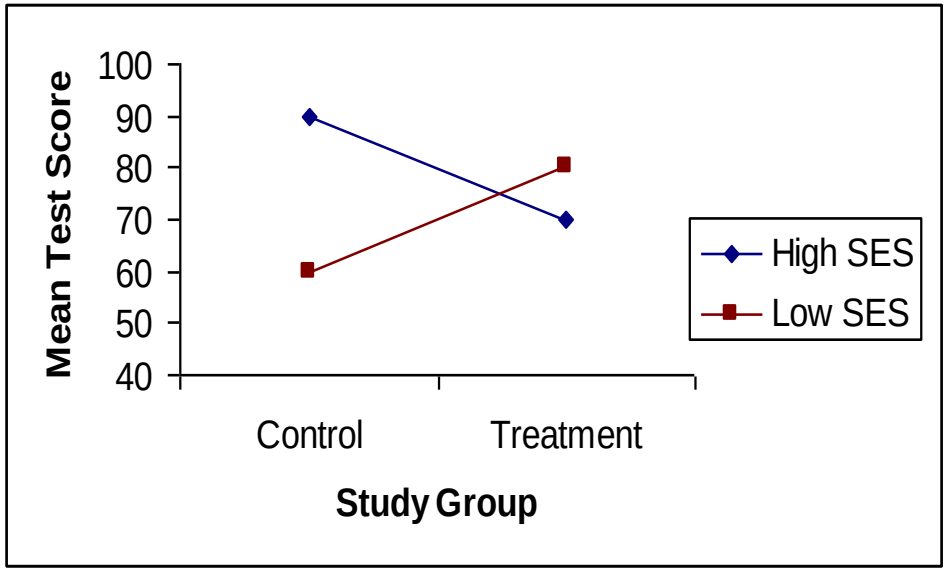
Ordinal

- Direction or order of effects is similar for different subgroups



Disordinal

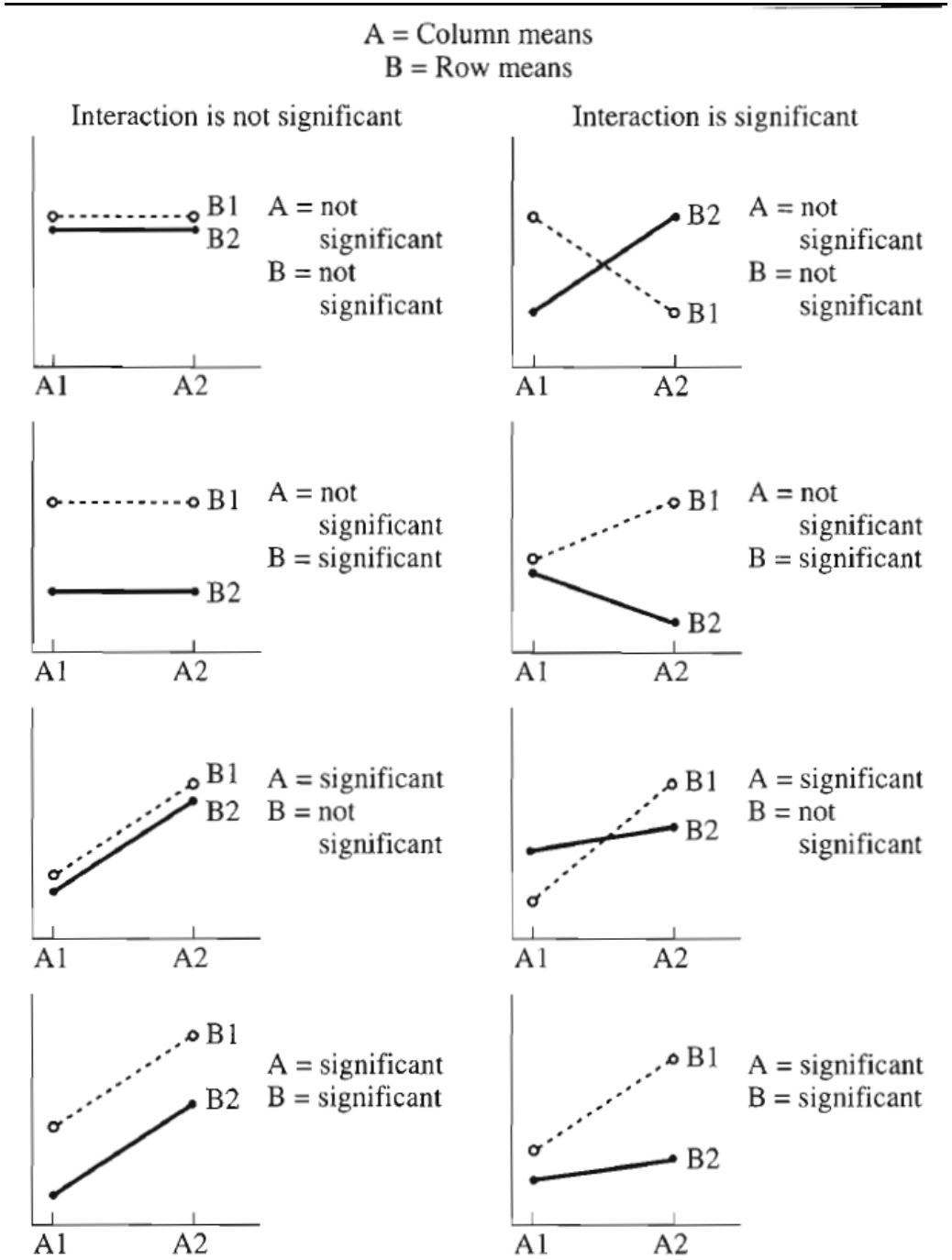
- Direction or order of effects is reversed for different subgroups



INTERACTIONS

- Significance of interaction always evaluated 1st
  - If significant, interpret interaction, not main effects
  - If non-significant, interpret main effects
- Once we know effects of 1 factor are tempered by or contingent on levels of another factor (as in an interaction), interpretation of either factor (main effect) alone is problematic
- Best interpreted through visualization
  - Cell means plot
  - Interactions exist if lines cross or will cross (non-parallel)
- Design graph to best illustrate
  - Outcome on y-axis
  - Select factor for x-axis
  - Other factor(s) represented by separate lines
    - Selection guides interpretation, can dictate whether plot is ordinal/disordinal

Jamison Fargo, PhD



INTERACTIONS

- Some recommend only interpreting significant main effects (Keppel & Wickens, 2004) …
  - When there is no significant interaction
  - (Cautiously) when there is a significant interaction, but 1) interaction effect size is small relative to that of main effects and 2) there is an ordinal pattern to the means

- However, must report all main and interaction effects regardless of statistically significance

NEED FOR TESTING INTERACTIONS

- Results may be distorted if additional factors are not included in analysis so that interactions are not tested
  - E.g., If experimental effects of a drug had opposite effects in men and women, the variable representing drug effects may appear to be ineffective (non-significant main effect) without including the variable for sex differences

- If interaction terms are non-significant, increased confidence that effect of key factor (e.g., drug treatment) is generalizable to all levels of other factors (e.g., sex)

EXAMPLE FROM TEXT

- Effect of sleep deprivation and compensating stimulation on performance of complex motor task
  - Outcome: Video game score simulating driving truck at night
  - Factor A (Row): Sleep deprivation
    1. Control: Normal sleep schedule
    2. Jet lag: Normal sleep amount, but during different hours
    3. Interrupted: Normal sleep amount, but only for 2 hours at a time
    4. Total Deprivation: No sleep for 4 days
  - Factor B (Column): Stimulation conditions
    1. Placebo: Told they are given a stimulant pill (really placebo)
    2. Caffeine: Told they are given a stimulant pill (really stimulant)
    3. Reward: Given mild electric shocks for mistakes and given a monetary reward for good performance

EXAMPLE

- H_0 Deprivation
  - $\mu_{control} = \mu_{jetlag} = \mu_{interrupted} = \mu_{deprive}$
- H_0 Stimulus
  - $\mu_{placebo} = \mu_{caffeine} = \mu_{reward}$
- H_0 Interaction
  - Effect of two factors is additive (no multiplicative or interaction effect)
  - Effect of 1 factor does NOT depend on level of other factor

EXAMPLE

		Stimulus Type		
		Placebo	Caffeine	Reward
Deprivation Type	Control	24	26	28
		20	22	23
		29	20	24
		20	30	30
		28	27	33
	Jet Lag	22	25	26
		18	31	20
		16	24	32
		25	27	23
		27	21	30
	Interrupt	16	23	16
		20	28	13
		11	26	12
		19	17	18
		14	19	19
	Total Deprivation	14	23	15
		17	16	11
		12	26	19
		18	18	11
		10	24	17

Score	Stimulus	Stimulus	Deprivation	Deprivation
24	1	Placebo	1	Control
20	1	Placebo	1	Control
29	1	Placebo	1	Control
20	1	Placebo	1	Control
28	1	Placebo	1	Control
22	1	Placebo	2	Jet Lag
18	1	Placebo	2	Jet Lag
16	1	Placebo	2	Jet Lag
25	1	Placebo	2	Jet Lag
27	1	Placebo	2	Jet Lag
16	1	Placebo	3	Interrupt
20	1	Placebo	3	Interrupt
11	1	Placebo	3	Interrupt
19	1	Placebo	3	Interrupt
14	1	Placebo	3	Interrupt
14	1	Placebo	4	Total Dep
17	1	Placebo	4	Total Dep
12	1	Placebo	4	Total Dep
18	1	Placebo	4	Total Dep
10	1	Placebo	4	Total Dep
26	2	Caffeine	1	Control
22	2	Caffeine	1	Control
20	2	Caffeine	1	Control
30	2	Caffeine	1	Control
27	2	Caffeine	1	Control

EFFECT SIZE

- Proportion of variation in outcome accounted for by a particular factor or interaction term

- **Eta-squared (η^2)**
  - **1-way ANOVA**
    - $SS_{Between} / SS_{Total}$
  - **2-way ANOVA**
    - **Row factor:** SS_R / SS_{Total}
    - **Column factor:** SS_C / SS_{Total}
    - **Interaction:** SS_{RC} / SS_{Total}

- Interpretation:
  - Range: 0 to 1
    - Small: .01 to .06
    - Medium: .06 to .14
    - Large: > .14

EFFECT SIZE

- η^2 are biased parameter estimates

- Should estimate omega squared (ω^2)
  - Substitute *SS* and *df* values

$$\omega^2 = \frac{SS_{Effect} - df_{Effect} * MS_{Within}}{SS_{Total} + MS_{Within}}$$

  - Same interpretation as η^2

EFFECT SIZE

- When <u>all</u> factors are experimental or when <u>many</u> factors are included in analysis, SS due to a factor or interaction will be small relative to SS_{Total}
- Partial effect size estimates are often reported
  - Proportion of variation in outcome accounted for by a particular factor or interaction term, excluding other main effects or interaction sources of variation

$$\text{Partial } \eta^2 = \frac{SS_{Effect}}{SS_{Effect} + SS_{Within}}$$

$$\text{Partial } \omega^2 = \frac{df_{Effect}(MS_{Effect} - MS_{Within})/N}{[df_{Effect}(MS_{Effect} - MS_{Within})/N] + MS_{Within}}$$

MULTIPLE COMPARISONS

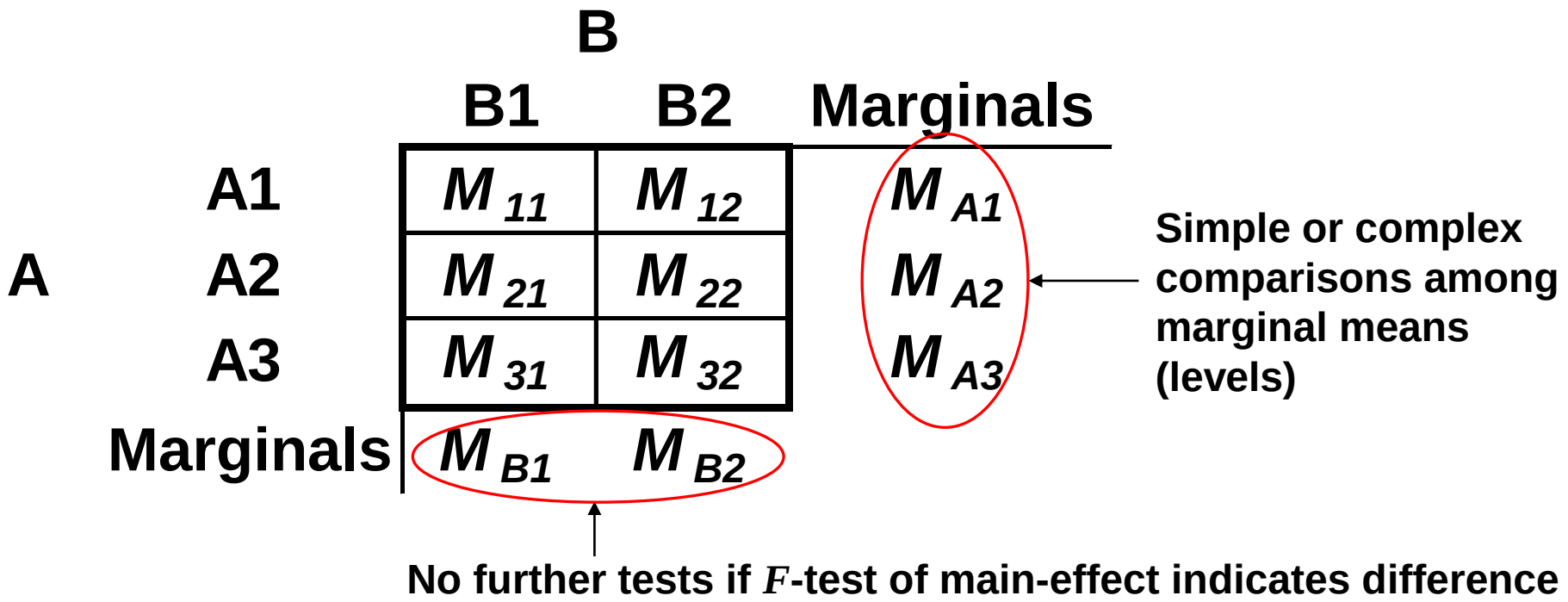
- Factorial ANOVA produces omnibus results
  - No indication of specific level (group) differences within or across factor(s)
- Multiple comparisons elucidate differences within significant main effects or interactions
- Pattern of results dictates approach
  - E.g., Significant main effects, but no interaction
- Each of the 3 F-tests in a 2-Way ANOVA represents a 'planned comparison'
  - No adjustment to α_{EW} necessary
- However, within each main-effect and interaction a separate family of possible multiple comparisons may be conducted
  - α_{EW} must be controlled within each 'family'

NON-SIGNIFICANT INTERACTION

- Evaluation of significant main effect(s)
  - Factors with 2 levels
    - No multiple comparisons required
  - Factors with > 2 levels
    - 2-way ANOVA is reduced to two 1-Way ANOVAs
    - Simple (pairwise) or complex (linear) contrasts are computed within individual significant main-effect(s) (ignoring others)

NON-SIGNIFICANT INTERACTION

Significant main-effects



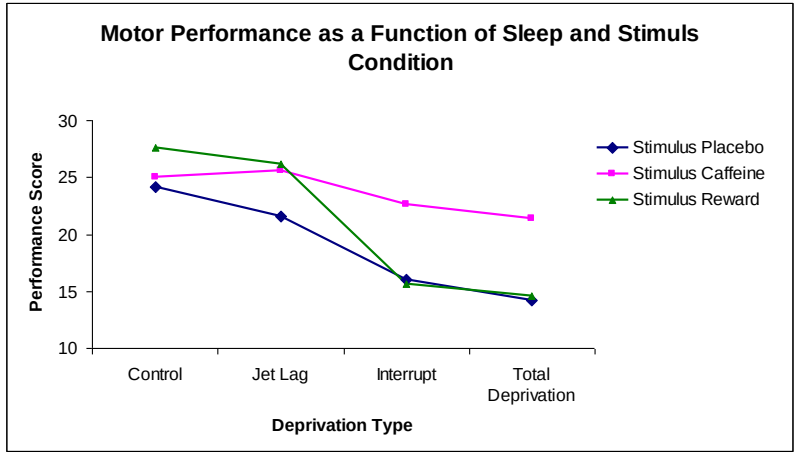
EXAMPLE 1: NON-SIGNIFICANT INTERACTION

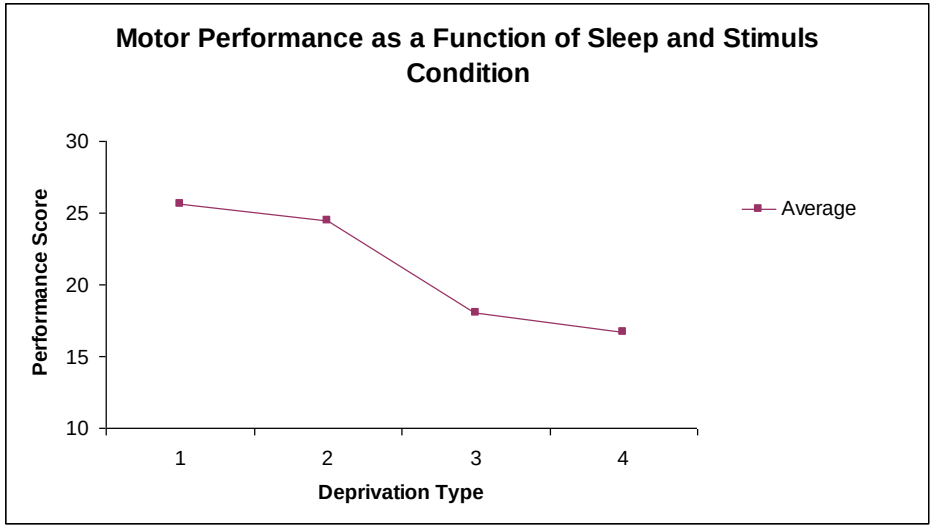
- Sleep deprivation, stimulant, and motor performance example

```
Anova Table (Type II tests)
Response: score
             Sum Sq Df   F value    Pr(>F)
deprivation   897.0  3   18.2406 4.896e-08 ***
Stimulus       217.6  2    6.6385  0.002849 **
Interaction   194.8  6    1.9803  0.087003 .
Residuals     786.8 48
```

- Non-significant interaction
- Both main-effects are significant
- Need to compare 'marginal means' for differences among levels

EXAMPLE 1: NON-SIGNIFICANT INTERACTION





- Figure on right indicates main effect for deprivation type collapsing across levels of stimulant type
  - 'Average' of simple (main) effects (lines in figure)
- When interaction is tested it is really a test of the H_0 that all simple effects are simialr

EXAMPLE 1: NON-SIGNIFICANT INTERACTION

- Run 1-Way ANOVA on main-effects deemed significant in 2-Way ANOVA
  - Optional

- Run multiple comparisons, controlling α_{EW} within each contrast
  - Pairwise: Tukey, Bonferroni
  - Linear contrasts: Contr.helmert

EXAMPLE 1: NON-SIGNIFICANT INTERACTION

- Conduct 1-Way ANOVA in R as before, select pairwise comparisons for Tukey tests

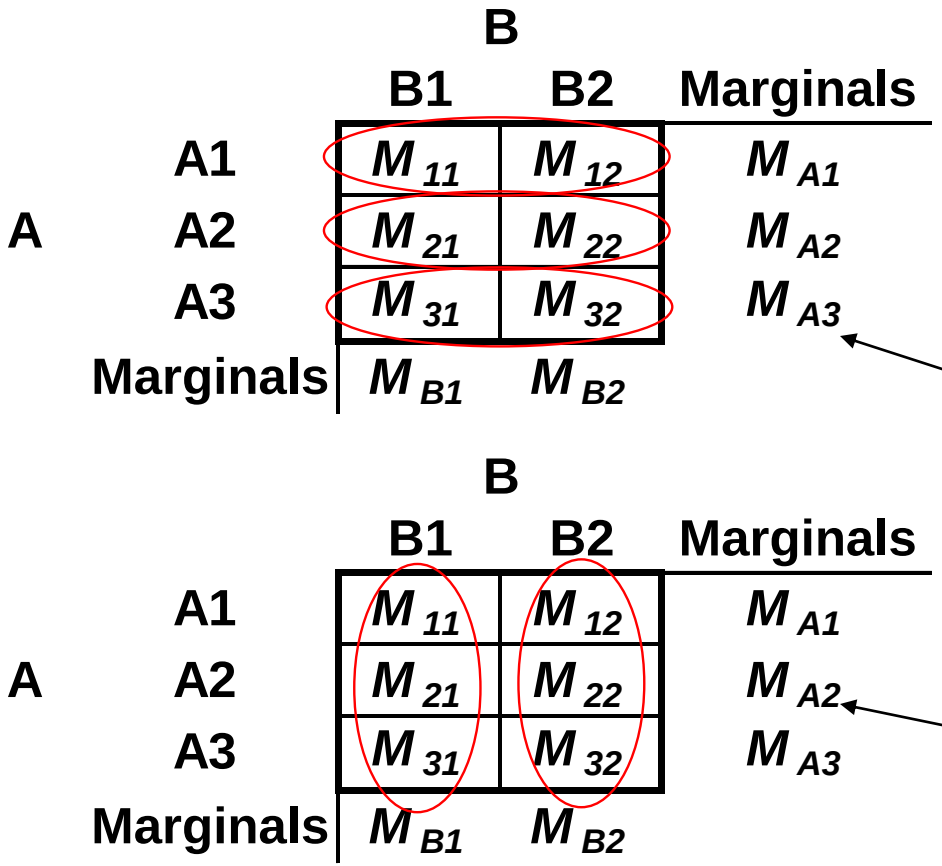
- Alternative 'by hand'; *p*-values close, not exactly the same
  - NOTE: 1-Way ANOVA object must be created using `aov` command

```
TukeyHSD(model.object, "dep_F", ordered = F)
TukeyHSD(model.object, "stim_F", ordered = F)
plot(TukeyHSD(fm1, "dep_f"))
plot(TukeyHSD(fm1, "stim_f"))
```

SIGNIFICANT INTERACTION

- Simple (main) effects of interaction are tested
- One factor is selected as stratifying factor
  - Similar to deciding which factor to put on x-axis in means plot
  - Let theory and research questions guide selection
  - Levels (cells) of other factor are compared within each level of stratified factor
  - Can redo analysis by reversing which factor is stratified and which is examined
- Comparing cell, rather than marginal, means

SIGNIFICANT INTERACTION



- Simple main effects generally tested within each level of stratifying factor
  - 2-levels
    - Simple, pairwise comparisons: Tukey HSD or *t*-tests with Bonferroni correction
  - > 2 levels
    - Modified 1-way ANOVA followed by simple or complex comparisons

SIGNIFICANT INTERACTION

- Modified 1-Way ANOVA tests of simple main effects often done 'by hand'
  - Obtain $MS_{Between}$ from standard 1-Way ANOVA
    - Comparing means across 1 level of 1 factor within 1 level of another factor
  - Obtain MS_{Within} from original 2-Way ANOVA
    - Ensure homogeneity of variance assumption is reasonably satisfied

$$F_{Simple\ Effect} = \frac{MS_{Between\ (1\text{-way ANOVA})}}{MS_{Within\ (2\text{-way ANOVA})}}$$

$$F_{Critical}(df_{Between}, df_{Within}) \text{ from 1-way ANOVA}$$

UNBALANCED DESIGNS

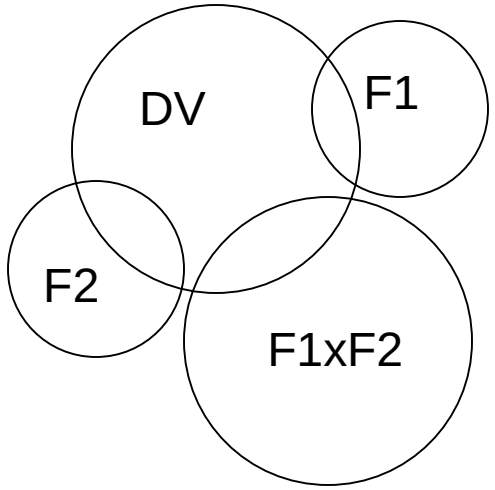
- Equal *n*s in each cell = Orthogonal design
  - Factors are independent/uncorrelated so that significance of any effect is independent of significance of other effects (including interaction)

- Most research consists of unbalanced data
  - As *n*s across cells become more unequal, factors become more dependent/correlated
    - Unbalanced: $SS_{Between} \neq SS_R + SS_C + SS_{RC}$
  - More difficult to determine independent effects of each factor

- Previous equations and R commands will not work correctly for unbalanced designs

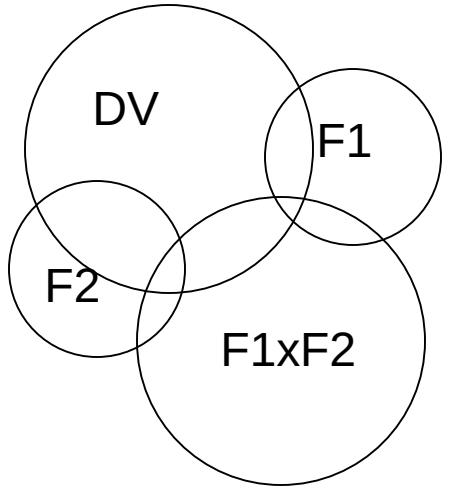
UNBALANCED DESIGNS

- Balanced



- Sum of areas where factors overlap with DV = SS_B
  - Remaining portion of DV = SS_W

- Unbalanced



- Sum of areas where factors overlap with DV ≠ SS_B
  - Some areas counted twice
  - Remaining portion of DV = SS_W

1. Equal cell sizes			
	Factor A		
Factor B	a1	a2	Row Marginal Means
b1	M = 100 n = 50	M = 150 n = 50	M = 125 n = 100
b2	M = 200 n = 50	M = 250 n = 50	M = 225 n = 100
Column Marginal Means	M = 150 n = 100	M = 200 n = 100	
2. Unequal cell sizes			
	Factor A		
Factor B	a1	a2	Row Marginal Means
b1	M = 100 n = 10	M = 150 n = 90	M = 145 n = 100
b2	M = 200 n = 90	M = 250 n = 10	M = 205 n = 100
Column Marginal Means	M = 190 n = 100	M = 160 n = 100	

Individual cell means and marginal *n*s are the same across both tables. Main effects (marginal means) differ across tables as a function of different cell *n*s. Conclusions from ANOVA may vastly differ.

UNBALANCED DESIGNS

- Reason for unequal *n*s should be random, not related to factor(s) themselves (more difficult with non-experimental studies)
  - If not so, validity of results is questionable when regular ANOVA procedures are employed

- Adjustments made to ANOVA to correct for unequal *n*s
  1. Analysis of weighted means: Non-recommended, but common, approach where imbalance is slight and imbalance is random
     1. Harmonic mean of cell *n*s is used in computation of various *MS*
     2. Total *N* is adjusted = Harmonic mean of all cell sizes x # cells
     3. MS_{Within} = Weighted average of cell variances
     4. Each row and column mean computed = Simple (non-weighted) average of cell means in a given row or column
  2. Alternate *SS* calculations to handle overlapping variation accounted for in outcome (Coming up next!)
  3. Regression analysis (Take EDUC/PSY 7610!)

$$\text{Harmonic mean} = \frac{\#cells}{\sum_{i=1}^{n} \frac{1}{n_i}}$$

ALTERNATIVE *SS* CALCULATIONS

- Several methods for partitioning or allocating variation between outcome and factor(s) to account for unbalanced designs
  - Commonly used
    - Type I SS: Sequential or Hierarchical
    - Type II SS: Partially Sequential
    - Type III SS: Simultaneous or Regression
  - Specialized and less commonly used
    - Type IV SS: Don't use
    - Type V SS: Used for fractional factorial designs
    - Type VI SS: Effective hypothesis tests though sigma-restricted coding

ALTERNATIVE *SS* RECOMMENDATIONS

- Type II or III SS recommended in most cases
  - Results should be fairly consistent
  - Type III is most commonly used
  - Nothing wrong with Type II
    - Considered by some to be more powerful, especially when testing main effects
    - Uncertainty of results when *n* are vastly unbalanced
- Not an issue when design is balanced
  - Type I-III yield same results
  - Even when unbalanced, interaction result same

INTERACTION CONTRASTS

- An alternative is to perform 'interaction contrasts', rather than immediately testing simple effects
  - With a 2x2 design, only tests of simple main effects are possible
  - With a 2x3 design, 3 separate 2x2 ANOVAs may be conducted
    - Interaction magnitude (and significance) can differ from one subset to another
    - Simple effects can be used following significant interaction subsets
    - MS_B for overall interaction = 'average' of $MS_{Interactions}$ for separate interaction subsets

IGNORING FACTORIAL DESIGN

- Treat each cell as a separate group (e.g., M/Rep, M/Dem, F/Rep, F/Dem) and run analysis as 1-Way ANOVA with RxC groups?
  - Results in same $SS_{Between}$ as factorial design ($SS_R + SS_C + SS_{RC}$; when study is balanced)
  - Cannot see patterns in data, as all levels of all factors are blended together in each group
    - Cannot as easily observe interaction effects
  - Limits identification of characteristics that uniquely differentiate participants
  - More cumbersome when many factors included
  - Less powerful

REPORTING RESULTS

- Marginal *M*s for main effects, cell *M*s for interactions and their *SD*s (or *SE*s) and *CI*s
- No need to report MS_W
- For each significant effect
  - $F(df_{Effect}, df_{Within}) = F_{stat}$, p-value, effect size (η^2 or ω^2)
- Results of post-hoc or planned comparisons
- Figures are *extremely* helpful!

CONCLUSIONS

- With a non-significant interaction
  - # of follow-up tests on main-effects needs to be kept low so as to not inflate α_{EW}, where each main-effect can contain a family of tests

- With a significant interaction
  - # of tests of simple effects or interaction contrasts should not exceed $df_{Interaction}$
  - For 2x2 ANOVA: # ≤ $(r\text{-}1)*(c\text{-}1)$
    - In `Conformity` data example = 2*1 = 2 tests
      - Did all 5 (should look back and use a more conservative α)

- Some forgo tests of simple main effects and compute all possible pairwise comparisons at cell level
  - Results in many, many tests

- Following a significant interaction and significant simple main effects
  - Not necessary to conduct all possible pairwise comparisons
  - Planned comparisons should be derived from theory or previous research and flow from research questions

- Significant unplanned interactions that do not conform to theory should be swallowed with a HIGH DEGREE OF SKEPTICISM