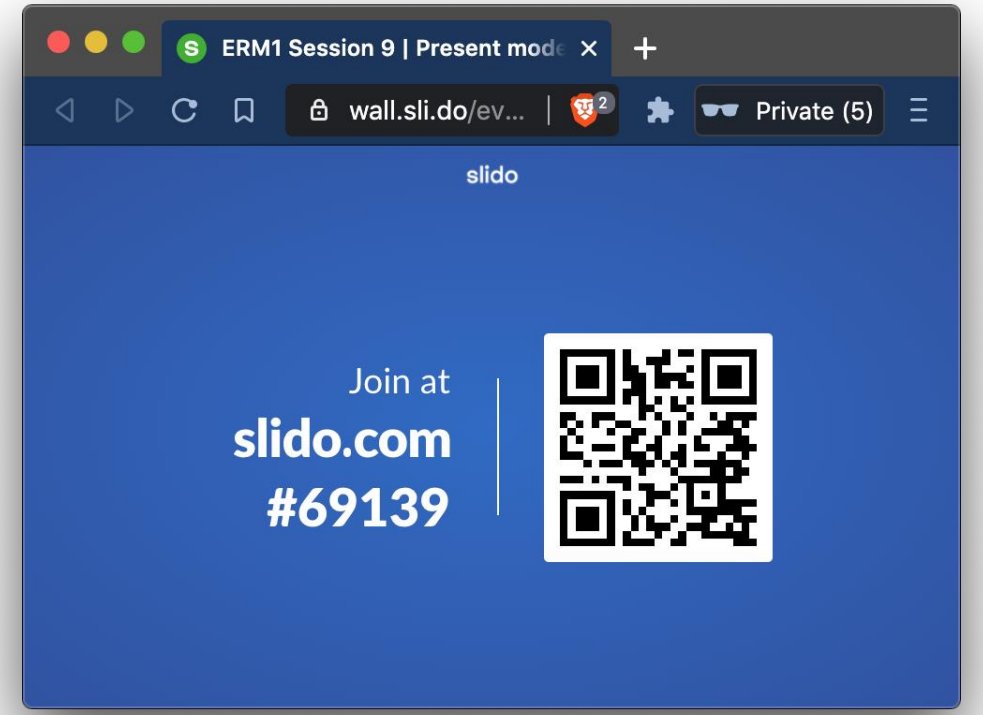
Session start question

https://app.sli.do/event/4rtwrkkj



Empirical Research Methods 1

ANOVA between groups (independent samples): One-way and Two-way ANOVA between groups

Miguel Rejón,

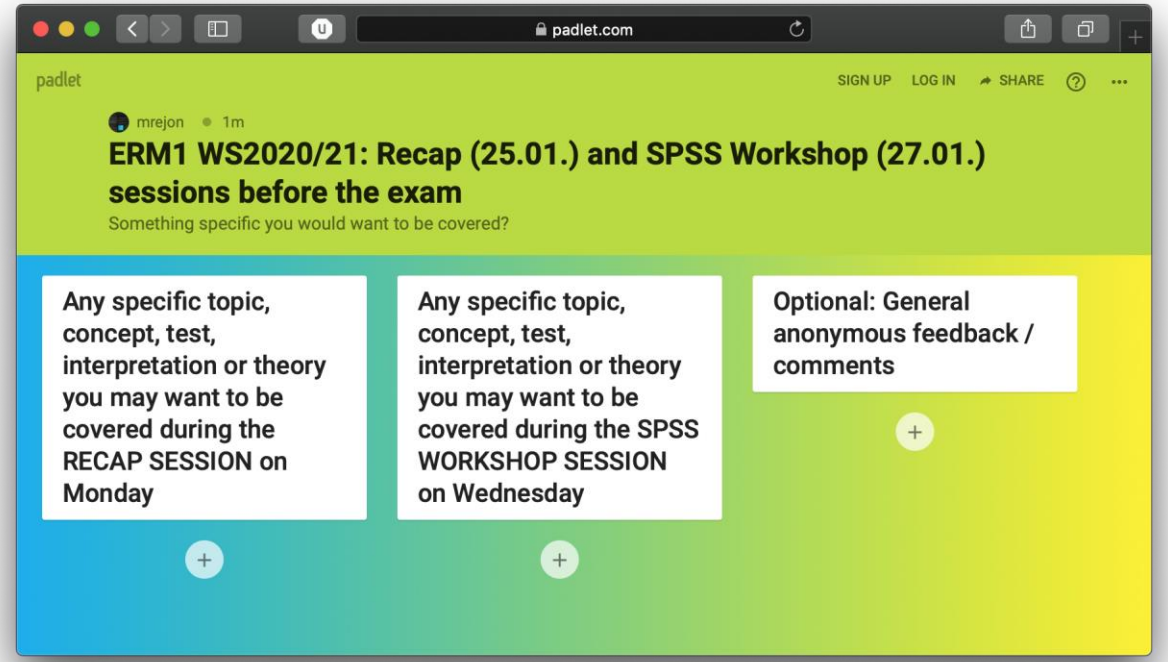
Room 1.14

m.rejon@edutech.uni-saarland.de

Before we start 1/2: Info and tools

- T-test interpretation: clearer now? Let me know! MS Teams, padlet or slido
- Exam: update coming soon. Very likely to be on 08.02. **in-person.** If not in SB, let me know!
- Slido for session start question and anonymous Q&A
- MS Teams for comments, questions, etc
- SPSS for exercise livecast

Before we start 2/2: ERM1 recap and SPSS workshop before the exam



https://padlet.com/mrejon/y71gbr4qt5yx720d

Contribute until Wednesday Jan 20

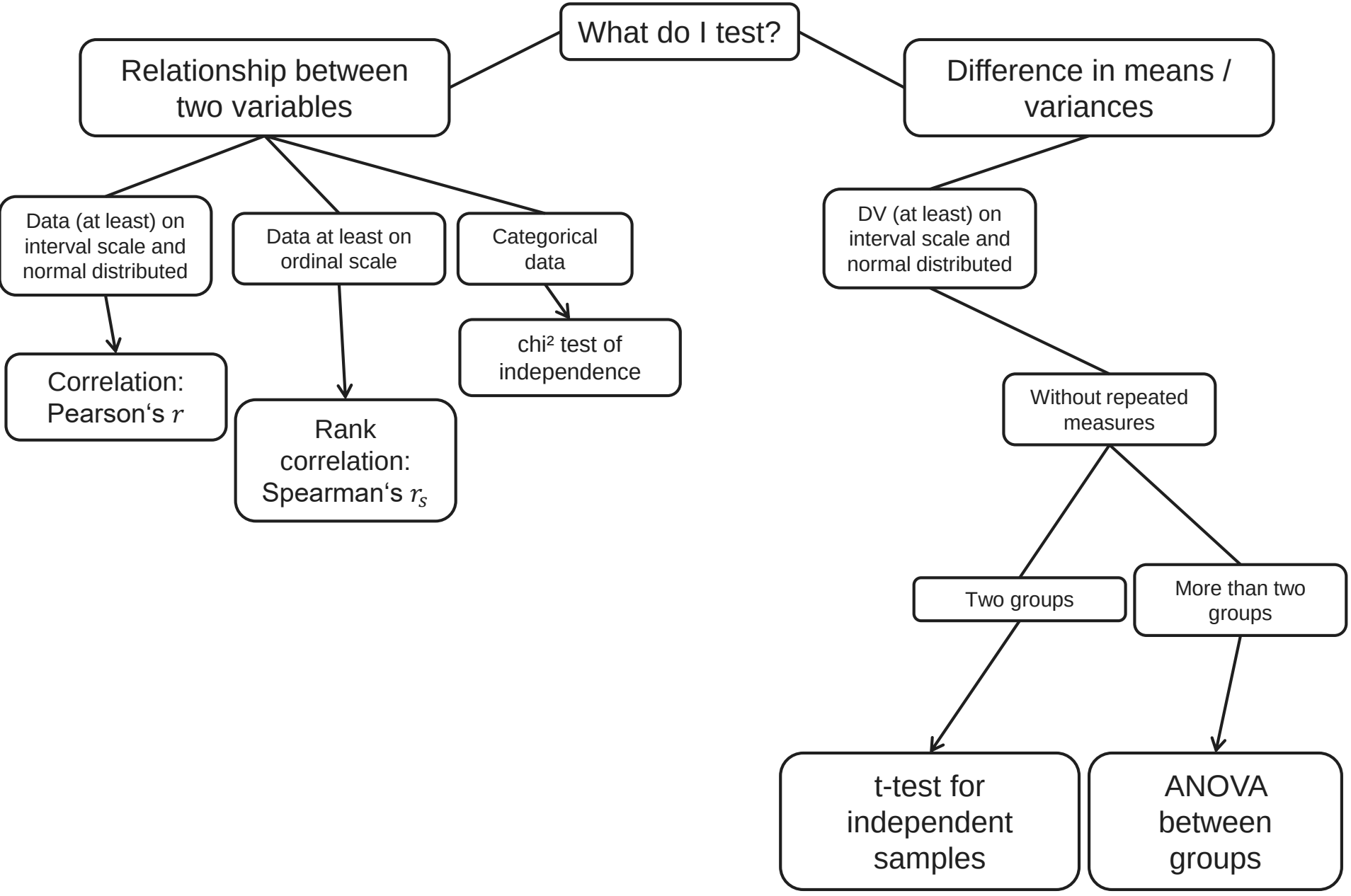
Add (“+” button), comment, or like posts

Agenda

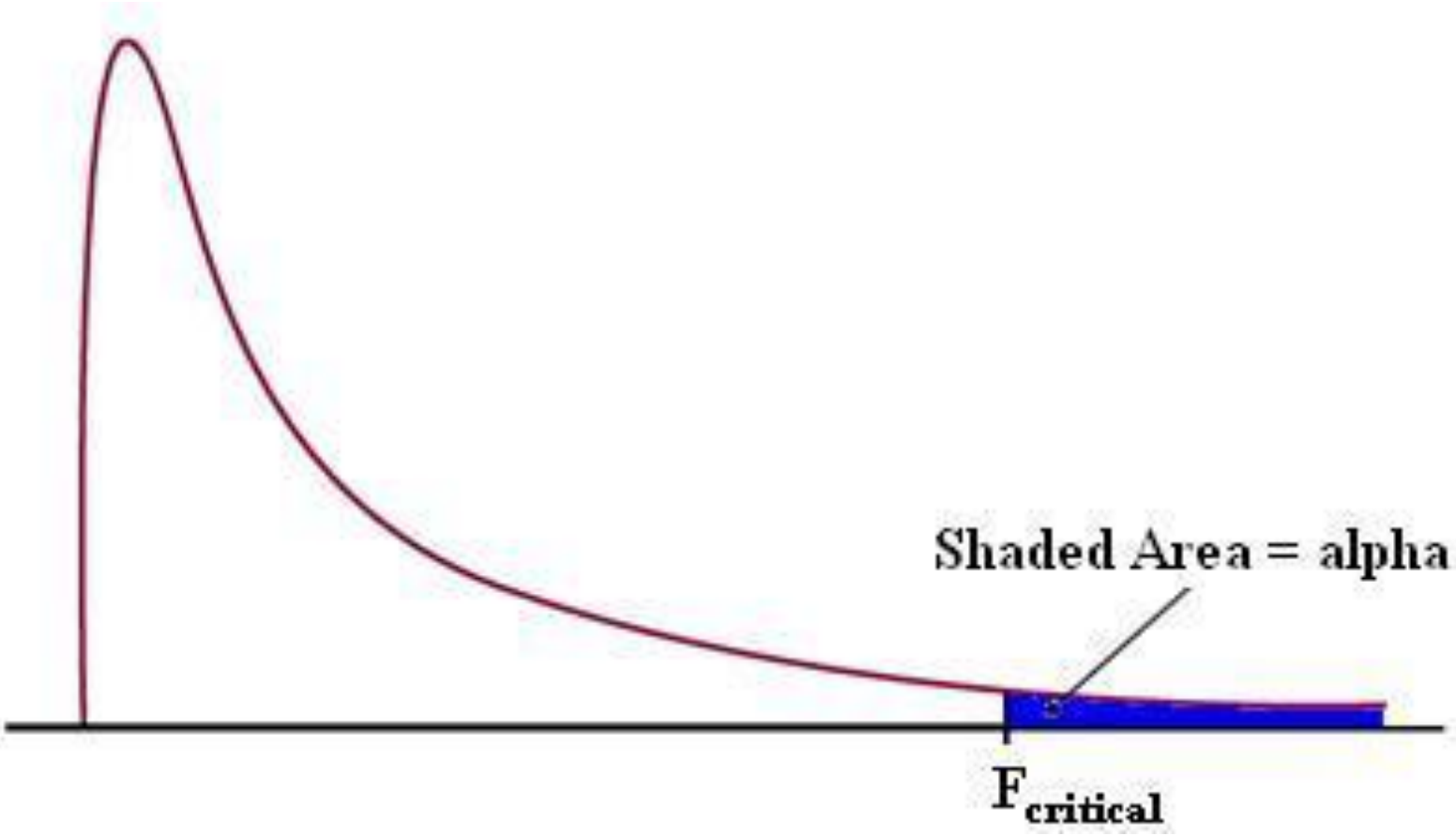
- Recap of previous concepts related to today's topic
- Updated "What do I test?" chart. F-Distribution
- T-test vs one-way ANOVA
- Two-way ANOVA. Multifactorial design and plot interpretation for main and interaction effects
- ANOVA extra info: Omnibus test; a priori and post hoc tests
- ANOVA essentials recap
- ANOVA on SPSS
- Self-paced exercises (for you)
- Qualis course evaluation survey

Recap

- Independent Variable: a.k.a. IV, factor, input variable, experimental variable, and predictor (variable). Is the variable that is manipulated in our experiment. We expect the IV to have an effect on the dependent variable
- Factors: IV
- Levels: the possible values within a factor. If a factor is a category, a level is a sub-category
- Dependent Variable: a.k.a. DV, outcome variable, response variable. Is the variable that we measure *within our experiment*. We expect the DV to be affected, or depend of, the independent variable
- Categorical Variable: Nominal scale variable with labels used for distinguishing the levels of a factor. The labels can be numerical but won't be treated as such!
- Null hypothesis: Hypothesis that states no difference or change between factors and/or levels
- Alternative hypothesis: Hypothesis that states the existence of a difference or change between factors and/or levels. Can either be directional or non-directional.
- Significance level: a.k.a. p-value, measure of the probability that an observed difference could have occurred just by random chance. That's why we want p-values to be as small as possible.


What do I test?
Relationship between two variables
Difference in means / variances
Data (at least) on interval scale and normal distributed
Data at least on ordinal scale
Categorical data
DV (at least) on interval scale and normal distributed
Correlation: Pearson‘s r
Rank correlation: Spearman‘s r_s
chi² test of independence
Without repeated measures
Two groups
More than two groups
t-test for independent samples
ANOVA between groups

F distribution: The probability distribution for ANOVAs



http://www.statisticshowto.com/wp-content/uploads/2013/09/f-table.jpg

t-test vs. One-way ANOVA

With two groups: both are applicable, but t-test is more common

Independent Samples Test

		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
alter	Equal variances assumed	,306	,581	1,937	148	,055
	Equal variances not assumed			2,125	133,057	,035

Tests of Between-Subjects Effects

Dependent Variable: alter

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	35,920[a]	1	35,920	3,751	,055
Intercept	65917,947	1	65917,947	6882,861	,000
sex	35,920	1	35,920	3,751	,055
Error	1417,413	148	9,577		
Total	73176,000	150			
Corrected Total	1453,333	149			

a. R Squared = ,025 (Adjusted R Squared = ,018)

t-test vs. One-way ANOVA

- With more than two groups:
  - Alpha inflation: The probability of making an alpha error is increasing when you calculate many t-tests
  - → ANOVA: efficient procedure to compare several groups, avoids alpha inflation
- With more than one factor:
  - → ANOVA: ability to test multi-factorial designs
    - Main effects
    - Interaction effects

Multi-factorial design for a 2-way ANOVA

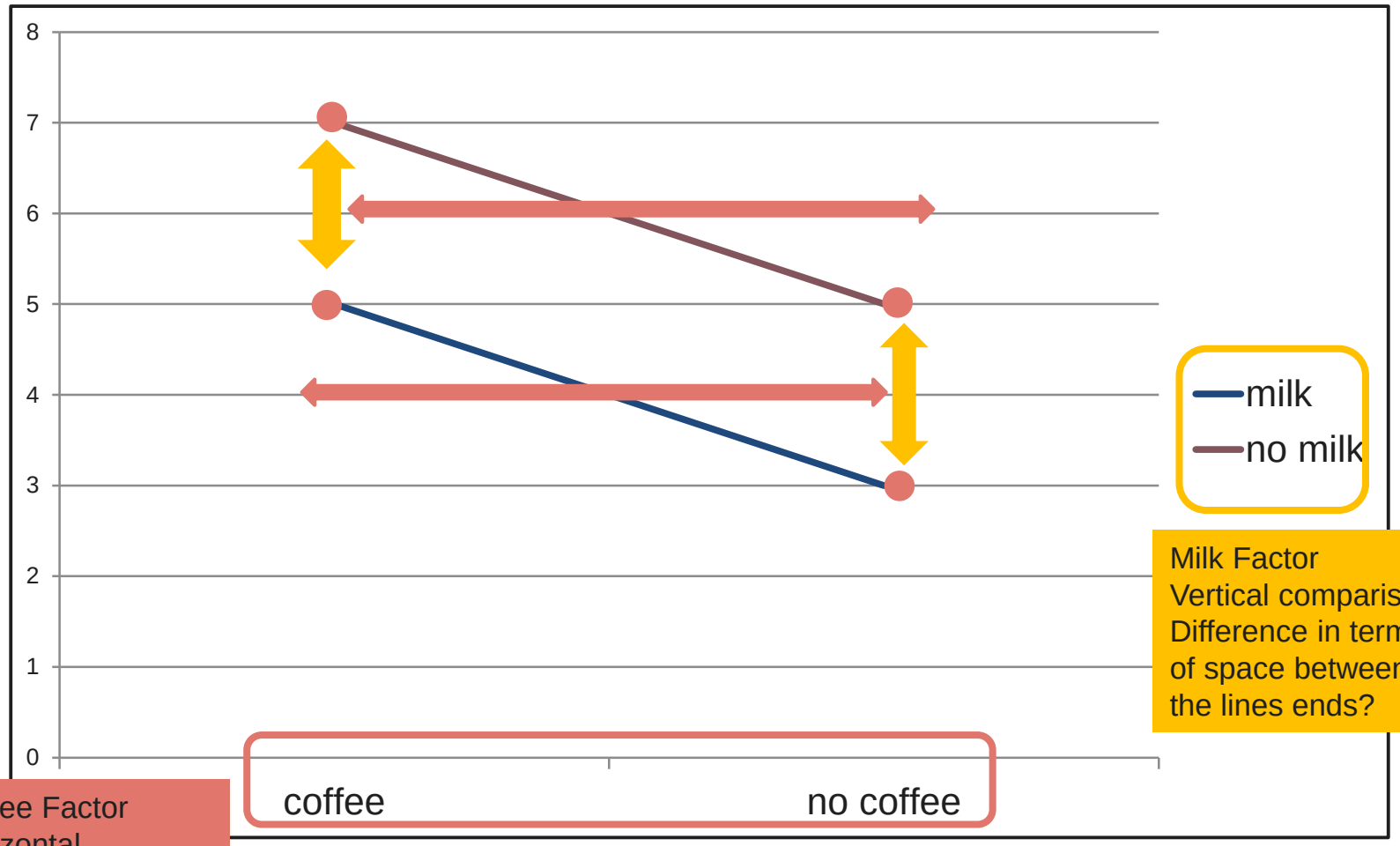
ANOVA calculates:

- Main effect for coffee
- Main effect for milk
- Interaction between coffee and milk

1- From the above factorial design, which are the factors and which are the levels?
2- How would a 1-way ANOVA factorial design look like?

http://versuch.file2.wcms.tu-dresden.de/w/images/thumb/9/98/Mehrfaktoriell_Beispiel_Glas.jpg/391px-Mehrfaktoriell_Beispiel_Glas.jpg

Main and interaction effects in Two-way ANOVA

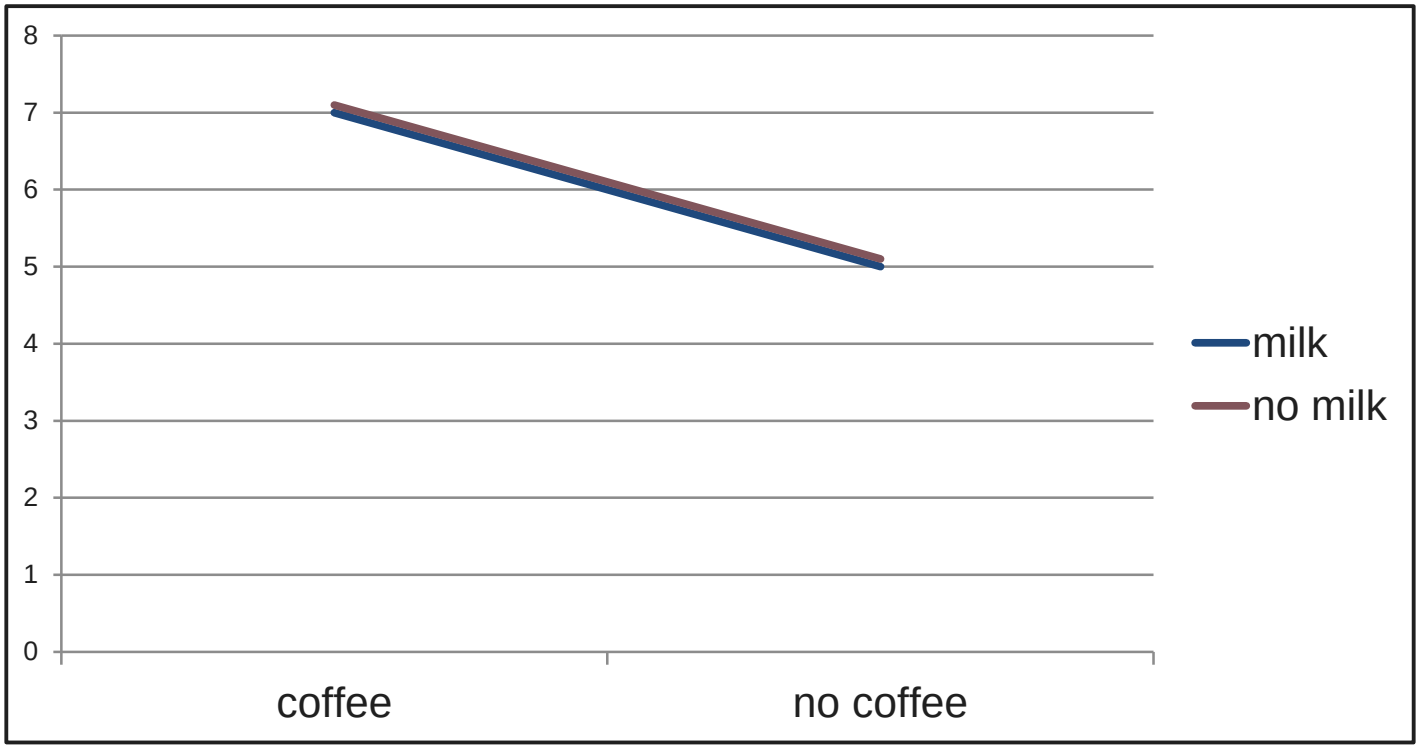


Milk Factor
Vertical comparison: Difference in terms of space between the lines ends?

Coffee Factor
Horizontal comparison: Difference in terms of inclination of the lines?

Main effect Factor milk. Main effect Factor coffee. No interaction effect between factors milk-coffee

Main and interaction effects in Two-way ANOVA

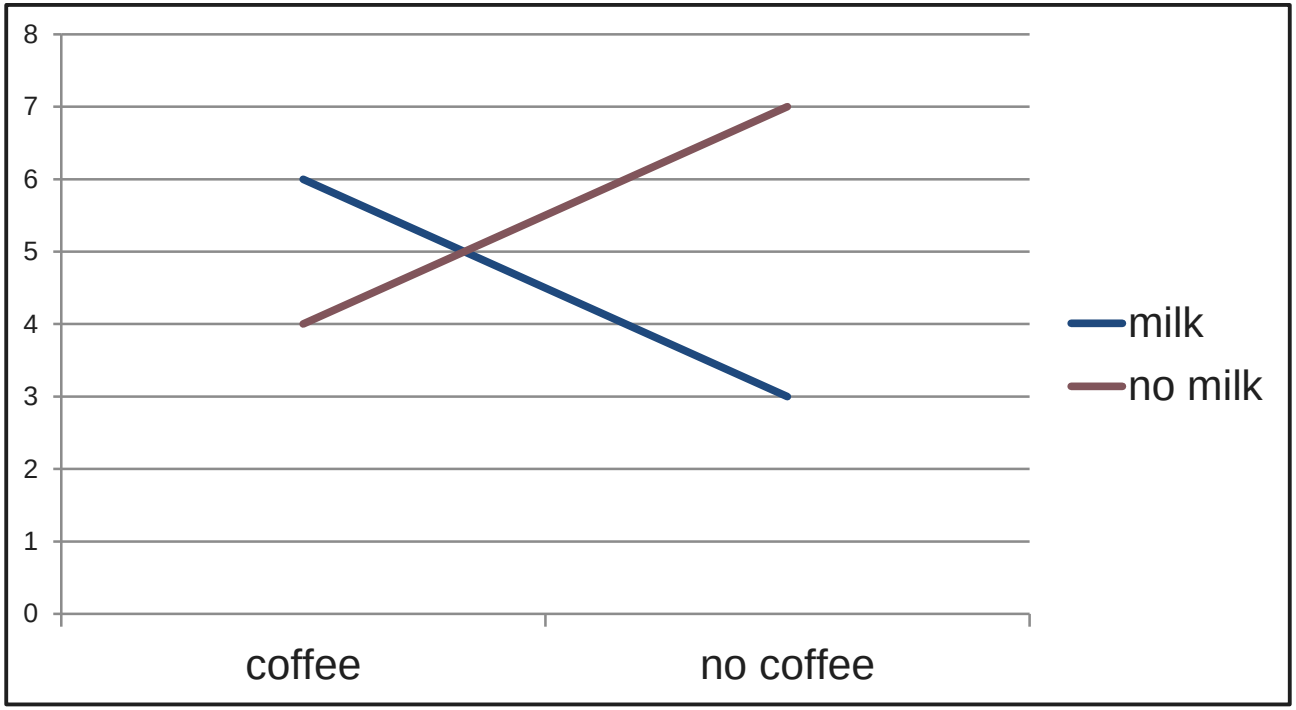


Milk factor: are the lines ends apart from each other? NO → No main effect Factor milk.

Coffee factor: are the lines inclined indicating a diff between the coffee factor levels? YES → Main effect Factor coffee.

Interaction effect: do the lines cross each other or are about to? NO
→ No interaction effect between factors milk-coffee

Main and interaction effects in Two-way ANOVA

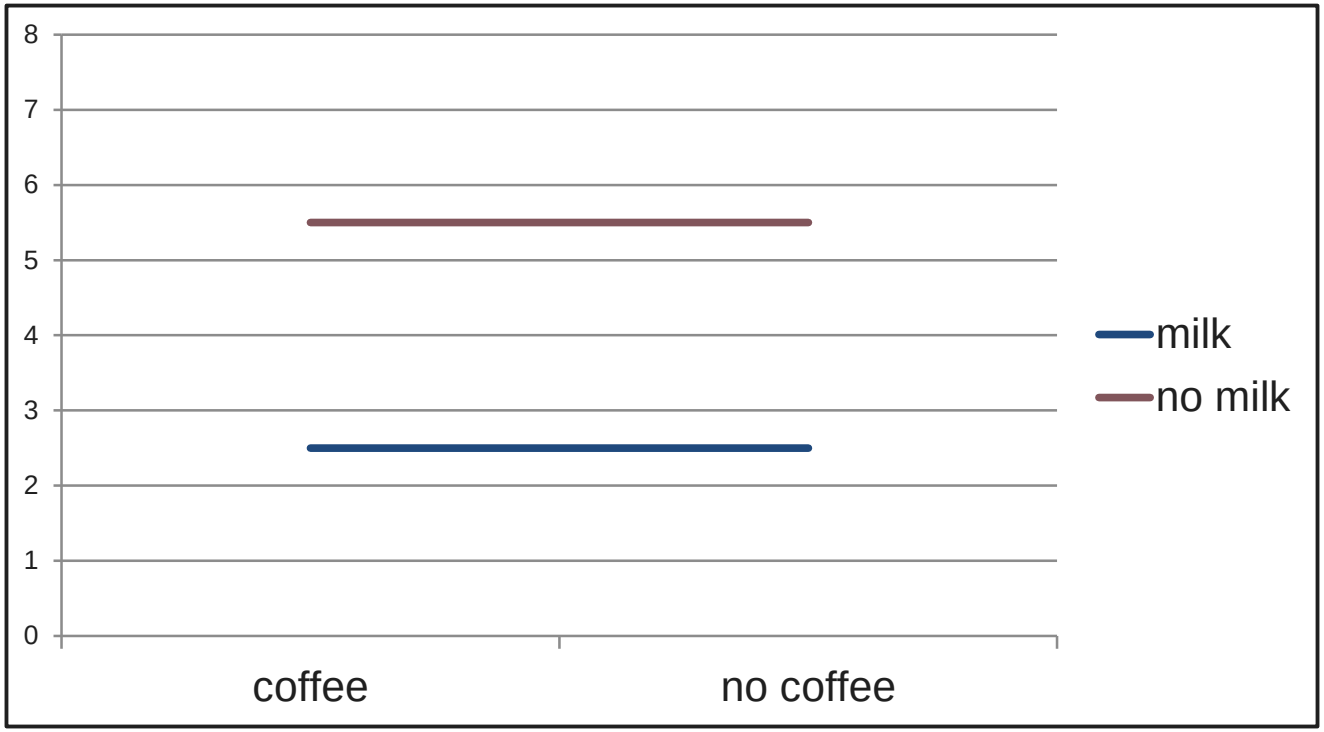


Milk factor: are the lines ends apart from each other? YES → Main effect Factor milk.

Coffee factor: are the lines inclined indicating a diff between the coffee factor levels? YES → Main effect Factor coffee.

Interaction effect: do the lines cross each other or are about to? YES
→ Interaction effect between factors milk-coffee

Main and interaction effects in Two-way ANOVA

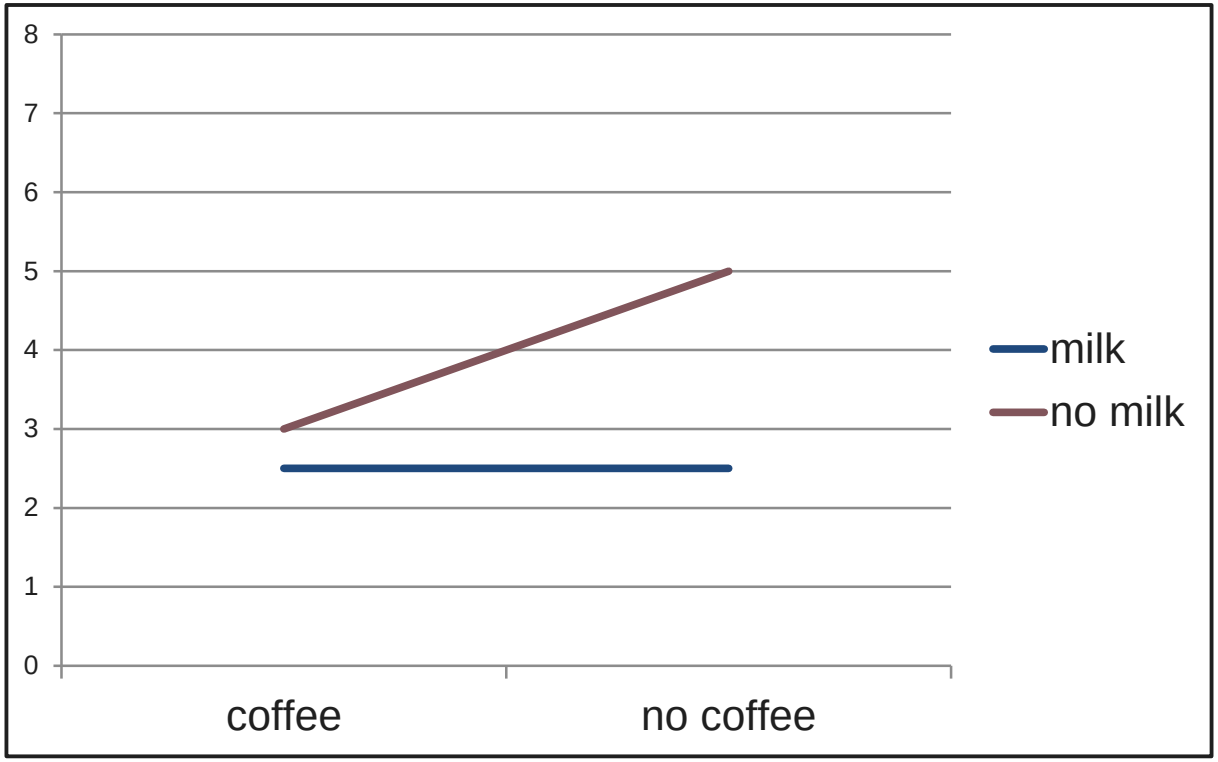


Milk factor: are the lines ends apart from each other? YES → Main effect Factor milk.

Coffee factor: are the lines inclined indicating a diff between the coffee factor levels? NO → No main effect Factor coffee.

Interaction effect: do the lines cross each other or are about to? NO
→ No interaction effect between factors milk-coffee

Main and interaction effects in Two-way ANOVA



Milk factor: are the lines apart from each other? Kinda → Posible main effect Factor milk.

Coffee factor: are the lines inclined indicating a diff between the coffee factor levels? Kinda → Possible main effect Factor coffee.

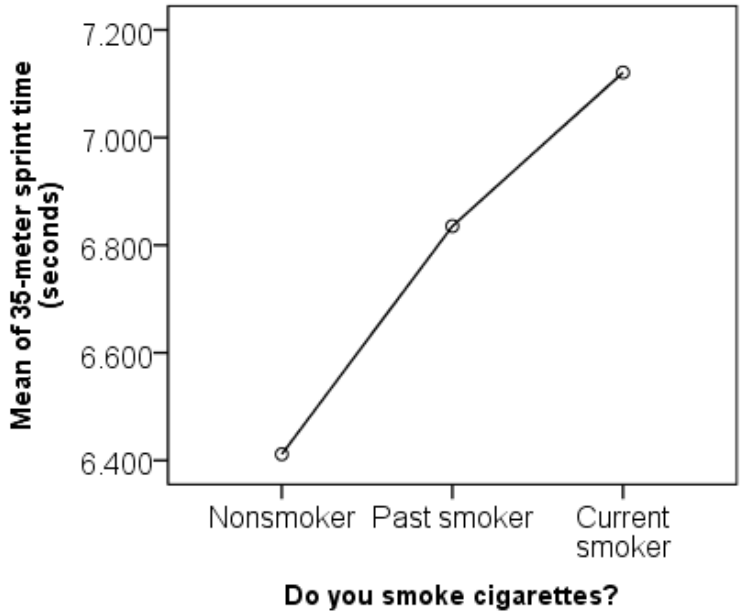
Interaction effect: do the lines cross each other or are about to? Almost→ Very likely interaction effect between factors milk-coffee
...that‘s why we always check the results table

ANOVA = “Omnibus“ test

- “Significant“ means:
  - There is a difference, somewhere. At least two groups differ significantly from each other.
- What remains unclear:
  - Where exactly is the difference? Which groups differ significantly from each other?

A priori and post-hoc ANOVA tests

- A priori / planned contrasts
  - If you have specific hypotheses, you can check them with contrasts
- Post-Hoc
  - You find something interesting in your data and want to analyze it in more detail afterwards



(You don‘t need to select and apply a priori or post-hoc test in the exam, but you need to be aware that they exist and when you would need one.)

Reporting ANOVA

"Comparing post-test means, we found a small negative significant main effect of individual preparation on knowledge outcomes, ***F*(1;124) = 5.121; *p* = .025; η_p^2 = .04**."

- *F* = 5.121 → F-value, result from the ANOVA
- (1;124) → df (n of groups minus 1; n of observations minus n of groups)
- *p* = .025 → result from the significance test
- η_p^2 = .04 → Effect size (next week)

Tsovaltzi, D., Judele, R., Puhl, T., Weinberger, A. (2015). Individual preparation and argumentation scripts in social networking sites. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine & S. Ludvigsen (Eds.). *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 2.* Gothenburg, Sweden: The International Society of the Learning Sciences, pp. 667-668.

ANOVA recap

- ANOVAs: belong to the F family of tests, therefore they produce a F-value. They are parametric tests normally used for determining if, based on their means, >2 samples are statistically significantly different
- One-way ANOVA: used when you have **only 1** factor/IV with >2 levels and **only 1** DV
- Two-way ANOVA: a.k.a. “Factorial ANOVA”, used when you have >1 factor (with multiple levels each) and **only 1** DV

ANOVA recap

		Independent Variables	
		1	>2
Dependent variables	1	One-way ANOVA	Factorial ANOVA
	>2	Multiple ANOVAS	MANOVA

Inspired from the table found at: https://www.statisticssolutions.com/conduct-interpret-factorial-anova/

Exercise 1: Two-way ANOVA. Problem description

Using Beispieldatensatz_S9, test the following (alternative*) hypotheses:

a) HA_1: There is an interaction effect between gender and condition on the number of remembered negative adjectives (negativ).

b) HA_2: There is a main effect of gender (sex) on the number of remembered negative adjectives (negativ).

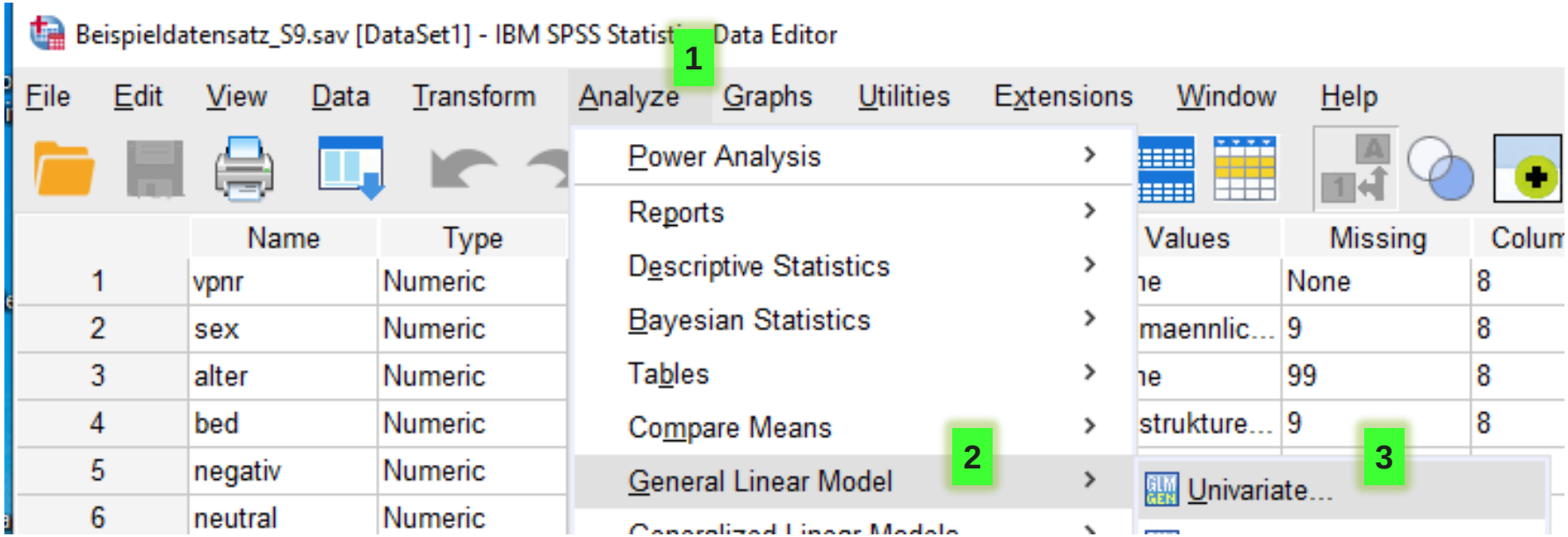
c) HA_3: There is a main effect of condition (bed) on the number of remembered negative adjectives (negativ).

*This means that for each HA, there is a corresponding H0

Exercise 1: Two-way ANOVA. Why not one-way?

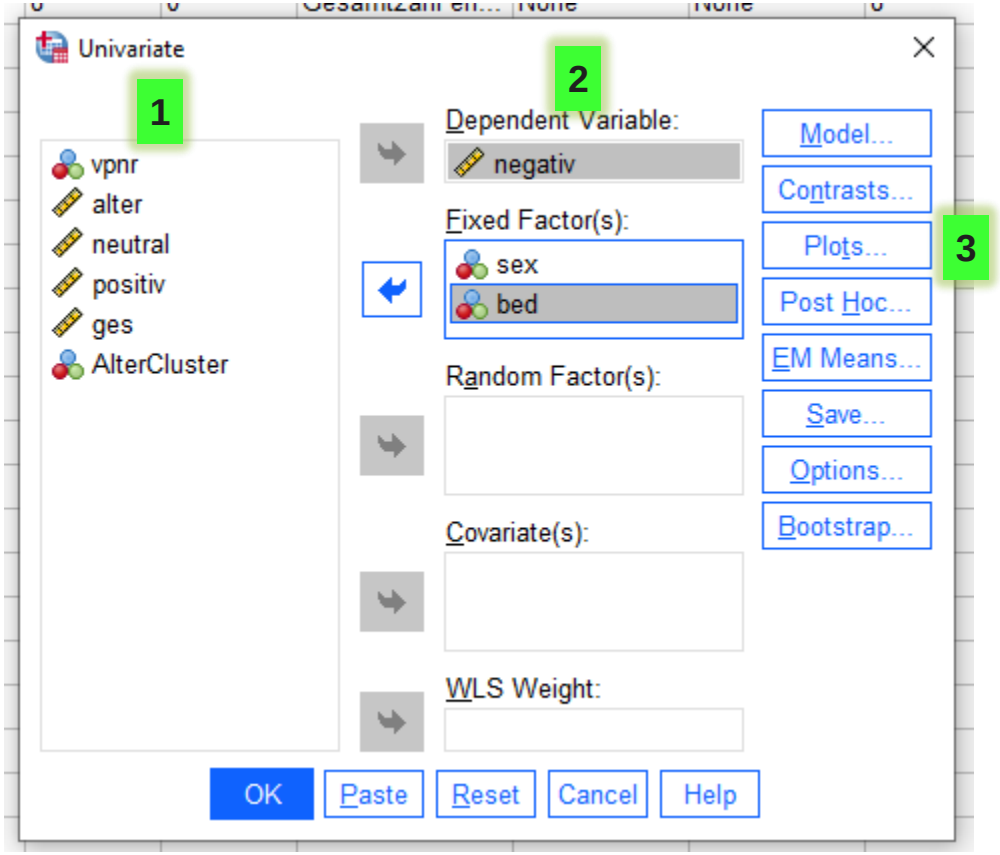
- We have 3 hypotheses: 2 main effect ones, 1 interaction effect one
  - Therefore, we have 2 IV (a.k.a. factors)
  - One-way ANOVA is only when you have 1 IV

Exercise 1: Two-way ANOVA. On SPSS



Exercise 1: Two-way ANOVA. On SPSS

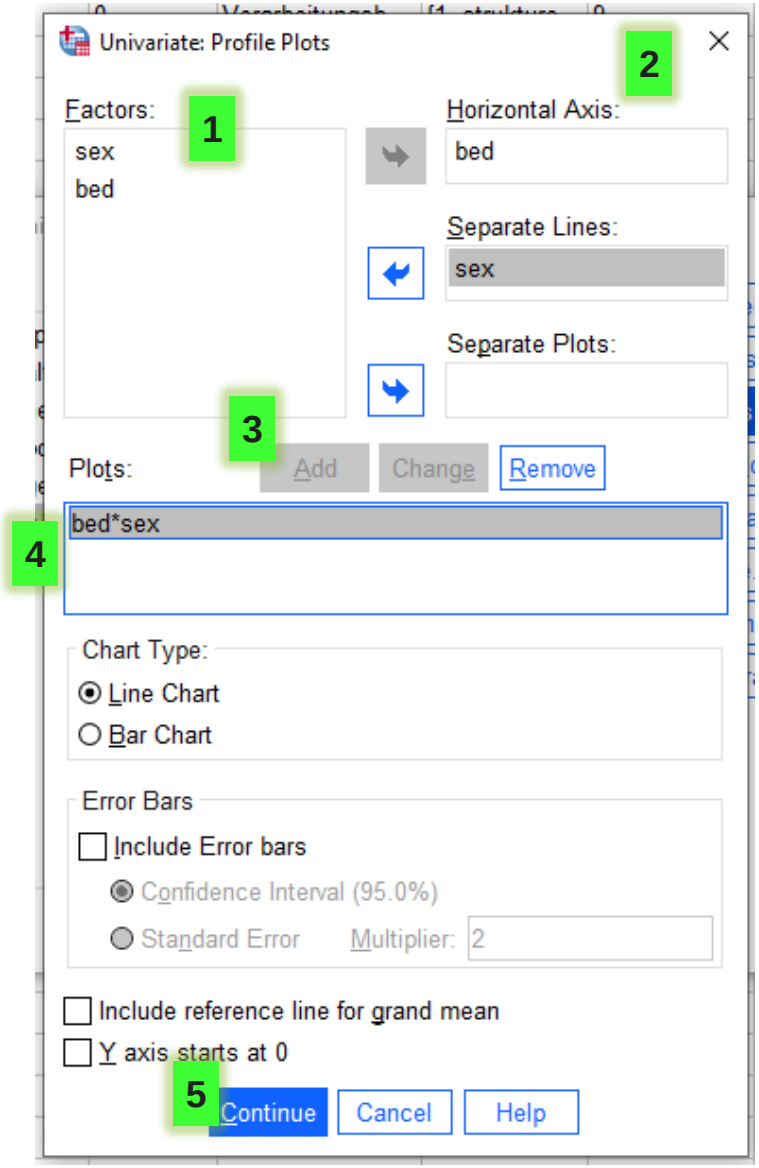
In the “Univariate” window, allocate the necessary variables from the left box [1] to the corresponding box in the right side [2]. Once completed, click on the “Plots” button [3]



Exercise 1: Two-way ANOVA. On SPSS

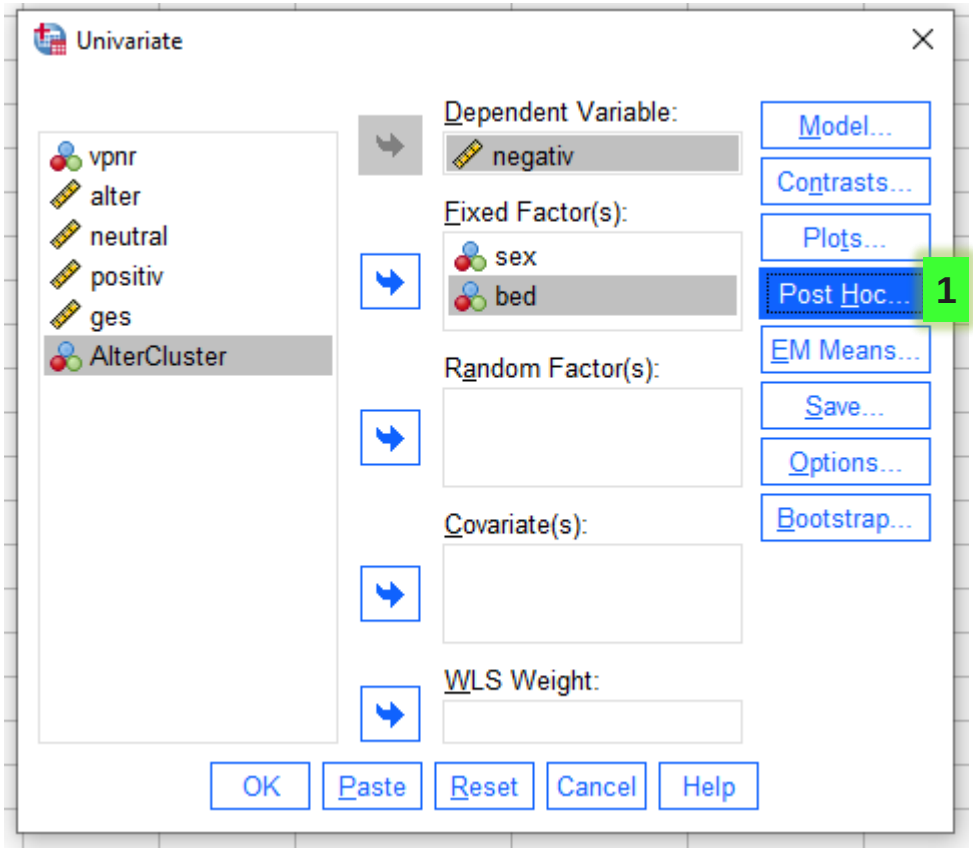
In the “Univariate: Profile plots” window, allocate the necessary variables from the left box [1] to the corresponding box in the right side [2]. Once completed, click on the “Add” button [3] which will show the to-be-created plot in the box below [4]. Ensure the other options are defined as shown.
Click on the “Continue” button [5] to close the window and go back to the main “Univariate” window.

Note: for extra fun, also Add a “sex*bed” plot



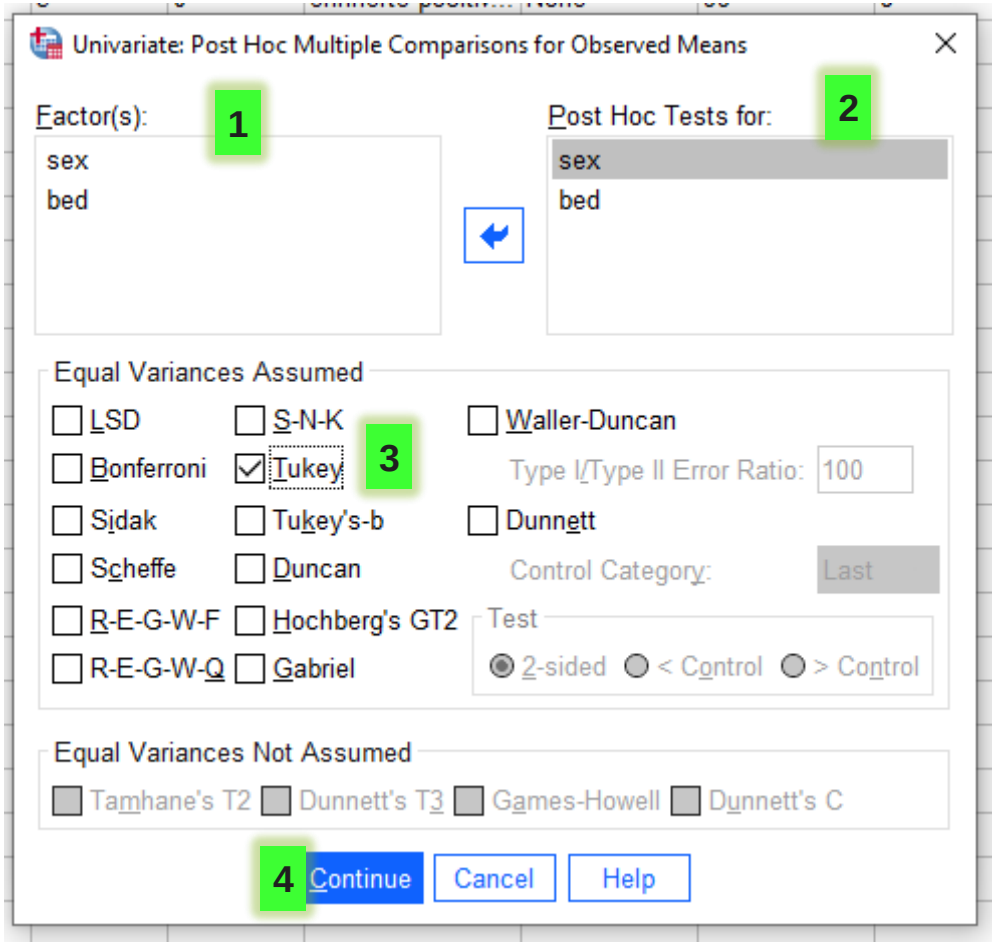
Exercise 1: Two-way ANOVA. On SPSS

Once back in the main “Univariate” window, click on the “Post-Hoc” button



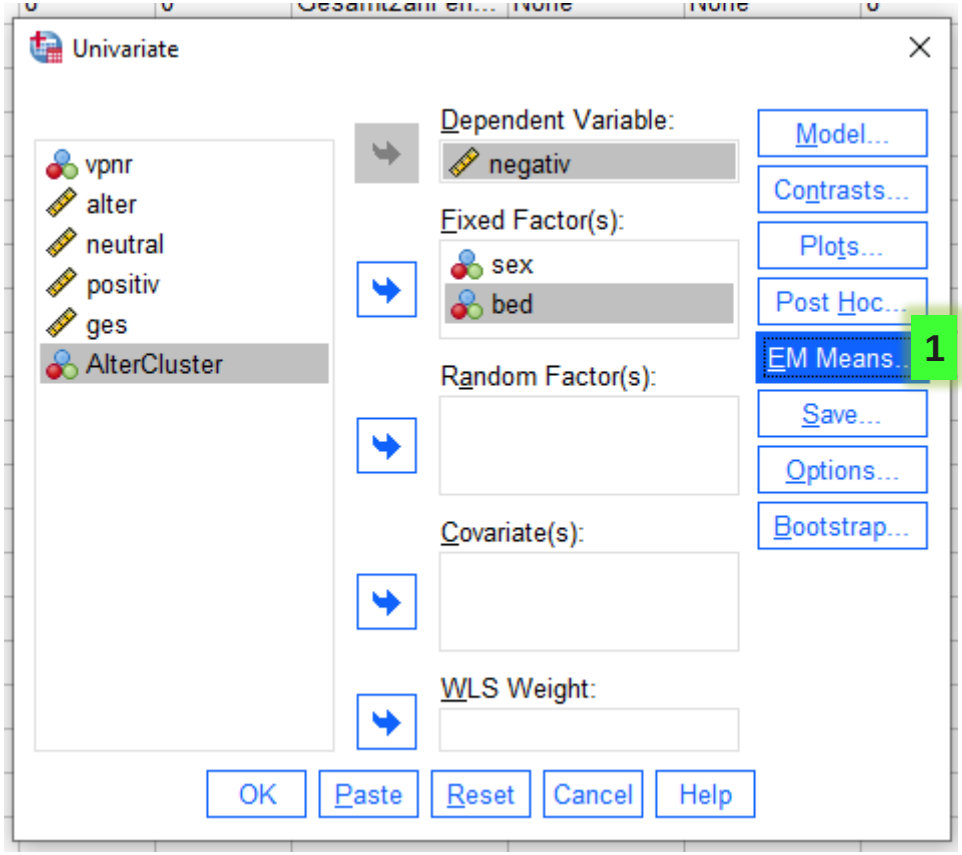
Exercise 1: Two-way ANOVA. On SPSS

In the “Univariate: Post-Hoc Multiple Comparisons for Observed Means” window, allocate the necessary variables from the left box [1] to the corresponding box in the right side [2]. Once completed, mark the “Tukey” checkbox [3]. Ensure the other options are defined as shown.
Click on the “Continue” button [4] to close the window and go back to the main “Univariate” window.



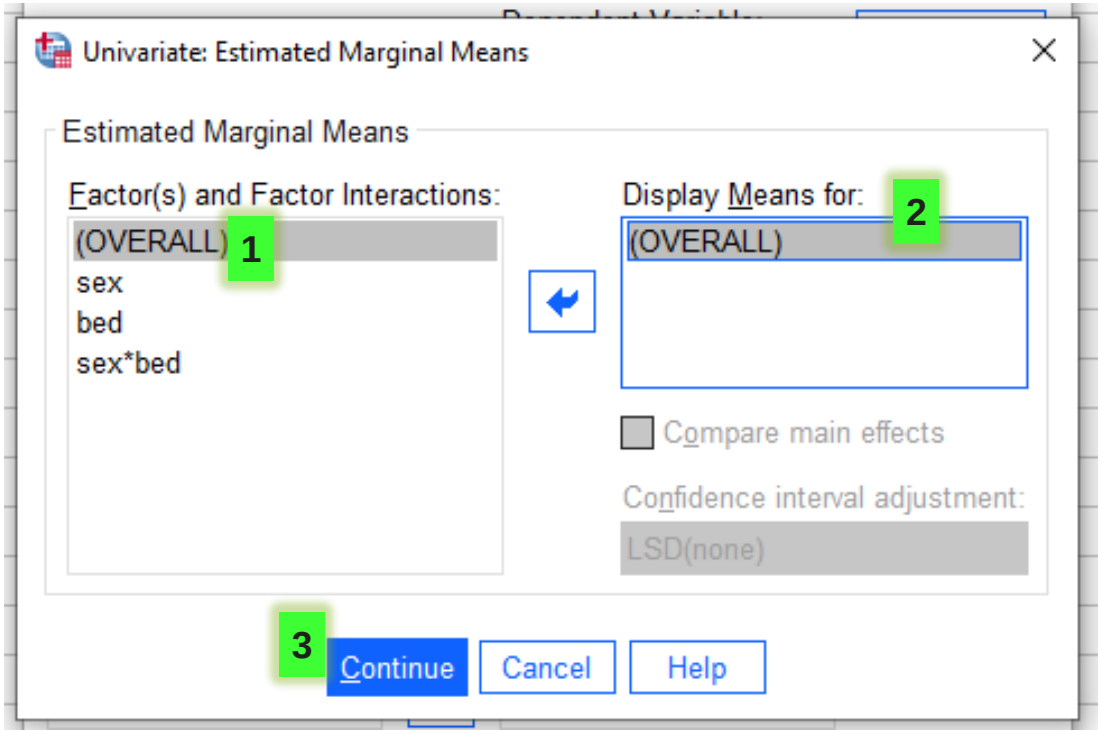
Exercise 1: Two-way ANOVA. On SPSS

Once back in the main "Univariate" window, click on the "EM Means" button



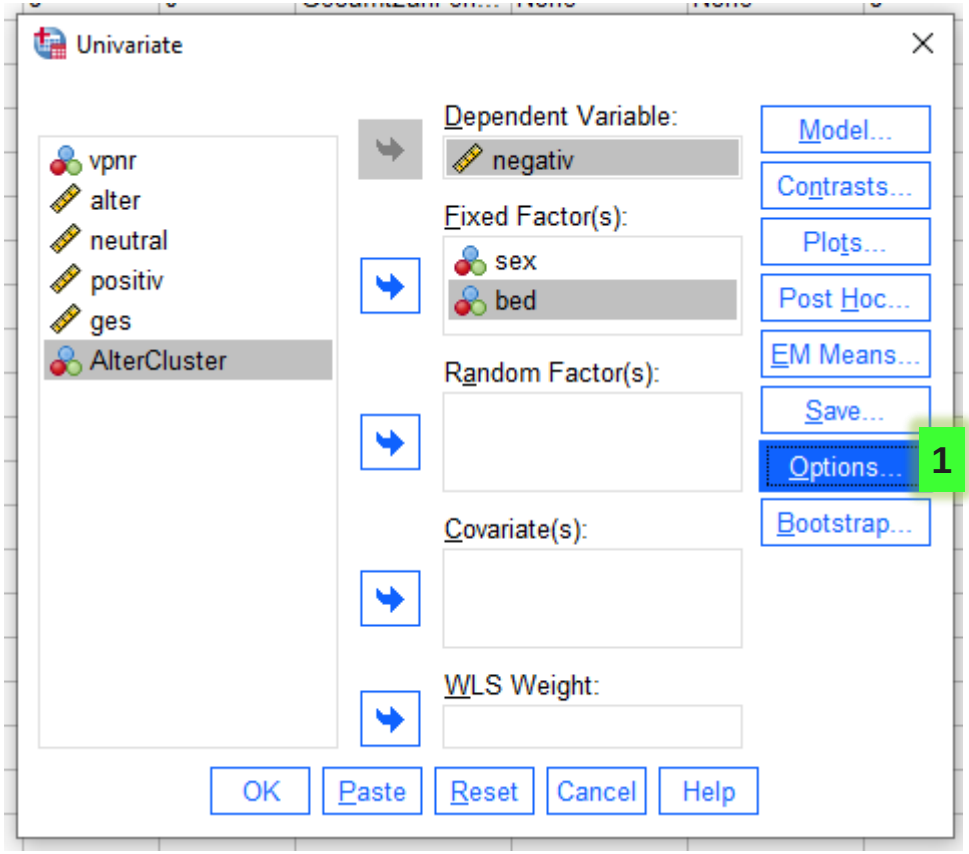
Exercise 1: Two-way ANOVA. On SPSS

In the “Univariate: Estimated Marginal Means” window, allocate only the “(OVERALL)” factor located in the left box [1] to the box in the right side [2]. Click on the “Continue” button [3] to close the window and go back to the main “Univariate” window.



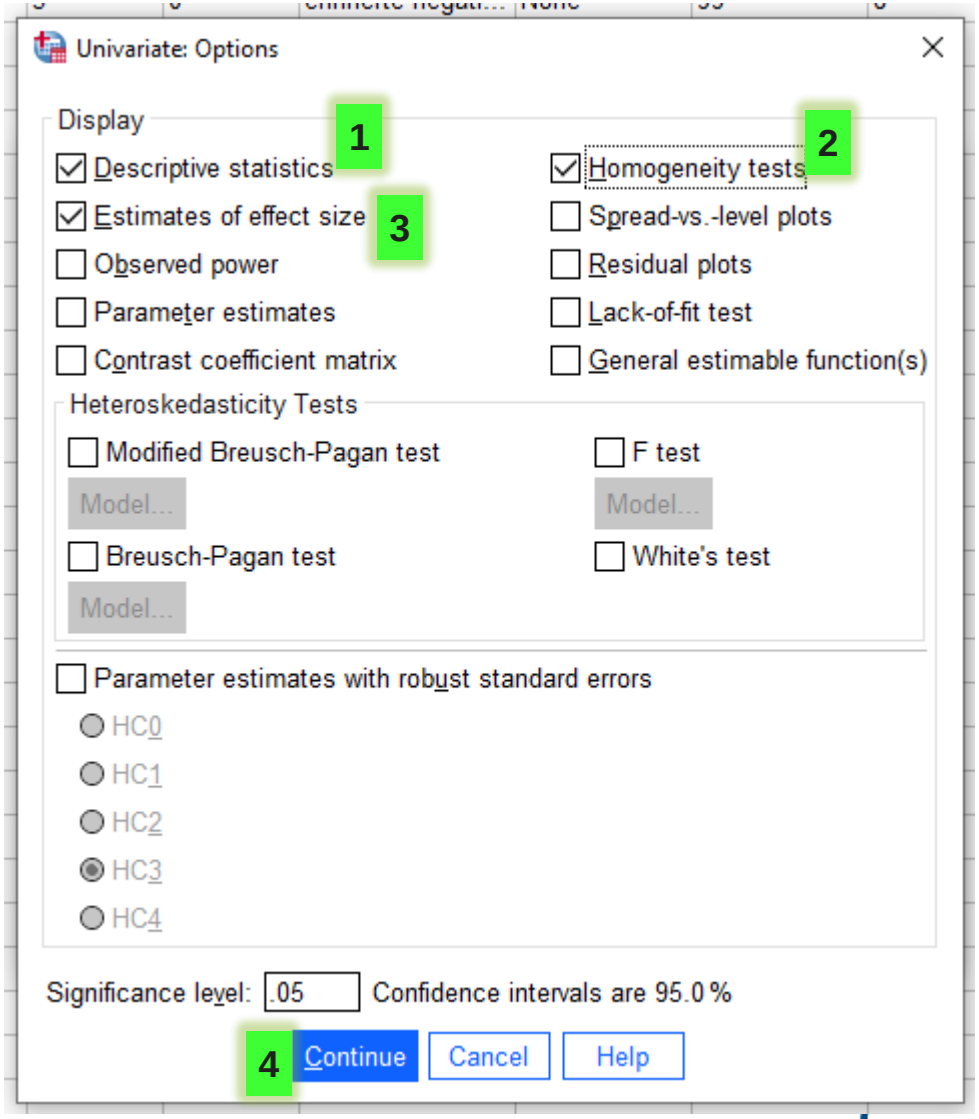
Exercise 1: Two-way ANOVA. On SPSS

Once back in the main "Univariate" window, click on the "Options" button



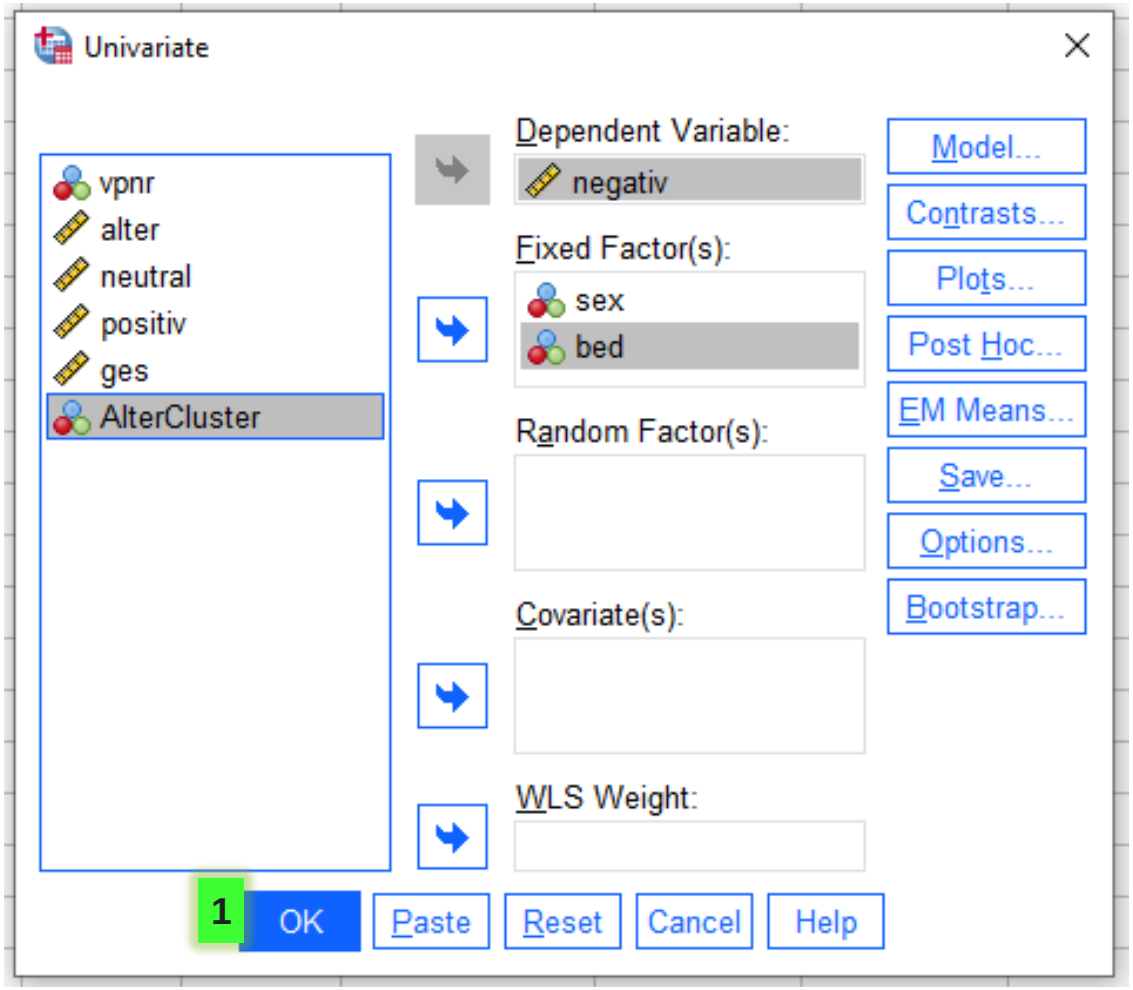
Exercise 1: Two-way ANOVA. On SPSS

In the “Univariate: Options” window, mark the indicated checkboxes [1], [2], and [3]. Leave the rest unchecked. Click on the “Continue” button [4] to close the window and go back to the main “Univariate” window.



Exercise 1: Two-way ANOVA. On SPSS

Once finally back in the main “Univariate” window, click on the “OK” button to start the analysis.



Exercise 1: Two-way ANOVA. Output and reporting

Tests of Between-Subjects Effects

Dependent Variable: negativ

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	41,917[a]	5	8,383	2,403	,040	,077
Intercept	1478,115	1	1478,115	423,700	,000	,746
sex	2,021	1	2,021	,579	,448	,004
bed	31,352	2	15,676	4,494	,013	,059
sex * bed	1,449	2	,725	,208	,813	,003
Error	502,357	144	3,489			
Total	2231,000	150				
Corrected Total	544,273	149				

a. R Squared = ,077 (Adjusted R Squared = ,045)

A

B

A * B

No sign. main effect of gender (sex): $F(1;144) = .579$, $p = .448$, $\eta_p^2 = .004$

Sign. main effect of condition (bed): $F(2;144) = 4.494$, $p = .013$, $\eta_p^2 = .059$

No sign. interaction between sex and bed: $F(2;144) = .208$, $p = .813$, $\eta_p^2 = .003$

…on the DV “negativ”

Exercise 1: Two-way ANOVA. Output and reporting

Interaction effect:

No sig. interaction between gender (sex) and conditon (bed): $F(2;144) = .208$, $p = .813$, $\eta_p^2 = .003$

Main effects:

No sig. main effect of gender (sex): $F(1;144) = .579$, $p = .448$, $\eta_p^2 = .004$

Sig. main effect of condition (bed): $F(2;144) = 4.494$, $p = .013$, $\eta_p^2 = .059$

This means that within the factors of the “Bed” variable, 1 or more of its levels should be sig. different to each other. To see this more in detail, we need to check the post-hoc results.

Exercise 1: Two-way ANOVA. Output and reporting

Multiple Comparisons

Dependent Variable: erinnerte negative Adjektive

Tukey HSD

(I) Verarbeitungsbedingung	(J) Verarbeitungsbedingung	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
strukturell	bildhaft	-1.10*	.374	.010	-1.98	-.22
	emotional	-1.04*	.374	.017	-1.92	-.16
bildhaft	strukturell	1.10*	.374	.010	.22	1.98
	emotional	.06	.374	.986	-.82	.94
emotional	strukturell	1.04*	.374	.017	.16	1.92
	bildhaft	-.06	.374	.986	-.94	.82

Based on observed means.
The error term is Mean Square(Error) = 3.489.

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

Verarbeitungsbedingung (bed) factor has 3 levels: strukturell, bildhaft, and emotional.

The mean of each level is compared to each other to look for significance, hence the 6 comparisons. Half of the comparisons are the same, therefore we focus on 3 of them.

Exercise 1: Two-way ANOVA. Output and reporting

Multiple Comparisons

Dependent Variable: erinnerte negative Adjektive

Tukey HSD

(I) Verarbeitungsbedingung	(J) Verarbeitungsbedingung	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
strukturell	bildhaft	-1.10*	.374	.010	-1.98	-.22
	emotional	-1.04*	.374	.017	-1.92	-.16
bildhaft	strukturell	1.10*	.374	.010	.22	1.98
	emotional	.06	.374	.986	-.82	.94
emotional	strukturell	1.04*	.374	.017	.16	1.92
	bildhaft	-.06	.374	.986	-.94	.82

Based on observed means.
The error term is Mean Square(Error) = 3.489.

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

From these post-hoc results we can conclude that, from the IV "Verarbeitungsbedingung", the levels:
Strukturell and Pictographic are significantly different (p = .010)
Strukturell and Emotional are significantly different (p = .017)
…in terms of the DV (remembered negative-adjectives)

Exercise 1: Two-way ANOVA. Problem description

Using Beispieldatensatz_S9, test the following (alternative) hypotheses:

a) HA_1: There is an interaction effect between gender and condition on the number of remembered negative adjectives. → Fail to reject $H0_1$

b) HA_2: There is a main effect of gender (sex) on the number of remembered negative adjectives (negativ). → Fail to reject $H0_2$

c) HA_3: There is a main effect of condition (bed) on the number of remembered negative adjectives (negativ). → Reject $H0_3$, accept HA_3

   a) Structural and Pictographic conditions are sig. diff. (p = .010)

   b) Structural and Emotional are significantly different (p = .017)

Extra resources

- http://wwwstage.valpo.edu/other/dabook/ch13/c13-1.htm
- https://www.statisticssolutions.com/conduct-interpret-factorial-anova/

SELF-PACED EXERCISES

Three self-paced exercises about ANOVAS including their answers and explanations

Self-paced Exercise 1. Problem description

Using Beispieldatensatz_S9, conduct an ANOVA to test the following hypotheses:

a) There is a main effect of age (AlterCluster) on the number of remembered neutral adjectives (neutral).

b) There is a main effect of condition (bed) on the number of remembered neutral adjectives (neutral).

c) There is an interaction effect between gender and condition on the number of remembered neutral adjectives.

Self-paced Exercise 1. Tasks

1. Which ANOVA do you need to conduct and why? One-way or Two-way?
2. On SPSS, why do you need to use the variable “AlterCluster” and not the “alter” variable?
3. Outline the factor(s) and their levels
4. Report the results

Note: next 2 slides contain the answers for the 3 tasks including the output and reported results

Self-paced Exercise 1. Task answers

1. Which ANOVA do you need to conduct and why? One-way or Two-way?

A1: A two-way ANOVA needs to be done because there are more than 2 factors that we will compare and we're interested on finding interaction effects

2. On SPSS, why do you need to use the variable "AlterCluster" and not the "alter" variable?

A2: Because "alter" is a continuous scale variable and with age being a factor in this study, we need a categorical variable

3. Outline the factor(s) and their levels:

A3: Factor 1: age (AlterCluster). Levels: 4 [1. "up to 19 years"; 2. "20-29 years"; 3. "30-39" years; 4. "40-49 years"]

Factor 2: Verarbeitungsbedingung (bed). Levels: 3 [1. strukturell (structural); 2. bildhaft (pictorial); 3. emotional (emotional)]

4. Report the results

A4: On the next slide

Self-paced Exercise 1. Output and task 4 answer

Tests of Between-Subjects Effects

Dependent Variable: erinnerte neutrale Adjektive

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	149.472[a]	7	21.353	7.039	.000	.258
Intercept	102.204	1	102.204	33.691	.000	.192
bed	57.132	2	28.566	9.417	.000	.117
AlterCluster	17.217	3	5.739	1.892	.134	.038
bed * AlterCluster	.510	2	.255	.084	.919	.001
Error	430.768	142	3.034			
Total	2194.000	150				
Corrected Total	580.240	149				

a. R Squared = .258 (Adjusted R Squared = .221)

There is a sign. main effect of bed on the DV “neutral”:
$F(2, 142) = 9.417$, $p = .000$, $\eta_p^2 = .117$
There is no sign. main effect of AlterCluster on the DV “neutral”:
$F(3, 142) = 1.892$, $p = .134$, $\eta_p^2 = .038$
There is no sign. interaction effect of bed*AlterCluster on the DV “neutral”:
$F(2, 142) = 0.084$, $p = .919$, $\eta_p^2 = .001$

Self-paced Exercise 2. Tasks

Next slide contains an ANOVA output. Based on that output, respond the following questions:

1. Determine if it's the output from a one-way or two-way ANOVA and explain why
2. Determine the factor and its levels, and the DV
3. Report the results
4. Determine if a post-hoc test would be needed and why.

Self-paced Exercise 2

Report the results:

Tests of Between-Subjects Effects

Dependent Variable: neutral

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	23,215[a]	3	7,738	2,028	,113	,040
Intercept	50,743	1	50,743	13,300	,000	,083
AlterCluster	23,215	3	7,738	2,028	,113	,040
Error	557,025	146	3,815			
Total	2194,000	150				
Corrected Total	580,240	149				

a. R Squared = ,040 (Adjusted R Squared = ,020)

Beware: Next 2 slides contain the answers and interpreted results!

Self-paced Exercise 2. Task answers

Next slide contains an ANOVA output. Based on that output, respond the following questions:

1. Determine if it's the output from a one-way or two-way ANOVA and explain why

A1: it's the output from a one-way ANOVA because there's only 1 row showing 1 factor and there's no row with an interaction effect (e.g. "IV1*IV2")

2. Determine the factor and its levels, and the DV

A2: Factor 1: AlterCluster. Levels: they are not indicated in the provided output screenshot. Assuming it's the same dataset as the one from the self-paced exercise 1, the levels could be: 4 [1. "up to 19 years"; 2. "20-29 years"; 3. "30-39" years; 4. "40-49 years"]. The dependent variable is called "neutral"

3. Report the results

A3: On the next slide

4. Determine if a post-hoc test would be needed and why.

A4: Due to the non-sig. results (p = .113), a post-hoc test would not be needed

Self-paced Exercise 2. Output and task 3 answer

Tests of Between-Subjects Effects

Dependent Variable: neutral

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	23,215[a]	3	7,738	2,028	,113	,040
Intercept	50,743	1	50,743	13,300	,000	,083
AlterCluster	23,215	3	7,738	2,028	,113	,040
Error	557,025	146	3,815			
Total	2194,000	150				
Corrected Total	580,240	149				

a. R Squared = ,040 (Adjusted R Squared = ,020)

There is no sign. main effect of AlterCluster on the DV “neutral”:
$F(3, 146) = 2.028$, $p = .113$, $\eta_p^2 = .040$

Self-paced Exercise 3. Tasks

Next slide contains an ANOVA output. Based on that output, respond the following questions:

1. Determine if it's the output from a one-way or two-way ANOVA and why
2. Determine the factor and its levels, and the DV
3. Report the results
4. Determine if a post-hoc test would be needed and why.

Self-paced Exercise 3

Report the results:

Tests of Between-Subjects Effects

Dependent Variable: Post_Qu_6

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	7,448[a]	3	2,483	2,562	,062	,098
Intercept	254,763	1	254,763	262,916	,000	,787
hand_N	,158	1	,158	,163	,687	,002
Own_MT	2,318	1	2,318	2,392	,126	,033
hand_N * Own_MT	5,609	1	5,609	5,788	,019	,075
Error	68,798	71	,969			
Total	1382,500	75				
Corrected Total	76,247	74				

a. R Squared = ,098 (Adjusted R Squared = ,060)

Beware: Next 2 slides contain the answers and interpreted results!

Self-paced Exercise 3. Task answers

Next slide contains an ANOVA output. Based on that output, respond the following questions:

1. Determine if it's the output from a one-way or two-way ANOVA and why

A1: It's the output from a two-way ANOVA because we can see 2 factor rows as well as a interaction effect row

2. Determine the factor and its levels, and the DV

A2: Factor 1: hand_N. Factor 2: Own_MT; DV: Post_Qu_6. The levels of the factors are not shown in the provided output screenshot

3. Report the results

A3: On the next slide

4. Determine if a post-hoc test would be needed and why.

A4: No, even though a statistically sig. interaction effect was obtained, the next step is to conduct a one-way ANOVA to further analyze the interaction effect. This procedure is out of scope of the course but if you're interested you can follow this tutorial: https://www.spss-tutorials.com/spss-two-way-anova-interaction-significant/

Self-paced Exercise 3 task 3 answers

Report the results:

Tests of Between-Subjects Effects

Dependent Variable: Post_Qu_6

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	7,448[a]	3	2,483	2,562	,062	,098
Intercept	254,763	1	254,763	262,916	,000	,787
hand_N	,158	1	,158	,163	,687	,002
Own_MT	2,318	1	2,318	2,392	,126	,033
hand_N * Own_MT	5,609	1	5,609	5,788	,019	,075
Error	68,798	71	,969			
Total	1382,500	75				
Corrected Total	76,247	74				

a. R Squared = ,098 (Adjusted R Squared = ,060)

- There is no sign. main effect of hand_N on the DV “Post_Qu_6”: F (1, 71) = .163, p = .687, η_p^2 = .002
- There is no sign. main effect of Own_MT on the DV “Post_Qu_6”: F (1, 71) = 2.392, p = .126, η_p^2 = .033
- There is a sign. interaction between hand_N and Own_MT on the DV “Post_Qu_6”: F(1, 71) = 5.788, p = .019, η_p^2 = .075

Qualis

- Official course evaluation survey from UdS
- Completely anonymous
- Evaluation link: https://qualis.uni-saarland.de/eva/?l=127357&p=pynh61