Research Methods II

Multi-Factorial Designs (Two-Way ANOVA and Beyond)

Stacey Coleman

Recap from last week:

- Multi-factorial designs
  - Multiple IV's, or more than 2 levels of an IV

- Analysis of Variance (ANOVA) can handle data from experiments that:
  - Involves **more than two** conditions (one-way within)
  - Involves **more than two** groups (one-way between)

- ANOVA allows us to investigate the effect of **more than one** IV
  - E.g., two-way, three-way, etc

- ANOVA allows us to investigate how multiple IV's **interact** to affect the DV

TERMS TO REMEMBER

- **FACTORS**
  - These are your independent variables (IVs)
- **LEVELS (of factors)**
  - Similar to conditions/groups
- **BETWEEN-SUBJECTS FACTOR**
  - Factors whose levels vary between participants
  - Each participant only experiences one level of a factor (e.g., male/female)
- **WITHIN-SUBJECTS FACTOR**
  - Factors whose various levels vary within a participant
  - Each participant experiences all levels of the factor (e.g., testing mood at 9am, 12pm and 3pm)

Two-way ANOVA

Between-subjects
Within-subjects
Mixed

Two-Way ANOVA

- Sometimes we are looking to ask a far more complex question, and we may want to examine *more than two* groups

- We may want to examine which of a number of variables **has the greatest influence over a given outcome…**

- Or whether one variable effects the DV **in light of another variable…**

- Or even explore the **interaction between TWO OR MORE variables** and determine how they effect the DV

Example…

- We are interested in exploring two types of teaching strategies (A & B) on learning
  - DV = reading scores
  - IV = teaching strategy A or B (2 separate groups)

Between participants design
Half of pupils in Strategy A, half in Strategy B

Results show that reading scores are higher for the students in Strategy B

Teaching example…

- What else could we could consider?
- Could original level of reading ability of children have an effect on their reading scores?
- How can we deal with this as a possible influencing factor?

Exploring multiple factors

- We want to look at more than one IV
- And we are interested in the **INTERACTION** between these
- Still looking at the effect on ONE DV (reading score)

Exploring multiple factors

IV 1	IV 2	DV
Teaching strategy	Original ability	Reading score after teaching
A or B	High or Low	

This is a **multi-factorial** design

Describing ANOVA's

- There are **THREE** things that you need to specify when describing the specific type of ANOVA you are using:

1. How many **factors** are involved? **2**
2. How many **levels** are there of each factor? **2 in each**
3. Are the factors **between-** or **within-subjects**?

Both between-subjects

Describing ANOVA's

- This gives us a **2 x 2 multi-factorial design**
- As both IV's are between-subjects, we can call the analysis a **two-way between-subjects ANOVA**
- Both IV's have two levels, so to test all combinations we need **4 conditions**

Multi-factorial Designs

- Both IV's have two levels, so we would have a 2 x 2 table of conditions (4 conditions)

	Strategy A	**Strategy B**
Low ability	Condition 1	Condition 2
High ability	Condition 3	Condition 4

Multi-factorial Designs

- Sample mean scores for the study:

	Strategy A	Strategy B	Means
Low ability	2	7	**4.5**
High ability	8	8	**8**
Means	**5**	**7.5**	

Main effects and interactions

- The first thing we can do here is assess the effect of each factor on their own

- We call this the **MAIN EFFECT:** this is used to describe the independent effect of each factor (IV).

- E.g., in this study, **2 main effects** would be reported:
  - Main effect of teaching strategy
  - Main effect of original ability

Main effects

- ‘Main effects’ are effects of ***ONE IV ALONE***
- Where is the main effect of ‘ability’ shown in the scores below?

	Strategy A	Strategy B	Means
Low ability	2	7	**4.5**
High ability	8	8	**8**
Means	**5**	**7.5**	

Main effects

- Where is the main effect of 'teaching strategy' shown in the scores below?

	Strategy A	Strategy B	Means
Low ability	2	7	**4.5**
High ability	8	8	**8**
Means	**5**	**7.5**	

Main effects

- The analysis can show us whether these ‘main effects’ are significant

- Effect of each IV in isolation on the DV score

  - Looking at whether there is a significant difference between conditions?

  - E.g., is the difference between A and B significant for teaching strategy?

And interaction effects?

Interaction effects

- If we have more than one IV then we are clearly interested in examining how the **TWO** IV's may influence the DV **together**

- So – we analyse any possible 'interaction' between IV's

  - Combination of IV's in effecting the DV
  - Differences dependent upon conditions of both IV's

Main Effects

- Compare means for **one level/condition** with the mean of the **other level(s)/condition(s)** of that factor

- In the same way as an experiment with only one IV
- You would simply have more than one main effect (2 IV's = 2 main effects, etc)

Interaction Effect

- Assess the combined effect of the factors (all the IV's)

Interaction effect

- The interaction effect that we are interested here is a **two-way interaction**

Teaching strategy x Ability

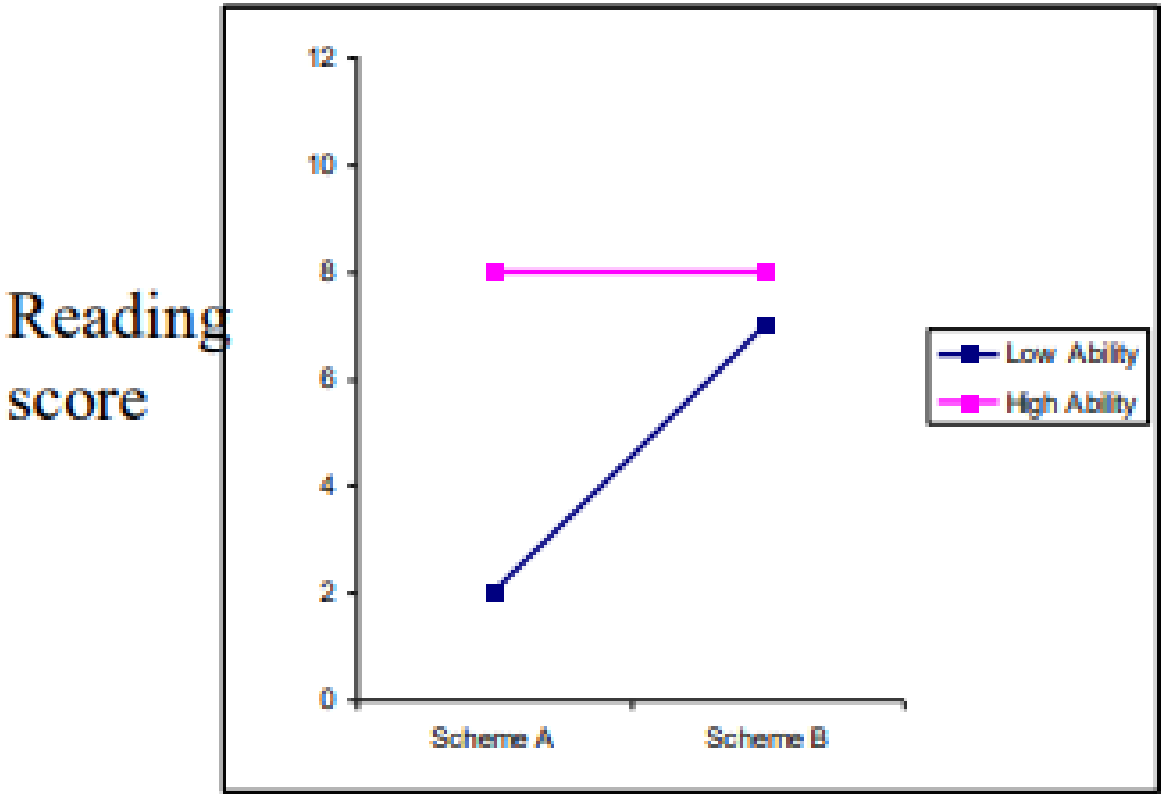
*(you will notice when running the analysis that this is shown as strategy*ability)*

Interaction effect

- How do the two IV's interact?
- What is the effect of **strategy** only for the **LOW** ability group?

	Strategy A	**Strategy B**	**Means**
Low ability	2	7	**4.5**
High ability	8	8	**8**
Means	**5**	**7.5**	

Using graphs to display interactions



- This graph shows an interaction – but how?

Effect of strategy IV (higher scores for B) but effect only shown for one ability (IV) group

Using graphs to display interactions

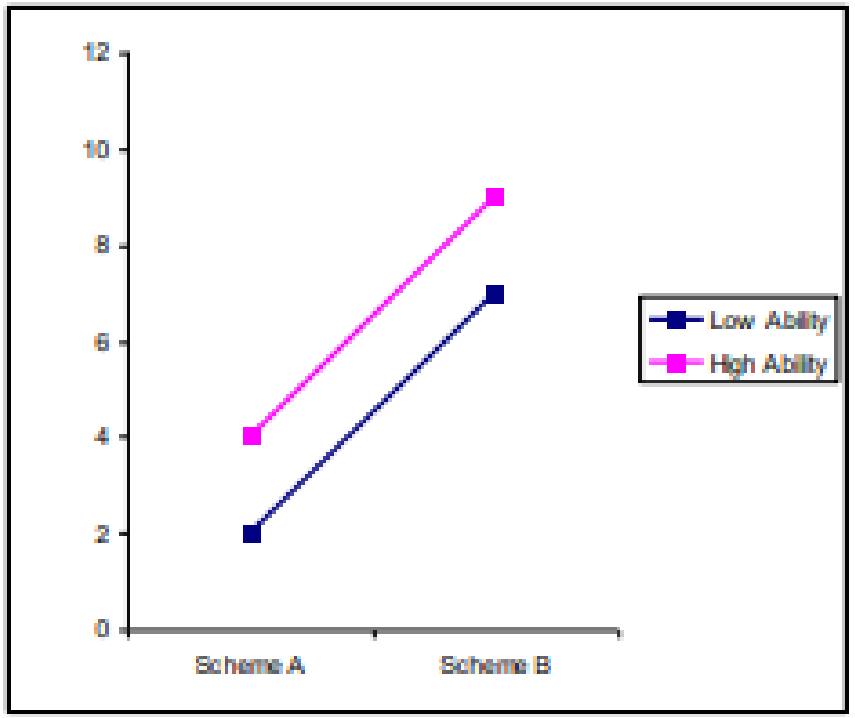
- The key to interpreting these graphs is to assess the relative **slope** of the lines

- Same/similar slope = likely **no interaction**

- Very different slopes indicate that there **may be an interaction** between the IV's

Using graphs to display interactions

Reading score



- No interaction
- Main effects of strategy IV and ability IV

Mixed Design

- Involves:
  - At least one between-subjects IV
    - This IV can have two or more levels
  - At least one within-subjects IV
    - This IV can have two or more levels

  - One continuous DV

  - Included together in one analysis

- A very common structure for studies

Reporting ANOVA statistics

- ANOVA will analyse:
  - Main effects (each IV on its own)
  - Interaction effects (combined effects of the IVs)
  - Main stats for each analysis
    - *F*-value, degrees of freedom, p-value (sig.)

  - All showing whether (or not) there are significant differences (and to what degree through the effect size statistic) between the means in each grouping of data

Interpretation of ANOVA Output

Rule 1: USE YOUR STATS BOOKS!

- Nothing can explain the steps to interpreting statistical outputs like a stats book; this is the only way you will understand it for your yourself!
- Tells you exactly how to run the analysis, what to interpret and how to report it
- Just make sure you are aware of exactly what TYPE of ANOVA analysis you are conducting based upon your experimental design…

Note 1…

- With ANOVA outputs, although the main values you will need to report results are the *F, df,* and *sig.* values, you still need to check other tables from the output (depending on the type of ANOVA) in order to ensure there have been no violations of assumptions and to ensure you have read the values you need from the correct row…

Interpreting ANOVA outputs: Between-subjects

- Report descriptives
- Check Levene's test for equality of error variances
  - Needs to be **GREATER THAN** 0.05 (i.e., non-significant)
- Report interaction effects
- Report main effects
- Report effect sizes
- Report planned or post-hoc tests (if needed)

Interpreting ANOVA outputs: Between-subjects

- Levene's test for equality of error variances
  - Needs to be **GREATER THAN** .05 (i.e., non-significant)

Levene's Test of Equality of Error Variances[a]

Dependent Variable: score

F	df1	df2	Sig.
.651	3	26	.589

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Check that the value in the Sig. box is bigger than 0.05 (i.e., **NOT** significant). This tells us that the assumption of homogeneity of variances has not been violated and we are safe to proceed with our interpretation of the ANOVA.

Interpreting ANOVA outputs: Within-subjects

- Report descriptives
- Check Mauchly's test for equality of error variances (only if any of the factors has ***more than two*** levels)
  - Needs to be **GREATER THAN** 0.05 (i.e., non-significant)
- Report interaction effects
- Report main effects
- Report effect sizes
- Report pairwise comparisons (if needed)

Interpreting ANOVA outputs: Within-subjects

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Time	.927	19.536	2	.000	.932	.942	.500

The assumption being tested is that the correlation between all the variables is roughly the same. The sig value here should ideally be bigger than 0.05 (i.e., **NOT** significant). If it is significant then the assumption of sphericity is violated and you will need to report values from either the 'Multivariate Tests' box or other rows in the 'Tests of within-subjects effects' box in the output.

**Note: If a within-subjects variable only has two levels, this box will be empty*

Interpreting ANOVA outputs: Mixed

- Same as that for within- and between-subjects
  - i.e., descriptives, assumptions, interaction effects, main effects, followup-tests if necessary
  - Need to ensure you check **both** Levene's test and Mauchly's test (if applicable)
  - However, also need to check Box's Test of Equality of Covariance Matrices (also called Box M).

Interpreting ANOVA outputs: Mixed

Check that the Sig. value here is bigger than .001 (i.e., **NOT** significant) to ensure that this assumption has not been violated

Box's Test of Equality of Covariance Matrices[a]

Box's M	8.555
F	1.408
df1	6
df2	473190.960
Sig.	.207

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + group
Within Subjects Design: Time

Note 2…

- There are **always** two values for the degrees of freedom (DF) with ANOVA:
  - The value from the row of your IV will provide you with the first df
  - The value from the top ‘error’ row will provide you with the second df

DF1 explanation…

- The number of cell means that are free to vary to get to a given grand mean
  - NB: marginal means with 2+ way ANOVA
- E.g., one way BS ANOVA with 3 levels
- DF = levels (3) -1 = **2**

	1	2	3	Grand mean
Mean	4	6	8	6

DF1 explanation…

- Two-way between-subjects designs?
- Marginal means and grand means come into play…
- Easy way to calculate..
- (factor 1 levels – 1) x (factor 2 levels – 1) = DF1
  - E.g., 2 x 2 design
  - (2 – 1) x (2 -1) = 1
  - **So DF1 would be 1.**

DF1 (between-subjects only)

- E.g., 2 x 2 design
- (2 – 1) x (2 – 1) = 1
- DF1 = 1

- E.g., 2 x 3 design
- (2 – 1) x (3 – 1) = 2
- DF1 = 2

DF2 explanation (between-subjects)

- DF2 relates to single observations in the cells and how these relate to the cell means…

- Total number of observations in all cell(s) **MINUS** the DFs lost because the cell means are set.

  - So – minus the number of cell means or groups/conditions

DF2 explanation (between-subjects)

- Easy way to calculate…
- DF2 = n – k
- Where n = participants (or total observations)
- And k = df's lost (or no* of conditions/groups)
- If we had 30 participants, and a 2 x 2 design (therefore 4 cells) then DF2 would be 30 – 4 = 26

- SPSS will always provide you with DF1 and DF2 for any ANOVA analysis that you conduct (DF2 labelled as [error])

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
colour	Sphericity Assumed	336.467	2	168.233	4.725	.022	.344
	Greenhouse-Geisser	336.467	1.517	221.761	4.725	.035	.344
	Huynh-Feldt	336.467	1.760	191.133	4.725	.028	.344
	Lower-bound	336.467	1.000	336.467	4.725	.058	.344
Error(colour)	Sphericity Assumed	640.867	18	35.604			
	Greenhouse-Geisser	640.867	13.655	46.932			
	Huynh-Feldt	640.867	15.843	40.450			
	Lower-bound	640.867	9.000	71.207			

Reporting of ANOVA statistics

- With two-way ANOVA (and beyond), we need to report the **interaction effects** and the **main effects**
  - For a two-way ANOVA, we would report **one** two-way interaction effect and **two** main effects

- General reporting formula is the same as with the one-way ANOVA

F (df1, df2) = F-value, p </> .05, η^2 = _____

F (df1, df2) = F-value, p = _____, η^2 = _____

Reporting Example

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
colour	Sphericity Assumed	336.467	2	168.233	4.725	.022	.344
	Greenhouse-Geisser	336.467	1.517	221.761	4.725	.035	.344
	Huynh-Feldt	336.467	1.760	191.133	4.725	.028	.344
	Lower-bound	336.467	1.000	336.467	4.725	.058	.344
Error(colour)	Sphericity Assumed	640.867	18	35.604			
	Greenhouse-Geisser	640.867	13.655	46.932			
	Huynh-Feldt	640.867	15.843	40.450			
	Lower-bound	640.867	9.000	71.207			

$$F(2, 18) = 4.73,\ p = .02,\ \eta^2 = .34$$

Mixed ANOVA

Mixed ANOVA

- Combines within-subjects variables and between-subjects variables in one analysis

- Requires (at the minimum);
  - One categorical between-subjects factor (2 or more levels)
  - One categorical within-subjects factor (2 or more levels(
  - One continuous DV

Mixed ANOVA: Study Example

Research question 1: the effect of subject of study (psychology and art) on end of year grades (DV)

- IV = between-subjects (2 levels; students either study psychology OR they study art)
- We find that scores were higher for psychology students

Mixed ANOVA: Study Example

- There are other factors that we may need to consider
- What about **year of study?**
  - i.e., assessing grades in year one, year two, and year three

- We now need to consider exploring multiple factors and also a **mixed design**

- Can investigate the **interaction** between factors
- Still looking at the effect on one DV (end of year grades)

New Study

- Factor 1: Subject (Psychology OR Art)
- Factor 2: Year of study (year one, year two, AND year three)
- DV: End of year grade

- This is a mixed multifactorial design

ANOVA: Main Effects and INTERACTIONS

- Mixed ANOVA allows us to investigate how the multiple IV's interact to affect the DV, and also each IV in isolation.
  - Interaction effect and main effects

- Using the previous example…

  - Psychology students may have better end of year grades than art students; students in year three may have better grades than they did in year one

  - OR – we may see an interaction
    - Psychology students have better grades in year three while art students score higher in year one

Study Design

- We have TWO IV's
  - We need to combine the conditions of these

- Subject: Psychology or Art (between-subjects – two levels)
- Year of study: Year One, Year Two, Year Three (within-subjects – three levels)

- This gives us a **2 x 3 mixed design**

Main Effects and Interactions

- In the same way as the two-way between-subjects ANOVA, we need to assess the **main effects** as well as the **interaction effect** between our two factors
- The interaction effect that we are interested in here is a two-way interaction

Subject x Year

*(you will notice when you run the analysis that this is shown as Subject*Year)*

SPSS Output

1. Descriptive Statistics

Descriptive Statistics

	Subject	Mean	Std. Deviation	N
Year1	Psychology	66.4500	7.98996	20
	Art	54.0000	9.13063	20
	Total	60.2250	10.55750	40
Year2	Psychology	68.2000	7.22277	20
	Art	55.1000	8.62615	20
	Total	61.6500	10.27955	40
Year3	Psychology	69.4000	8.22320	20
	Art	58.0000	8.35401	20
	Total	63.7000	10.01332	40

SPSS Output

2. Assumptions

Check that the Sig. value here is bigger than .001 (i.e., **NOT** significant) to ensure that this assumption has not been violated

Box's Test of Equality of Covariance Matrices[a]

Box's M	12.768
F	1.945
df1	6
df2	10462.189
Sig.	.070

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + Subject Within Subjects Design: Year

SPSS Output

Only displayed if within-subjects factor has at least three levels

2. Assumptions

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Year	.930	2.690	2	.260	.934	1.000	.500

The assumption being tested is that the correlation between all the variables is roughly the same. The sig value here should ideally be bigger than .05 (i.e., **NOT** significant).

If it is significant, then the assumption of sphericity has been violated and you will need to report values either from the 'Multivariate Tests' box or other rows in the 'Tests of Within-Subjects Effects' box in the output.

SPSS Output

2. Assumptions

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Year1	.195	1	38	.661
Year2	1.325	1	38	.257
Year3	.007	1	38	.932

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Subject
Within Subjects Design: Year

Check that the Sig. value for each conditions in this box is larger than .05 (i.e., **NOT** significant). We need to make sure that all are not significant because this tells us that the assumption of homogeneity of variances has not been violated, and so we are safe to proceed with our interpretation of the ANOVA

SPSS Output

Be careful when reading the columns
The last column 'Partial eta squared' shows the estimate of the effect size

3. Interaction effect

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Year	Sphericity Assumed	244.117	2	122.058	22.783	.000	.375
	Greenhouse-Geisser	244.117	1.869	130.618	22.783	.000	.375
	Huynh-Feldt	244.117	2.000	122.058	22.783	.000	.375
	Lower-bound	244.117	1.000	244.117	22.783	.000	.375
Year * Subject	Sphericity Assumed	14.717	2	7.358	1.373	.259	.035
	Greenhouse-Geisser	14.717	1.869	7.874	1.373	.259	.035
	Huynh-Feldt	14.717	2.000	7.358	1.373	.259	.035
	Lower-bound	14.717	1.000	14.717	1.373	.249	.035
Error(Year)	Sphericity Assumed	407.167	76	5.357			
	Greenhouse-Geisser	407.167	71.019	5.733			
	Huynh-Feldt	407.167	76.000	5.357			
	Lower-bound	407.167	38.000	10.715			

a. Computed using alpha = .05

No significant interaction!

SPSS Output

4. Main effects

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Year	Sphericity Assumed	244.117	2	122.058	22.783	.000	.375
	Greenhouse-Geisser	244.117	1.869	130.618	22.783	.000	.375
	Huynh-Feldt	244.117	2.000	122.058	22.783	.000	.375
	Lower-bound	244.117	1.000	244.117	22.783	.000	.375
Year * Subject	Sphericity Assumed	14.717	2	7.358	1.373	.259	.035
	Greenhouse-Geisser	14.717	1.869	7.874	1.373	.259	.035
	Huynh-Feldt	14.717	2.000	7.358	1.373	.259	.035
	Lower-bound	14.717	1.000	14.717	1.373	.249	.035
Error(Year)	Sphericity Assumed	407.167	76	5.357			
	Greenhouse-Geisser	407.167	71.019	5.733			
	Huynh-Feldt	407.167	76.000	5.357			
	Lower-bound	407.167	38.000	10.715			

a. Computed using alpha = .05

Significant main effect of Year

SPSS Output

4. Main effects

Tests of Between-Subjects Effects

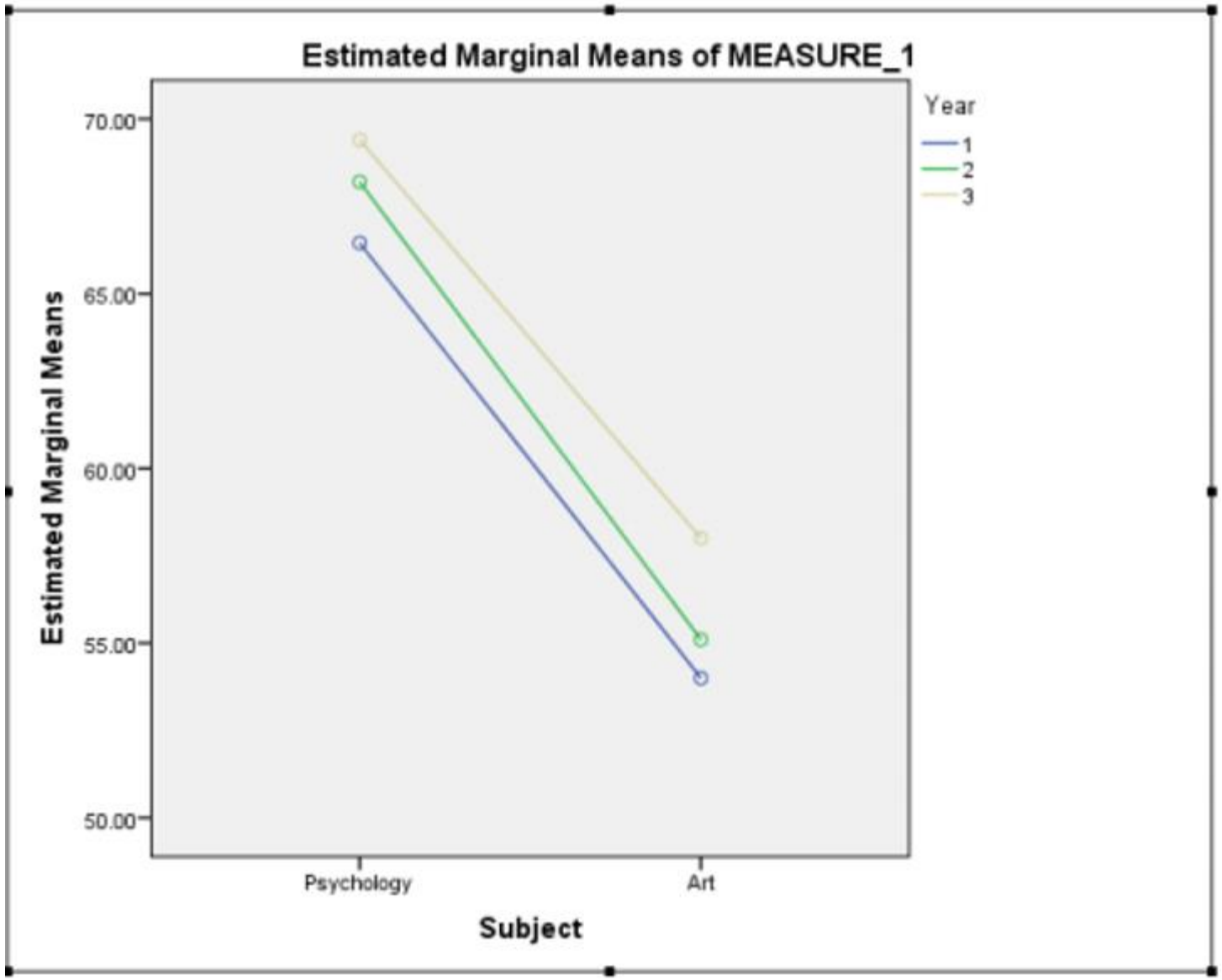
Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	459174.408	1	459174.408	2356.145	.000	.984
Subject	4551.008	1	4551.008	23.352	.000	.381
Error	7405.583	38	194.884			

a. Computed using alpha = .05

Significant main effect of Subject



No interaction effect

Slopes for all three lines are relatively similar

Significant Main Effect of Three-Level Within-Subjects Factor

- We have noted a significant main effects of our IV 'Year' on end of year grades
  - Within-subjects IV with three levels

- What the sig. main effects **does not** tells us is between which two years this difference lies…

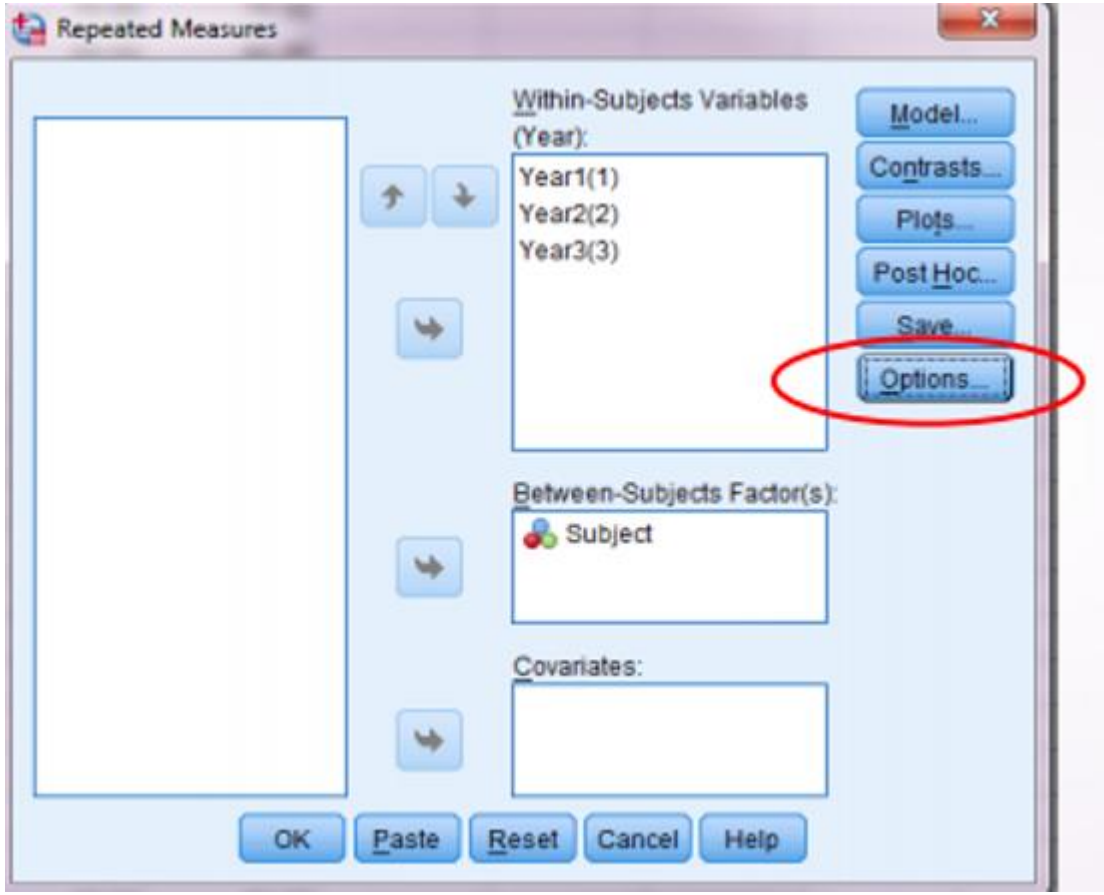
Exploring Significant Main Effects when there are <u>more than</u> two levels

- Are you expecting differences at the start of your study?
  - Planned comparisons (*a priori)* – lower the alpha, or conducted Bonferroni corrected t-tests

- Are you wanting to explore non-hypothesised differences
  - Post-hoc comparisons for between-subjects factors (e.g., Tukey, Scheffe)
  - Comparing main effects for within-subjects factors (e.g., Bonferroni – pairwise comparisons)

Recording on Planned / Post-hoc comparisons!

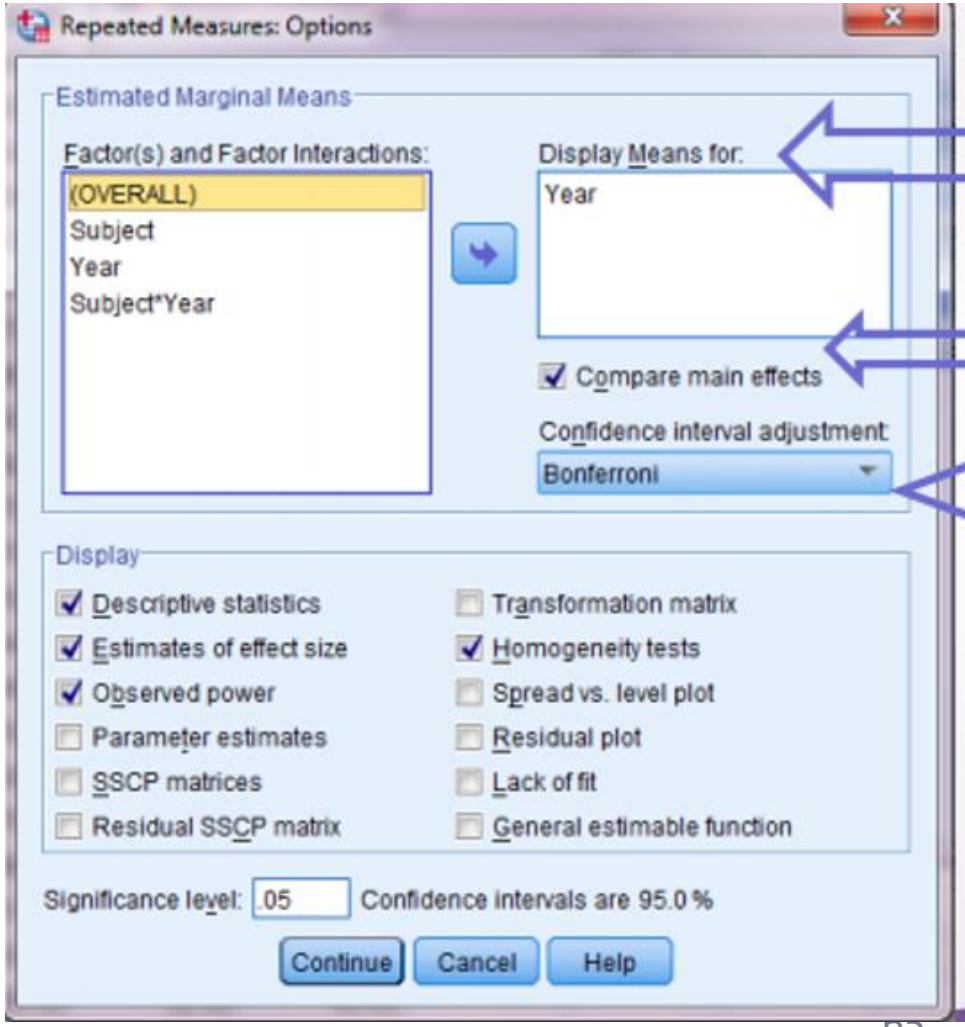
Comparing Main Effects



- The IV with more than two-levels was a within-subjects factor, therefore we are looking to **compare main effects**

- Use SPSS to perform ***Pairwise Comparisons***

Comparing Main Effects



Move the variable of interest into the 'Display means for' box

Tick the 'Compare main effects' box.

Change 'Confidence interval adjustment' to 'Bonferroni'.

SPSS Output

5. (if needed) Pairwise Comparisons

Pairwise Comparisons

Measure: MEASURE_1

(I) Year	(J) Year	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-1.425*	.463	.012	-2.584	-.266
	3	-3.475*	.505	.000	-4.739	-2.211
2	1	1.425*	.463	.012	.266	2.584
	3	-2.050*	.579	.003	-3.499	-.601
3	1	3.475*	.505	.000	2.211	4.739
	2	2.050*	.579	.003	.601	3.499

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

The Sig. values in this table show us that there is a significant difference in each of year grades between years 1 and 2, years 1 and 2, AND years 2 and 3.

Reporting Results

- ANOVA will generate information for:
  - Interaction effect(s)
  - Main effects

- Main statistics for each analysis:
  - F-value, degrees of freedom (df), and p-value (Sig.) and partial eta squared (if selected)
  - Need two values for the df: one from the row of the IV and one from the ‘Error’ row
  - Sig. value < .05 is significant

Interpreting ANOVA outputs: Mixed

- Same as that for within- and between-subjects
  - i.e., descriptives, assumptions, interaction effects, main effects, followup-tests if necessary
  - Need to ensure you check **both** Levene's test and Mauchly's test (if applicable)
  - However, also need to check Box's Test of Equality of Covariance Matrices (also called Box M).

Basic Results Section for The Study Example

*This does not include the descriptive statistics – ***DO NOT*** forget to include the descriptives at the beginning*

A 2 x 3 mixed ANOVA was carried out to explore the effect of the subject being studied at university and the year of study on end of year grades in university students. Results show that there was no significant interaction between the subject (Psychology or Art) and year (1, 2 or 3), $\boldsymbol{F\ (2, 76) = 1.37, p = .25, \eta^2 = .04}$. There was a significant main effect of the subject studied **[$\boldsymbol{F\ (1, 38) = 23.35, p < .001, \eta^2 = .38}$],** with psychology students achieving significantly higher grades overall than art students. The main effect of year of study was also significant; $\boldsymbol{F\ (2, 76) = 22.78, p < .001, \eta^2 = .38}$. Pairwise comparisons indicate that there was a significant difference in end of year grades between year 1 and 2 **($\boldsymbol{p < .05}$),** year 2 and 3 **($\boldsymbol{p < .001}$),** and year 1 and 3 **($\boldsymbol{p < .005}$).**

Final Reminders and Tips for Good Practice

- The main ANOVA statistics needed are the *F*, df's, and Sig. values, but you **MUST** check other tables in the output (depending on the type of ANOVA) to ensure no violations of the assumptions and that you are reading the values from the correct rows
- Always double check your data
- Reduce decimal places to 0 if appropriate
- Ensure you select to generate 'Estimates of effect size'
- Start reporting exact p-values