Empirical Research Methods
Repetition

Lara Kataja

Room 1.13

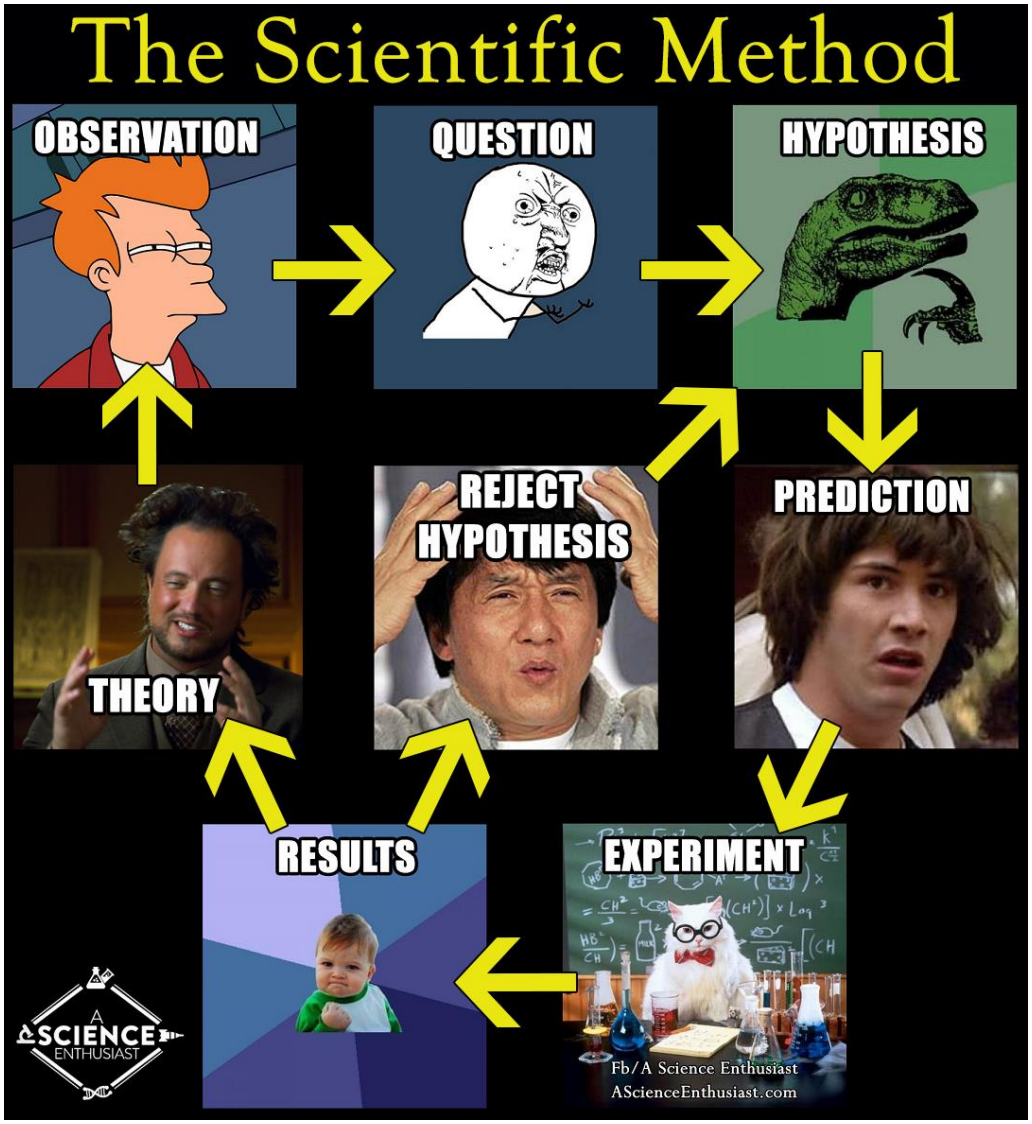
l.kataja@edutech.uni-saarland.de

Tutoring

- Interested in becoming the next tutor?

→ Email me

Basic concepts of experimental research

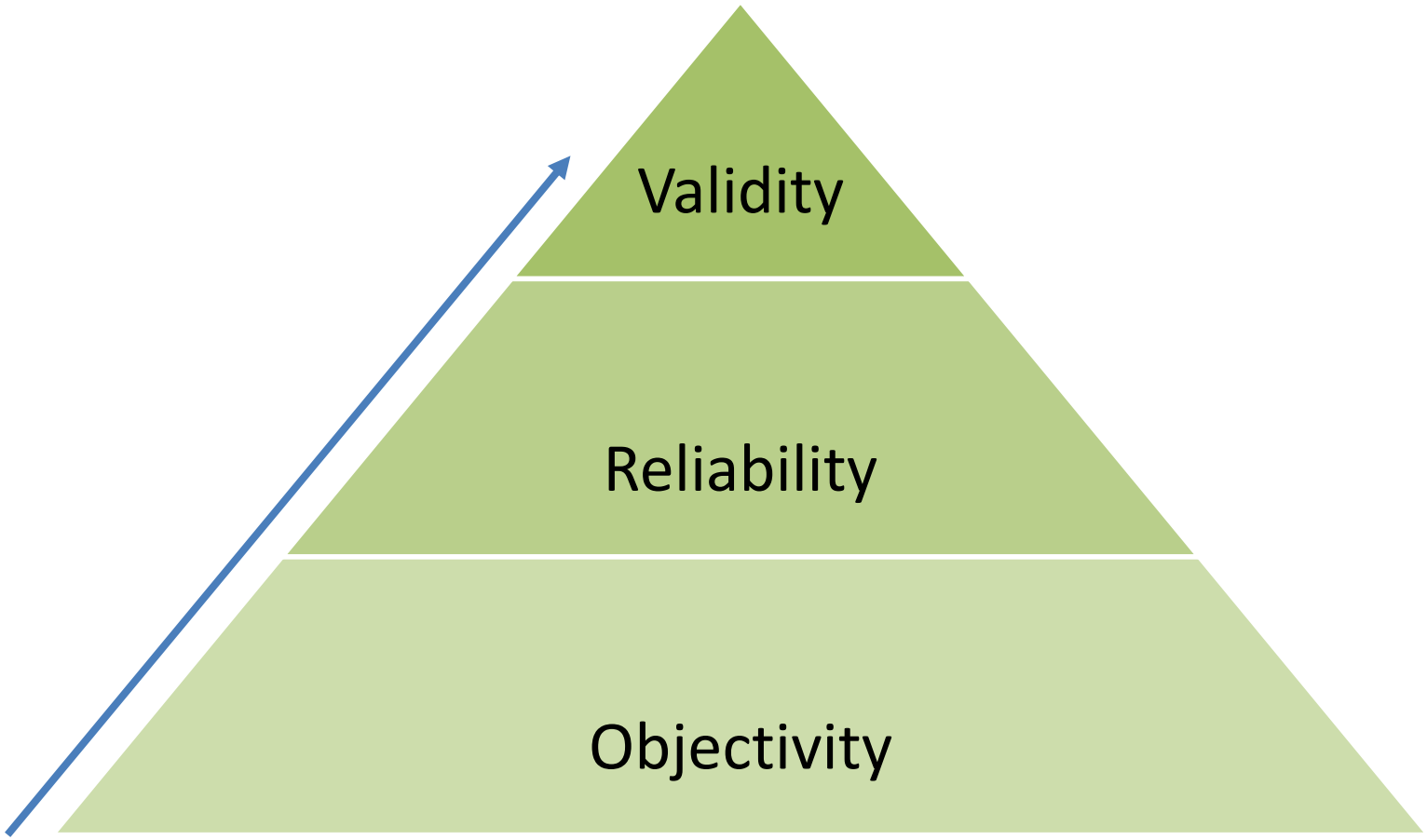


https://ascienceenthusiast.com/steps-of-the-scientific-method/

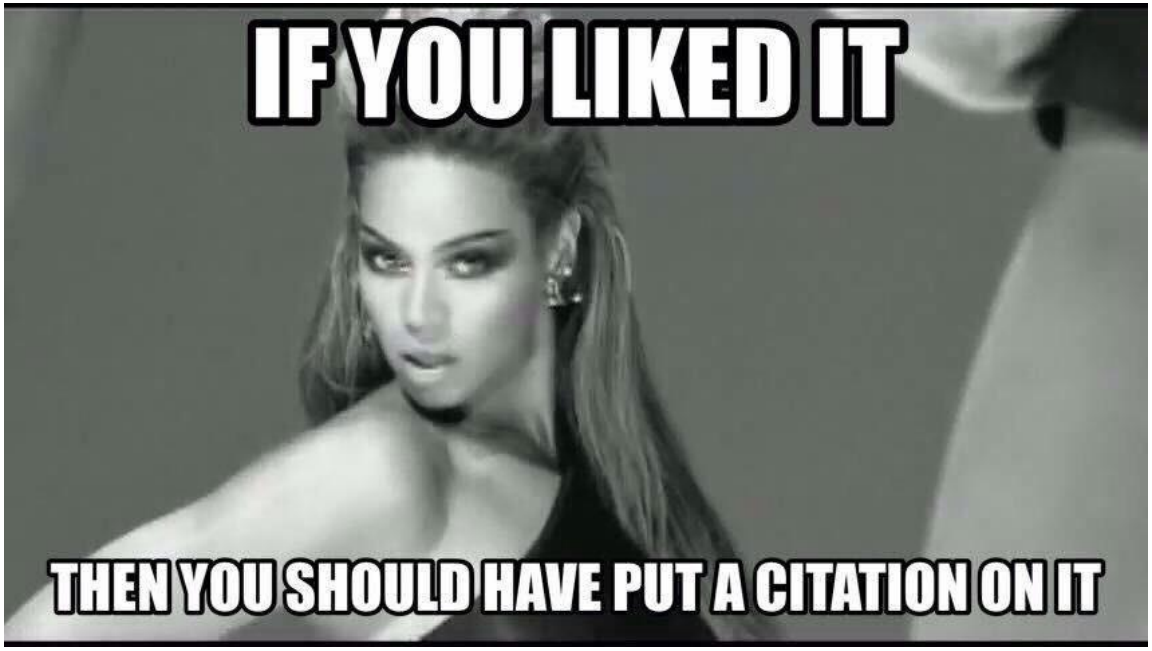
Explain:

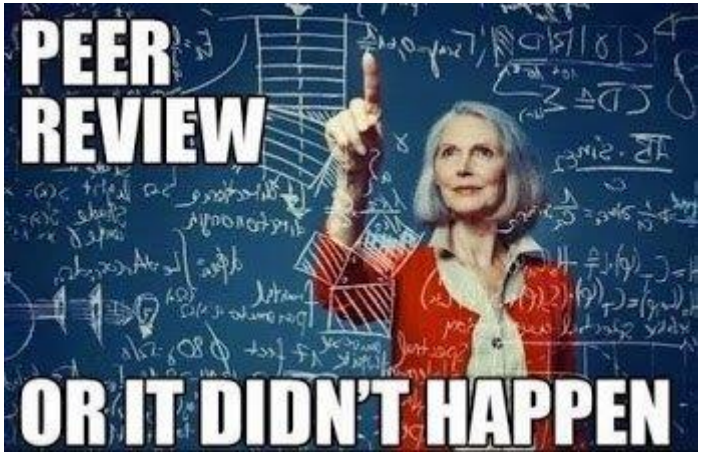
- H0 vs. H1
- Sample vs. population
- Dependent vs. independent variable
- Levels of measurement
- Lab vs. field experiment
- Independent vs. dependent (paired) samples (between vs. within)
- Factors and factor levels

Quality criteria



Plagiarism and literature search





Good sources for your scientific work:

- Peer reviewed journal articles, books and book chapters, conference papers
- You can start your search e.g. here:
  - https://www.sciencedirect.com/
  - https://scholar.google.de/
  - https://www.sulb.uni-saarland.de/

https://memegenerator.net/Peer-Review-Or-It-DidnT-Happen
https://www.arthurterry.bham.sch.uk/library/citations/

Questionnaires and tests

Explain:

- Open vs. closed item format
- How to formulate good items
- Using a validated questionnaire vs. constructing your own

SPSS

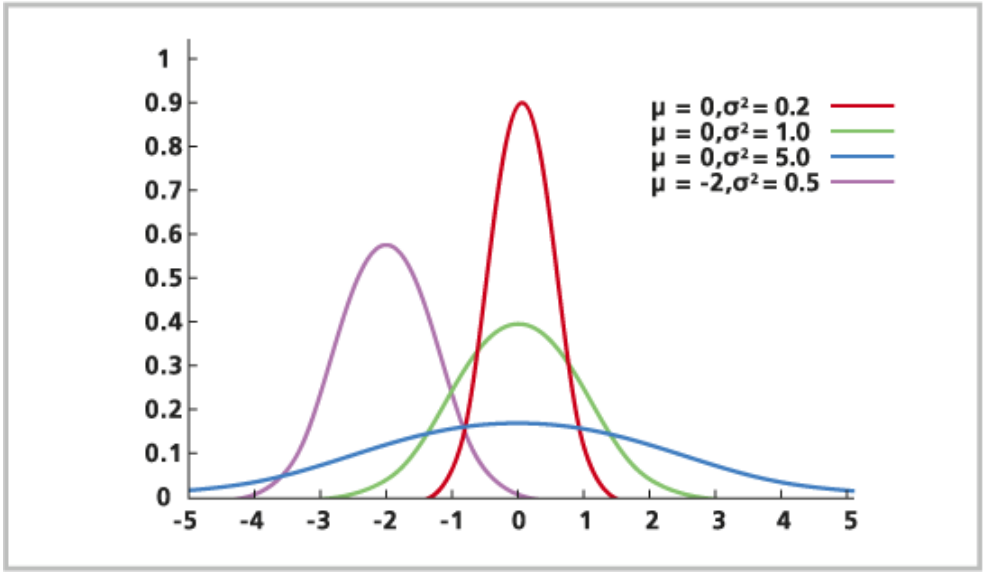
Explain:

- Variable vs. data view
- Syntax
- How to name variables
- Missing values

Descriptive statistics

Explain:

- Measures of central tendency:
  - Median
  - Mode
  - Mean
- Measures of dispersion:
  - Standard deviation
  - Variance
  - (…)



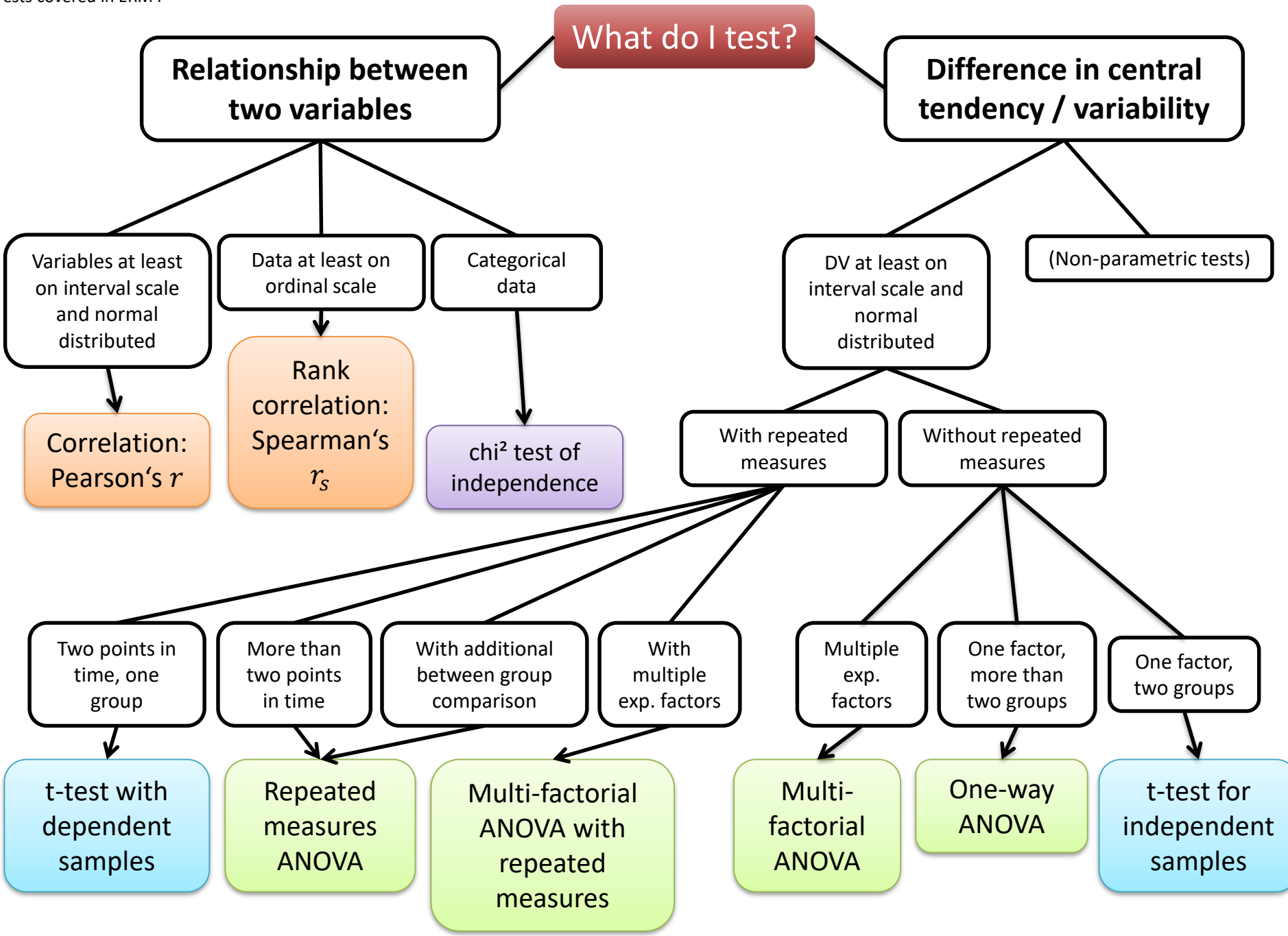
http://memebomb.net/standard-deviation-memes-1/
http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/process_simulations_sensitivity_analysis_and_error_analysis_modeling/distributions_for_assigning_random_values.htm

What is my research like?

- Descriptive
- Inferential
- Qualitative
- Quantitative

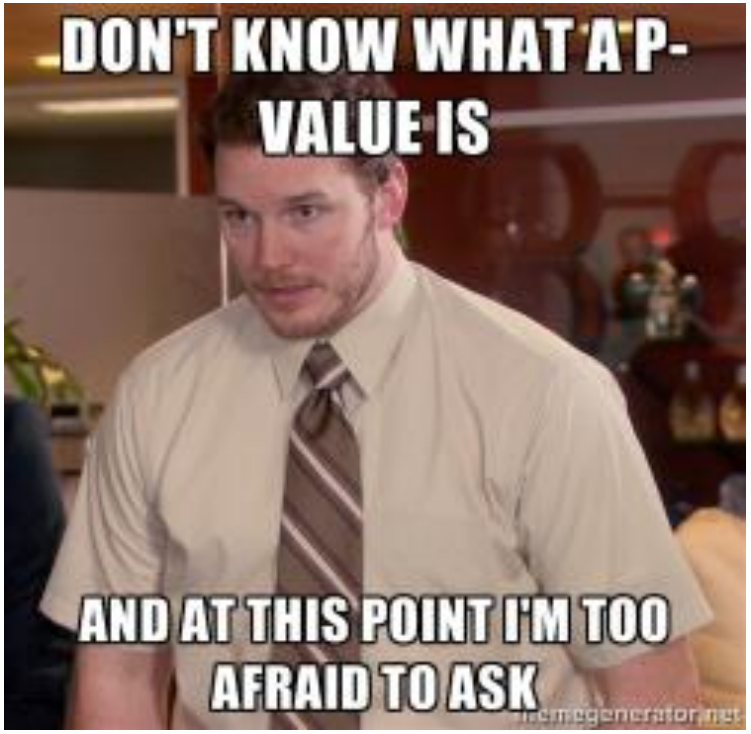
→ Explain the differences

Tests covered in ERM I



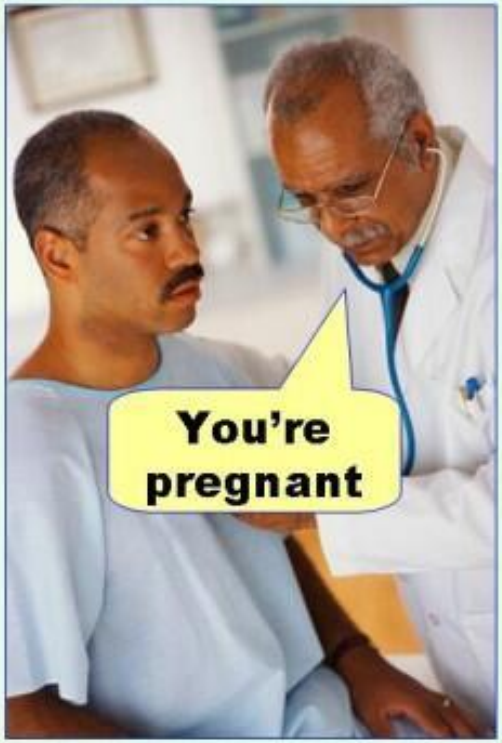
Understanding *p* and α

- *p* = probability of our data assuming that H_0 is true
- α level (= .05) sets the threshold for keeping or rejecting the H_0
- If $p \leq \alpha \rightarrow$ reject H_0 and assume H_1
- If $p > \alpha \rightarrow$ keep the H_0

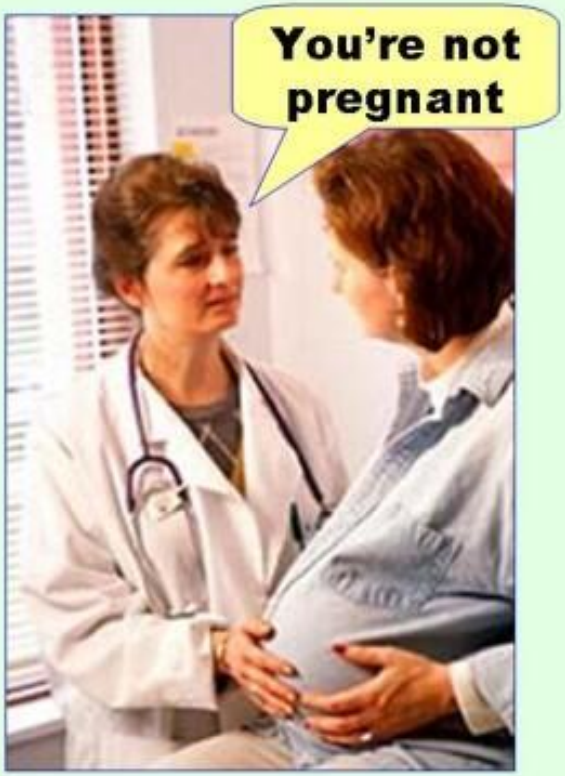


Decision errors

Type I error
(false positive)

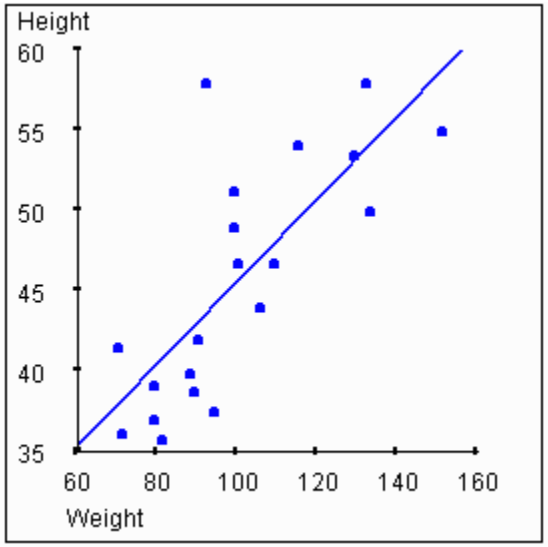


Type II error
(false negative)



https://stats.stackexchange.com/questions/1610/is-there-a-way-to-remember-the-definitions-of-type-i-and-type-ii-errors

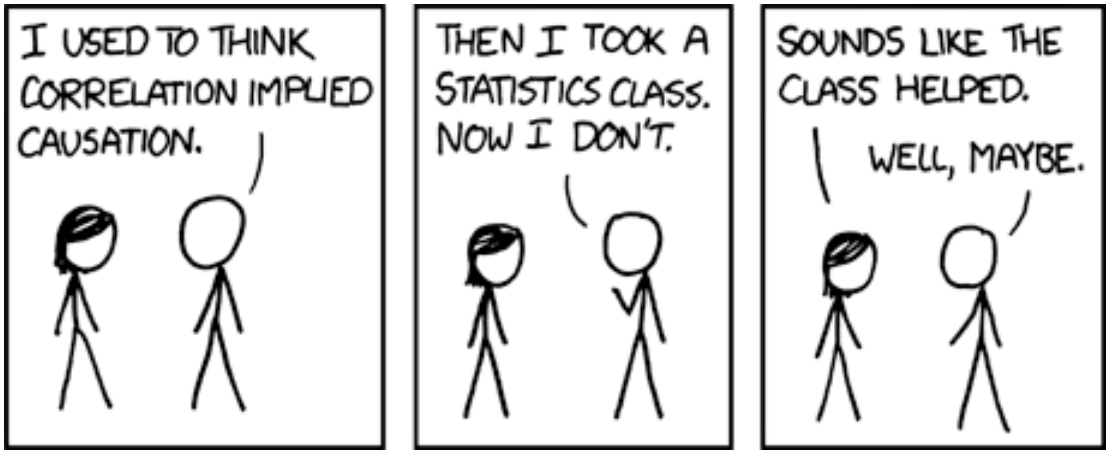
Correlation



Explain:

- What does ‘correlation’ mean?
- Name two different correlation measures
  - Spearman
  - Pearson

  → When do you use which one?



https://simon.cs.vt.edu/SoSci/converted/Correlation/hwcrop1.gif
https://xkcd.com/552/

Correlation: How to report

- Name the parameters:
  - *r*
    - Strength of the relationship (-1 to +1)
  - df (*n*-2)
    - Degrees of freedom
  - *p*
    - Probability → smaller than α level?

Correlation: SPSS Output

Correlations

		Res_pre	Res_post
Res_pre	Pearson Correlation	1	,687**
	Sig. (2-tailed)		,003
	N	16	16
Res_post	Pearson Correlation	,687**	1
	Sig. (2-tailed)	,003	
	N	16	16

**. Correlation is significant at the 0.01 level (2-tailed).

There was a significant and large positive correlation between Res_pre and Res_post with r (14) = .687, p = .003

chi²

- Explain:
- What do we test with the chi² test of independence?
- What is an alternative for the chi² test (for small samples)?

chi²: How to report

- Name the parameters:
  - χ^2 (chi²)
  - df
  - p

chi²: SPSS output

F1_FAC * Gender Crosstabulation

			Gender		
			Männlich	weiblich	Total
F1_FAC	NO FAC	Count	42	38	80
		Expected Count	40,0	40,0	80,0
	FAC	Count	39	43	82
		Expected Count	41,0	41,0	82,0
Total		Count	81	81	162
		Expected Count	81,0	81,0	162,0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,395[a]	1	,530		
Continuity Correction[b]	,222	1	,637		
Likelihood Ratio	,395	1	,530		
Fisher's Exact Test				,637	,319
N of Valid Cases	162				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 40,00.

b. Computed only for a 2x2 table

There was no significant association between F1_FAC and Gender: χ^2 (1) = .395, p = .530

t-test for independent samples

- Explain:
  - When do you use a t-test for independent samples?
  - Give an example for a study design

t-test for independent samples: How to report

- Name the parameters
  - *t*
  - df
  - *p*
  - *d*
    - Effect size

t-test for independent samples: SPSS output

Group Statistics

	F1_Media	N	Mean	Std. Deviation	Std. Error Mean
Mot_04	Tablet	8	2,50	1,512	,535
	Book	9	2,56	1,333	,444

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Mot_04	Equal variances assumed	,091	,767	-,081	15	,937	-,056	,690	-1,526	1,415
	Equal variances not assumed			-,080	14,120	,937	-,056	,695	-1,545	1,434

There was no significant difference between the conditions tablet and book regarding Mot_04: t (15) = -.081, p = .937, d = .042

t-test for dependent samples (paired samples t-test)

- Explain:
  - When do you use a t-test for dependent samples?
  - Give an example for a study design

t-test for dependent samples: How to report

- Name the parameters
  - *t*
  - df
  - *p*
  - *d*

t-test for dependent samples: SPSS output

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Res_pre	55,9412	17	20,95970	5,08347
	Res_post	59,5882	17	21,65369	5,25179

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Res_pre & Res_post	17	,684	,002

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Res_pre - Res_post	-3,64706	16,93717	4,10787	-12,35535	5,06123	-,888	16	,388

There was no significant change from pre- to post-test (over time): t (16) = -.888, p = .388, d = .017

ANOVA

- Explain: When do you use…
  - An one-way ANOVA?
  - A two-way ANOVA?
  - A repeated measures ANOVA? (one-way or two-way)

- Give an example for a study design each

ANOVA: How to report

- Name the parameters
  - F
  - df
  - p
  - η_p^2
    - Effect size

One-way ANOVA: SPSS output

Descriptive Statistics

Dependent Variable: Res_post

Exp_Condition	Mean	Std. Deviation	N
Tablet+Joy	73,7500	7,36546	4
Tablet+Sadness	40,2500	23,94960	4
Book+Joy	74,4000	17,52997	5
Book+Sadness	46,2500	9,70824	4
Total	59,5882	21,65369	17

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

Dependent Variable: Res_post

F	df1	df2	Sig.
1,738	3	13	,209

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

a. Design: Intercept + Exp_Condition

Tests of Between-Subjects Effects

Dependent Variable: Res_post

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	4106,668[a]	3	1368,889	5,241	,014	,547
Intercept	57958,550	1	57958,550	221,903	,000	,945
Exp_Condition	4106,668	3	1368,889	5,241	,014	,547
Error	3395,450	13	261,188			
Total	67865,000	17				
Corrected Total	7502,118	16				

a. R Squared = ,547 (Adjusted R Squared = ,443)

There was a significant difference between the experimental conditions regarding Res_post with a large effect size:

F (3, 13) = 5.241,
p = .014, η_p^2 = .547

Two-way ANOVA: SPSS output

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

Dependent Variable: Mot_mean

F	df1	df2	Sig.
3,335	3	13	,053

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

a. Design: Intercept + F1_Media + F2_Ind_Emo + F1_Media * F2_Ind_Emo

Tests of Between-Subjects Effects

Dependent Variable: Mot_mean

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	6,080[a]	3	2,027	3,008	,069	,410
Intercept	125,753	1	125,753	186,662	,000	,935
F1_Media	,296	1	,296	,439	,519	,033
F2_Ind_Emo	5,235	1	5,235	7,770	,015	,374
F1_Media * F2_Ind_Emo	,624	1	,624	,926	,353	,067
Error	8,758	13	,674			
Total	144,503	17				
Corrected Total	14,838	16				

a. R Squared = ,410 (Adjusted R Squared = ,274)

There was no significant main effect of media on motivation:
F (1, 13) = .439, p = .519, η_p^2 = .033

There was a significant main effect of Ind_Emo on motivation with a large effect size:
F (1, 13) = 7.770, p = .015, η_p^2 = .374

There was no significant interaction between media and Ind_Emo on motivation:
F (1, 13) = .926, p = .353, η_p^2 = .067

Two-way ANOVA with repeated measures: SPSS output

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
LearningGainPrePost	Sphericity Assumed	247,860	1	247,860	1,789	,204	,121
	Greenhouse-Geisser	247,860	1,000	247,860	1,789	,204	,121
	Huynh-Feldt	247,860	1,000	247,860	1,789	,204	,121
	Lower-bound	247,860	1,000	247,860	1,789	,204	,121
LearningGainPrePost * F2_Ind_Emo	Sphericity Assumed	32,021	1	32,021	,231	,639	,017
	Greenhouse-Geisser	32,021	1,000	32,021	,231	,639	,017
	Huynh-Feldt	32,021	1,000	32,021	,231	,639	,017
	Lower-bound	32,021	1,000	32,021	,231	,639	,017
LearningGainPrePost * Gender	Sphericity Assumed	53,869	1	53,869	,389	,544	,029
	Greenhouse-Geisser	53,869	1,000	53,869	,389	,544	,029
	Huynh-Feldt	53,869	1,000	53,869	,389	,544	,029
	Lower-bound	53,869	1,000	53,869	,389	,544	,029
LearningGainPrePost * F2_Ind_Emo * Gender	Sphericity Assumed	194,622	1	194,622	1,405	,257	,098
	Greenhouse-Geisser	194,622	1,000	194,622	1,405	,257	,098
	Huynh-Feldt	194,622	1,000	194,622	1,405	,257	,098
	Lower-bound	194,622	1,000	194,622	1,405	,257	,098
Error (LearningGainPrePost)	Sphericity Assumed	1801,100	13	138,546			
	Greenhouse-Geisser	1801,100	13,000	138,546			
	Huynh-Feldt	1801,100	13,000	138,546			
	Lower-bound	1801,100	13,000	138,546			

(solutions on next slide)

Two-way ANOVA with repeated measures: SPSS output

Ind_Emo did not significantly affect learning gain (no interaction of Ind_Emo with change over time):
F (1, 13) = .231, p = .639, η_p^2 = .017

Gender did not significantly affect learning gain (no interaction of gender with change over time):
F (1, 13) = .389, p = .544, η_p^2 = .029

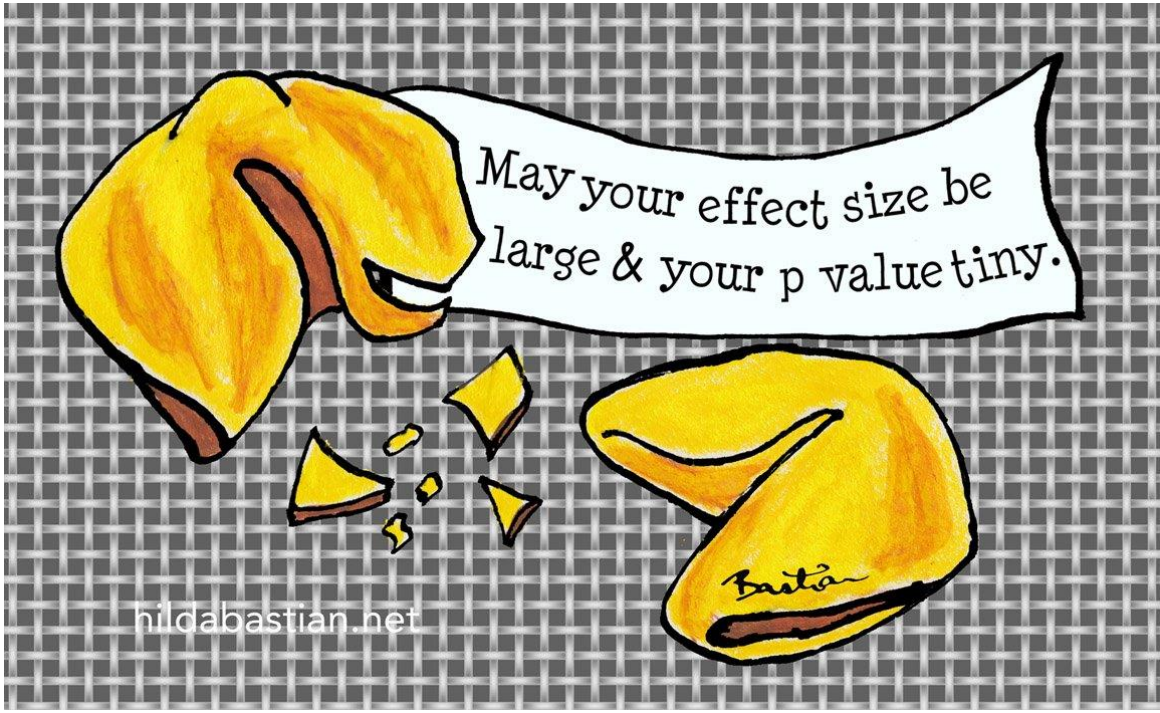
There was also no triple interaction between Ind_Emo, gender, and learning gain:
F (1, 13) = 1.405, p = .257, η_p^2 = .098

Effect size

- Explain
  - Why do we need to report it?

- Interpretation of η_p^2
  - Small ≥ .01
  - Medium ≥ .06
  - Large ≥ .14

- Interpretation of *d*
  - Small ≥ .2
  - Medium ≥ .5
  - Large ≥ .8

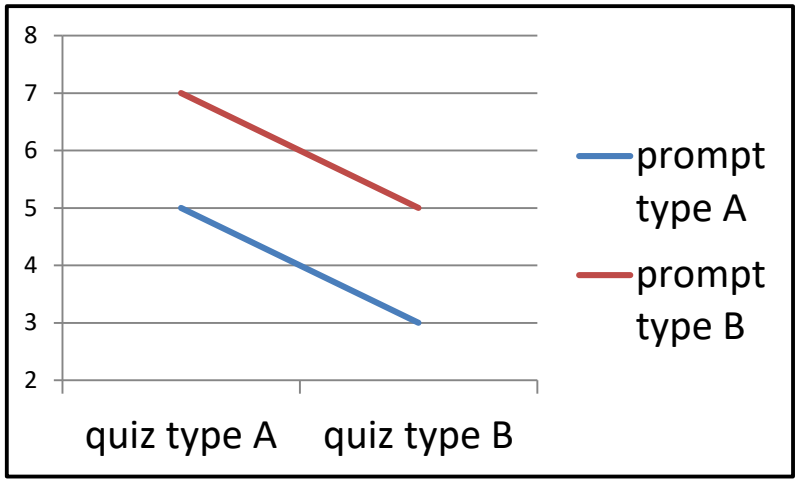


https://statistically-funny.blogspot.com/2017/12/a-science-fortune-cookie.html

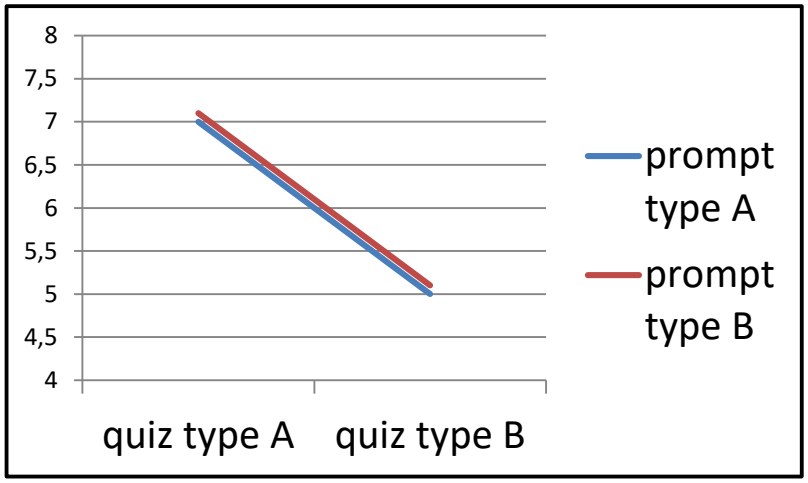
Quiz: Which test do you use?

- You want to know if the level of knowledge changes over time (t1, t2, t3).
  - Repeated measures ANOVA
- You want to know if pulse and level of reported anxiety (0-100) are related.
  - Pearson correlation
- You want to know if there is a difference in exam results (max. 70 points) between the students that practiced with the exercise sheets and the ones that did not.
  - t-test for independent samples
- You want to know if there is an interaction between study program (psychology vs. computer science) and level of studies (bachelor vs. master) on the amount of credit points obtained in one semester.
  - Two-way ANOVA

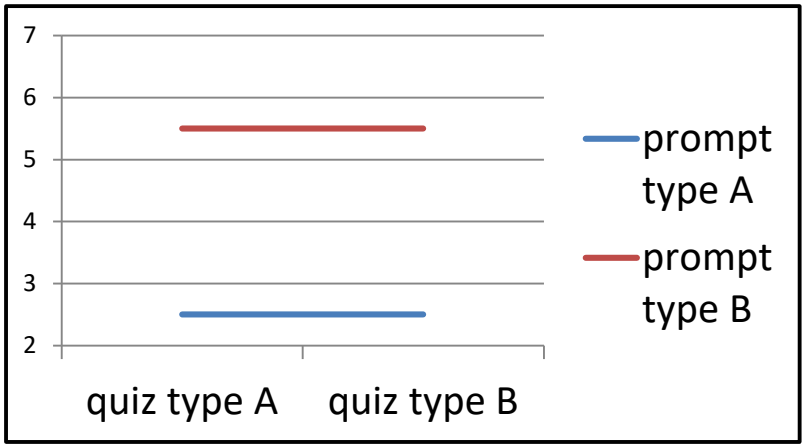
Main effects and interactions



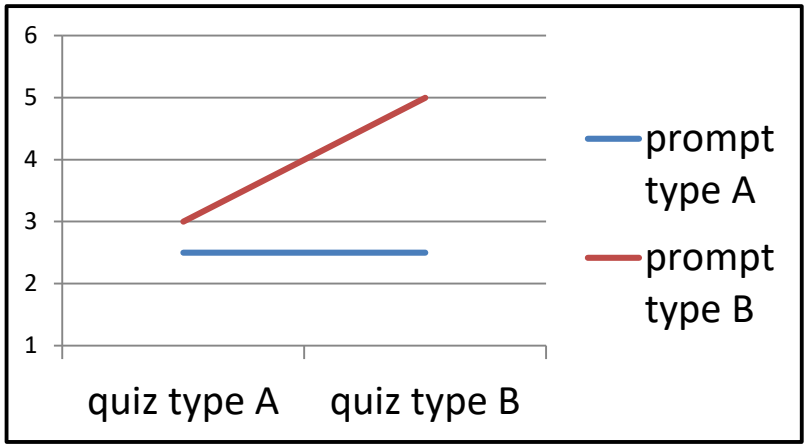
Main effect of prompt type
Main effect of quiz type



Main effect of quiz type



Main effect of prompt type



Main effect of prompt type
Interaction prompt type X quiz type

Exam:

- Next week (February 04)
- 10:15-11:45
- Graded:
  - Pass: 1.0-4.0
  - Fail: 5.0

Good luck ☺

https://knowyourmeme.com/photos/665819-i-have-no-idea-what-im-doing