Discovering Statistics Using JASP

JASP

2025 Workshop

UNIVERSITEIT VAN AMSTERDAM

Outline

- JASP Intro
- Correlation
- Regression
- PROCESS
- T-test
- ANOVA's
- Free-for-all

edu.nl/knhfd

Goals of this Workshop

- Get you familiar with JASP
- Show JASP workflow
- Know how to get in touch
- Have ran your favorite analysis in JASP

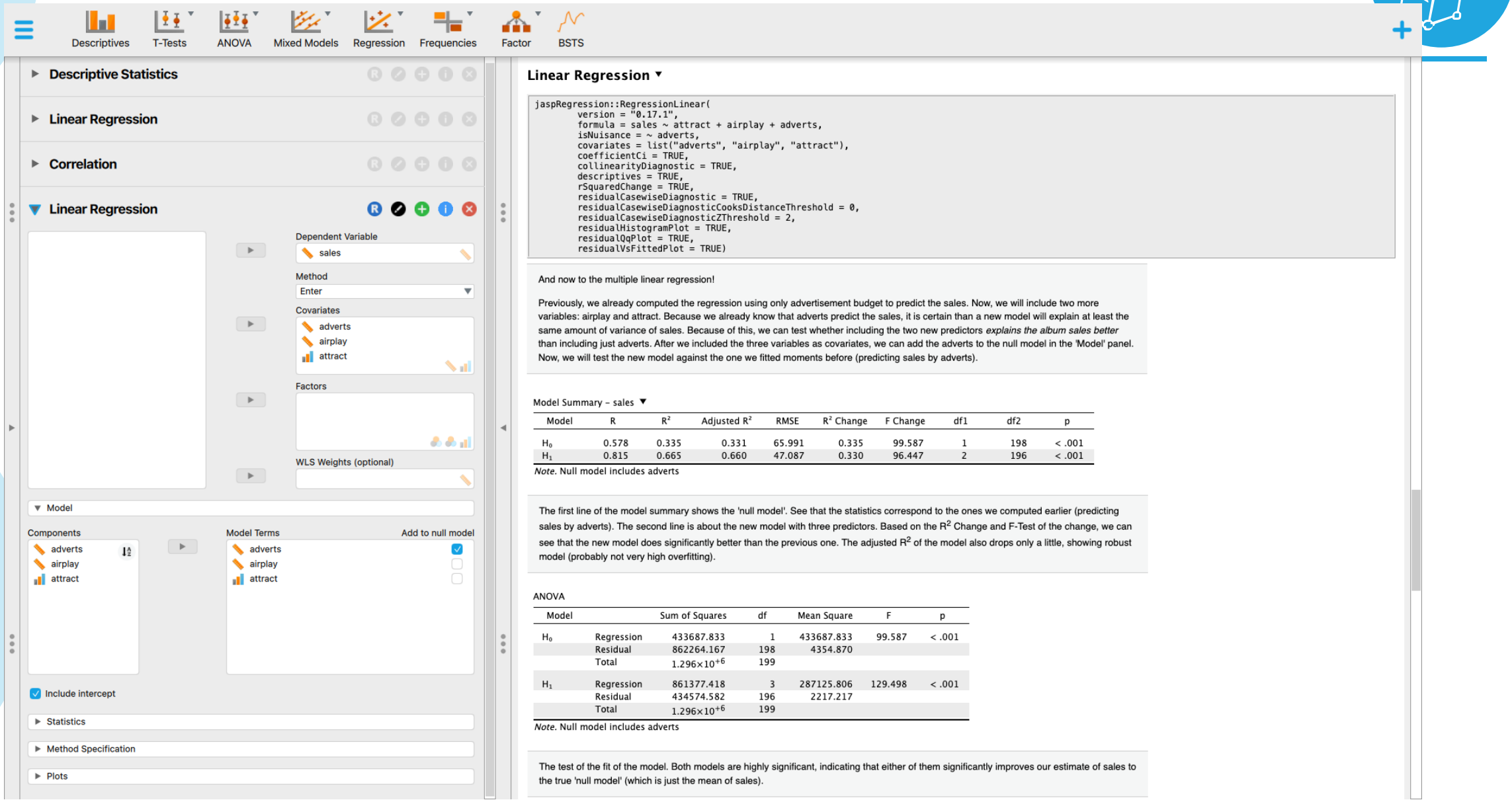
What is JASP?

Developed at UvA over the past 10 years, funded by research grants (NWO/EU)

Graphical user interface for conducting frequentist and Bayesian statistics

https://jasp-stats.org/

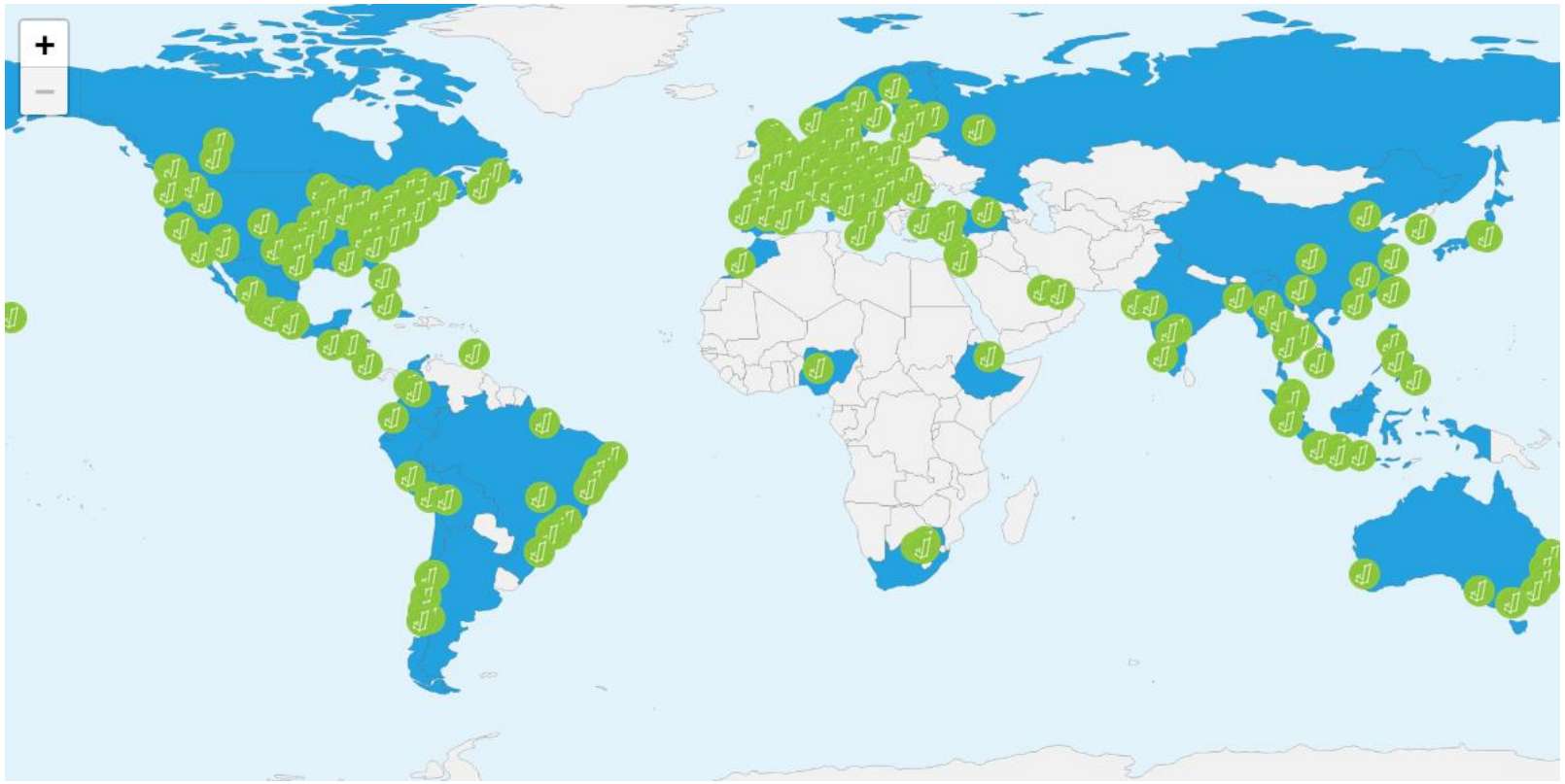
What is JASP?


Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
BSTS
Descriptive Statistics
Linear Regression
Correlation
Linear Regression
Dependent Variable
sales
Method
Enter
Covariates
adverts
airplay
attract
Factors
WLS Weights (optional)
Model
Components
adverts
airplay
attract
Model Terms
Add to null model
adverts
airplay
attract
Include intercept
Statistics
Method Specification
Plots
Linear Regression
jaspRegression::RegressionLinear(
version = "0.17.1",
formula = sales ~ attract + airplay + adverts,
isNuisance = ~ adverts,
covariates = list("adverts", "airplay", "attract"),
coefficientCi = TRUE,
collinearityDiagnostic = TRUE,
descriptives = TRUE,
rSquaredChange = TRUE,
residualCasewiseDiagnostic = TRUE,
residualCasewiseDiagnosticCooksDistanceThreshold = 0,
residualCasewiseDiagnosticZThreshold = 2,
residualHistogramPlot = TRUE,
residualQqPlot = TRUE,
residualVsFittedPlot = TRUE)
And now to the multiple linear regression!
Previously, we already computed the regression using only advertisement budget to predict the sales. Now, we will include two more variables: airplay and attract. Because we already know that adverts predict the sales, it is certain than a new model will explain at least the same amount of variance of sales. Because of this, we can test whether including the two new predictors explains the album sales better than including just adverts. After we included the three variables as covariates, we can add the adverts to the null model in the 'Model' panel. Now, we will test the new model against the one we fitted moments before (predicting sales by adverts).
Model Summary - sales
Model R R² Adjusted R² RMSE R² Change F Change df1 df2 p
H₀ 0.578 0.335 0.331 65.991 0.335 99.587 1 198 < .001
H₁ 0.815 0.665 0.660 47.087 0.330 96.447 2 196 < .001
Note. Null model includes adverts
The first line of the model summary shows the 'null model'. See that the statistics correspond to the ones we computed earlier (predicting sales by adverts). The second line is about the new model with three predictors. Based on the R² Change and F-Test of the change, we can see that the new model does significantly better than the previous one. The adjusted R² of the model also drops only a little, showing robust model (probably not very high overfitting).
ANOVA
Model Sum of Squares df Mean Square F p
H₀ Regression 433687.833 1 433687.833 99.587 < .001
Residual 862264.167 198 4354.870
Total 1.296×10^+6 199
H₁ Regression 861377.418 3 287125.806 129.498 < .001
Residual 434574.582 196 2217.217
Total 1.296×10^+6 199
Note. Null model includes adverts
The test of the fit of the model. Both models are highly significant, indicating that either of them significantly improves our estimate of sales to the true 'null model' (which is just the mean of sales).

What is JASP?

Used at 271 universities across 64 countries

80,000 monthly downloads



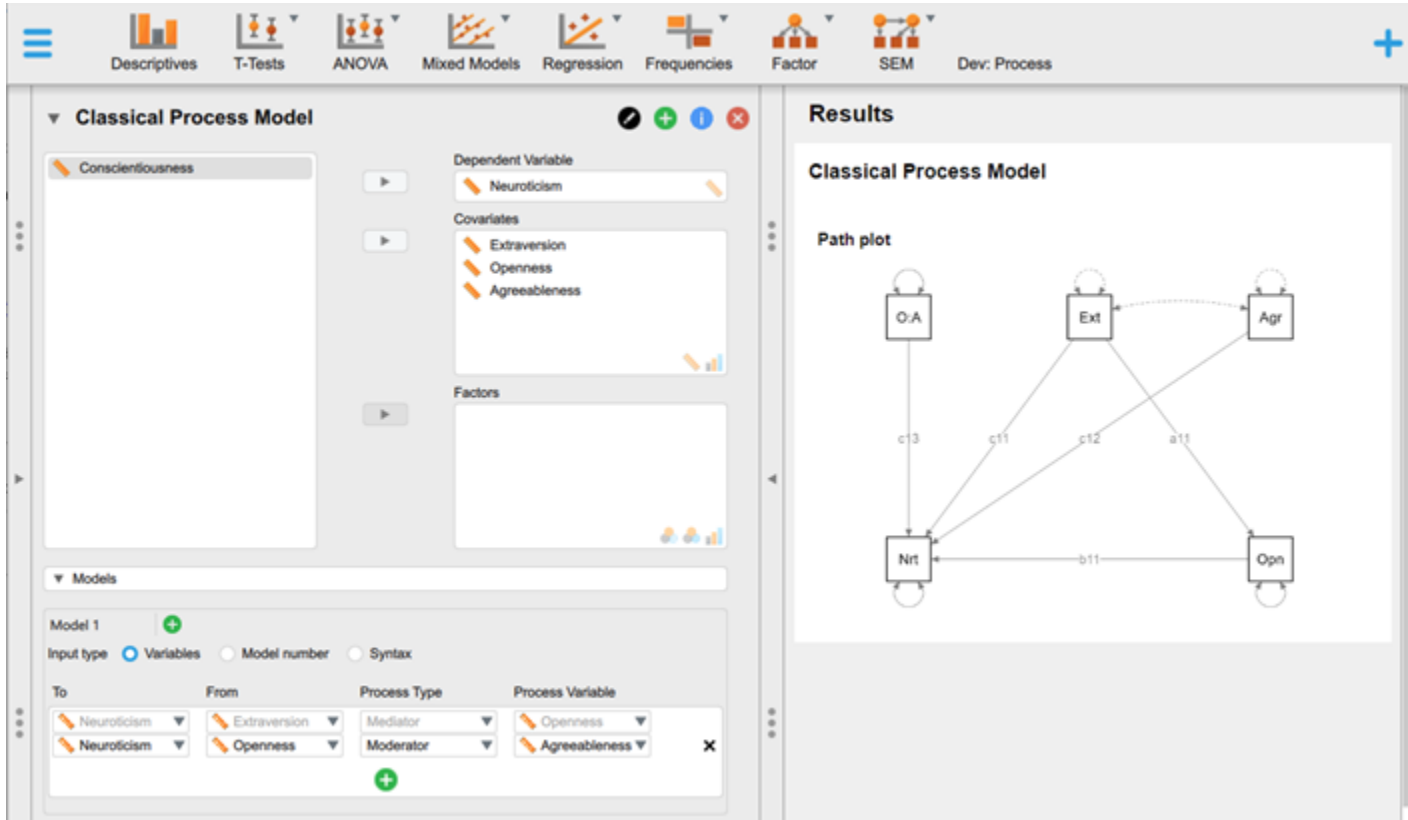
https://jasp-stats.org/teaching-with-jasp/#worldmap

Features

- Website overview
- JASP vs. SPSS feature comparison
- Data formats: .sav, .xls, .txt, .csv, .ods, .tsv, .dta, .por, .sas7bdat, .sas7bcat, and the .jasp format
- APA tables
- OSF integration
- R console
- Compute columns
- Filtering

Feature Roadmap

- Full syntax mode (blog about the first implementation)
- More data manipulation
- Select filters



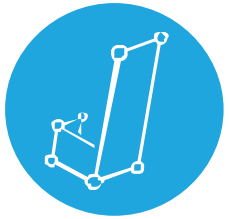
Other Handy Resources

- How to Use JASP – Inventory of blogs/videos/gifs for frequentist and Bayesian analyses
- JASP YouTube page
- The JASP Video Library
- Step By Step Guide: 1. Bayesian One-Way ANOVA and the full playlist
- JASP on Bluesky - https://bsky.app/profile/jaspstats.bsky.social
- JASP forum - https://forum.cogsci.nl/index.php?p=/categories/jasp-bayesfactor
- Found a bug? Please report on Github: https://github.com/jasp-stats/jasp-issues/issues
- JASP Verification Project
- More JASP workshops: https://jasp-stats.org/workshop/

JASP Literature

- The JASP Data Library
- Discovering Statistics Using JASP
- Learning Statistics with JASP: A Tutorial for Psychology Students and Other Beginners by Danielle J. Navarro, David R. Foxcroft, and Thomas J. Faulkenberry
- Statistics of Doom by Erin Buchanan
- Statistical Analysis in JASP. A Guide for Students by Mark Goss-Sampson
- Quantitative Analysis with JASP open-source software by Chris Halter (amazon)

Some Examples/Demos



From my own course (Research Methods & Statistics)

- https://johnnydoorn.github.io/IntroductionBayesianInference/06-exercises.html

OSF integration

- https://osf.io/u2e9d/files/osfstorage

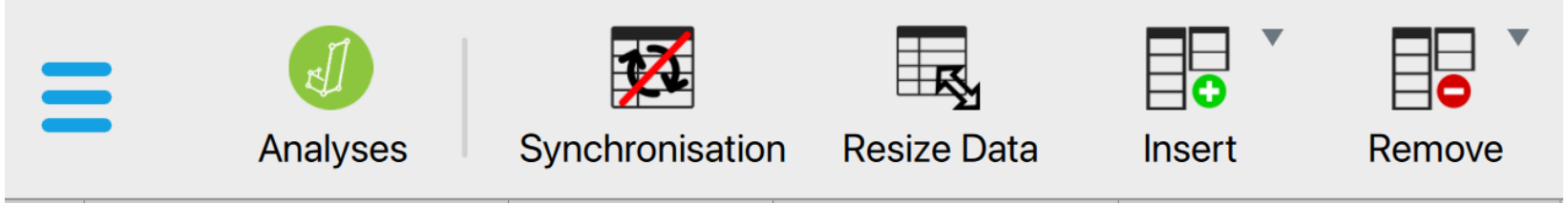
Discover JASP

- https://discoverjasp.com/

JASP

Data Management

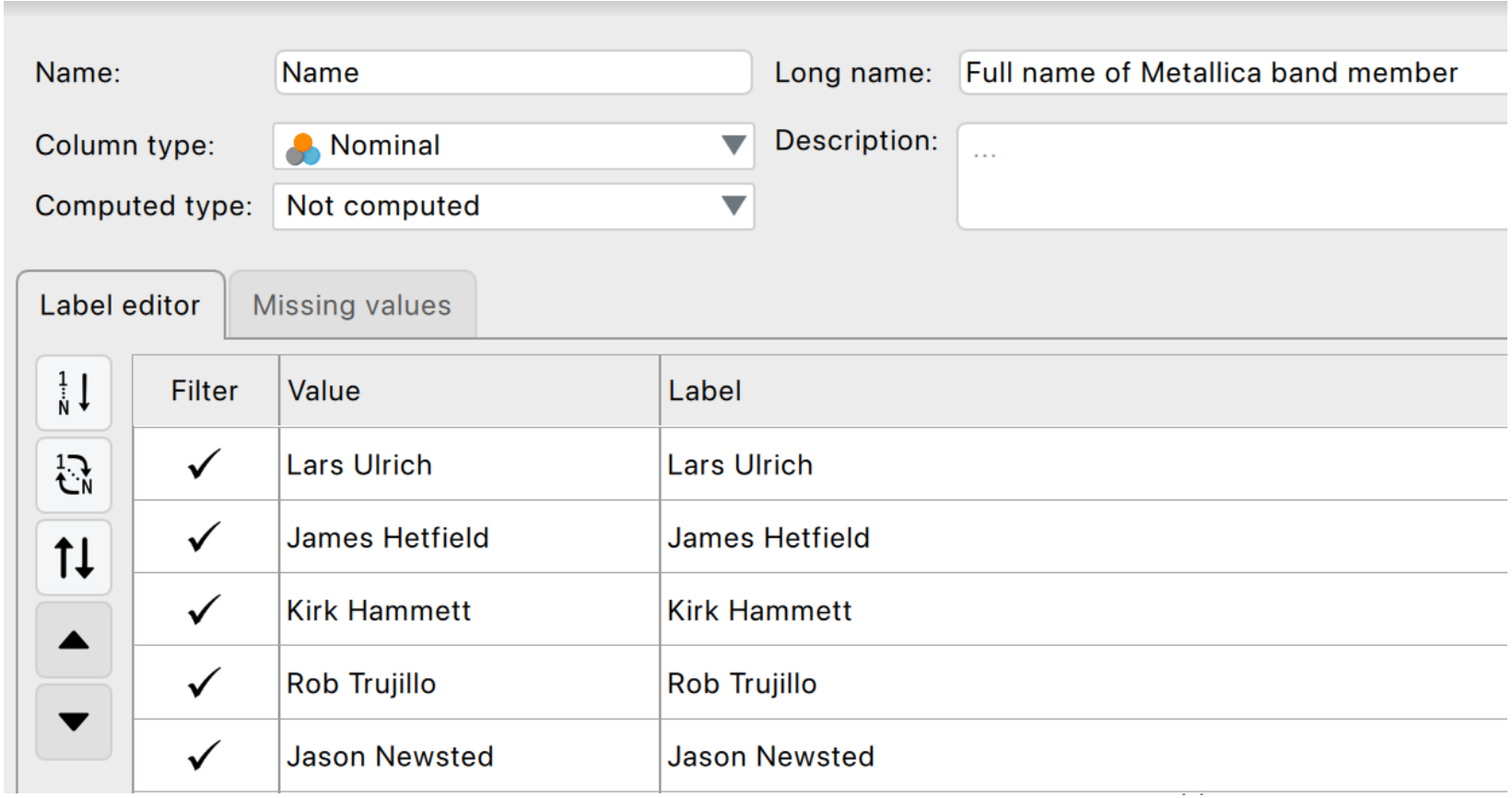
The JASP data editor for the Metallica data



	Name	Instrument	Current member	Headbanging intensity
1	Lars Ulrich	Drums	Yes 1	Light 1
2	James Hetfield	Guitar	Yes 1	Heavy 3
3	Kirk Hammett	Guitar	Yes 1	Light 1
4	Rob Trujillo	Bass	Yes 1	Moderate 2
5	Jason Newsted	Bass	No 0	Heavy 3

The Variable View

Figure 4.6 The variable settings for ‘Name’



Variable Types

Scale 

- Numbers (e.g., 7, 0, 120, 8.5)

Nominal 

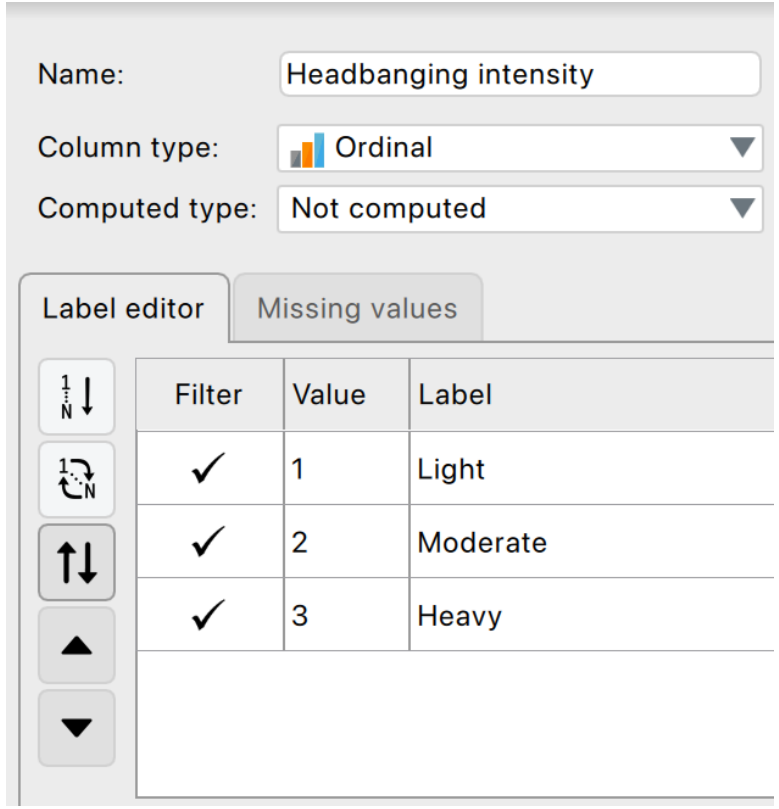
- Categories (e.g., ‘Control group’, ‘Experimental group’)

Ordinal 

- Ordered values (e.g., ‘Dislike’, ‘Neutral’, ‘Like’)

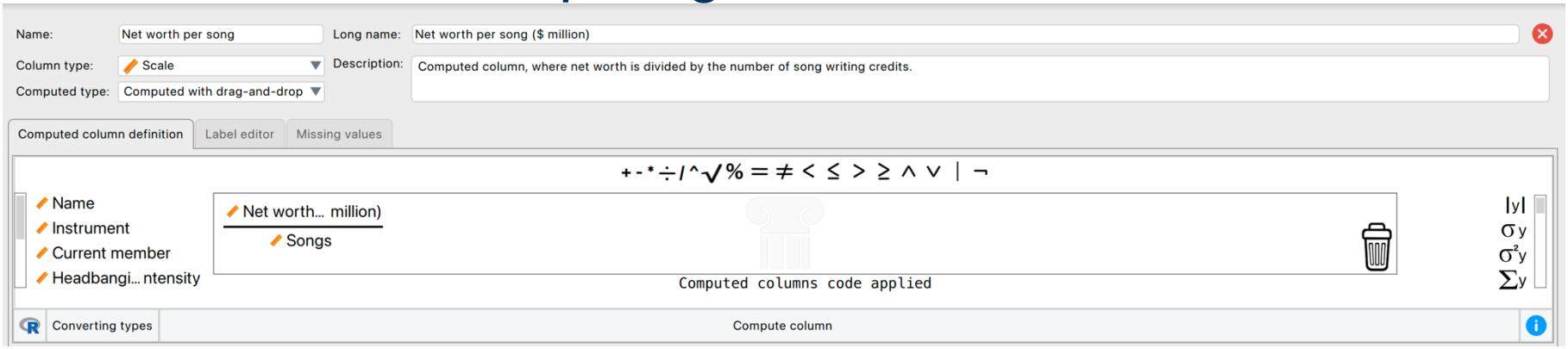
Variable Settings

Figure 4.7 Specifying the values for an ordinal variable



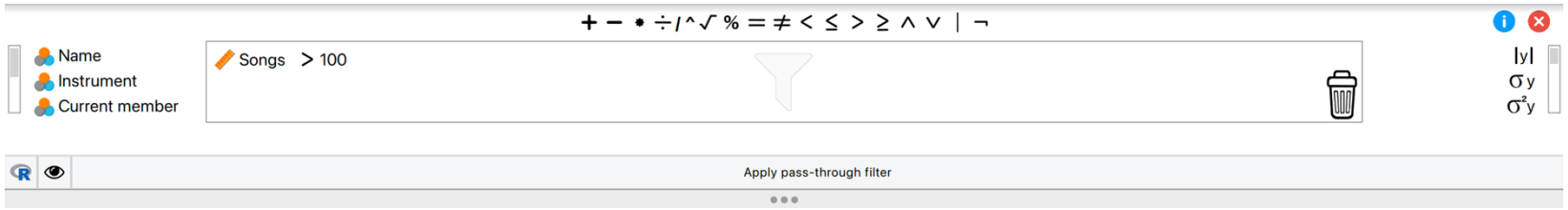
Computing a New Variable

Figure 4.8 The drag and drop interface for computing a new variable

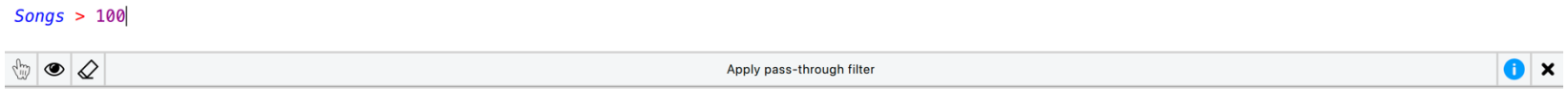


Filtering Data

- Using Variable Settings
- Using the Filter functionality
  - Drag and drop

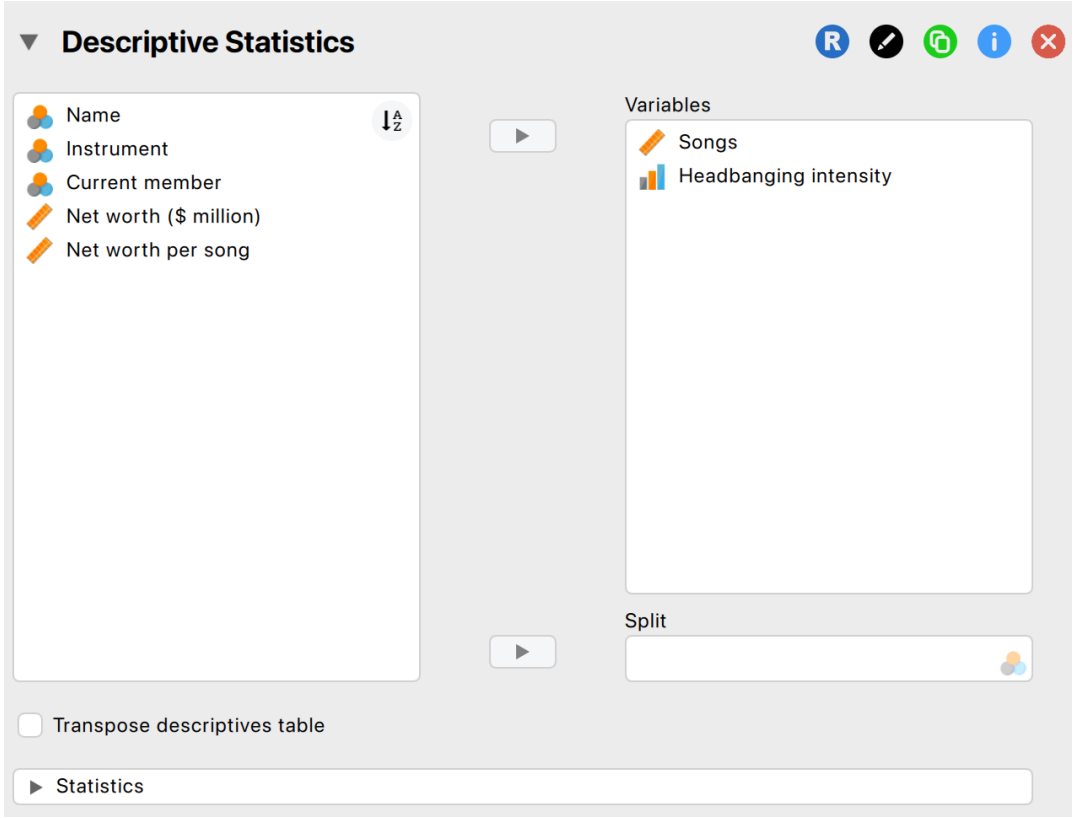


  - R-mode



Descriptives

Figure 4.10 Input window for the Descriptives module



Output Window in JASP

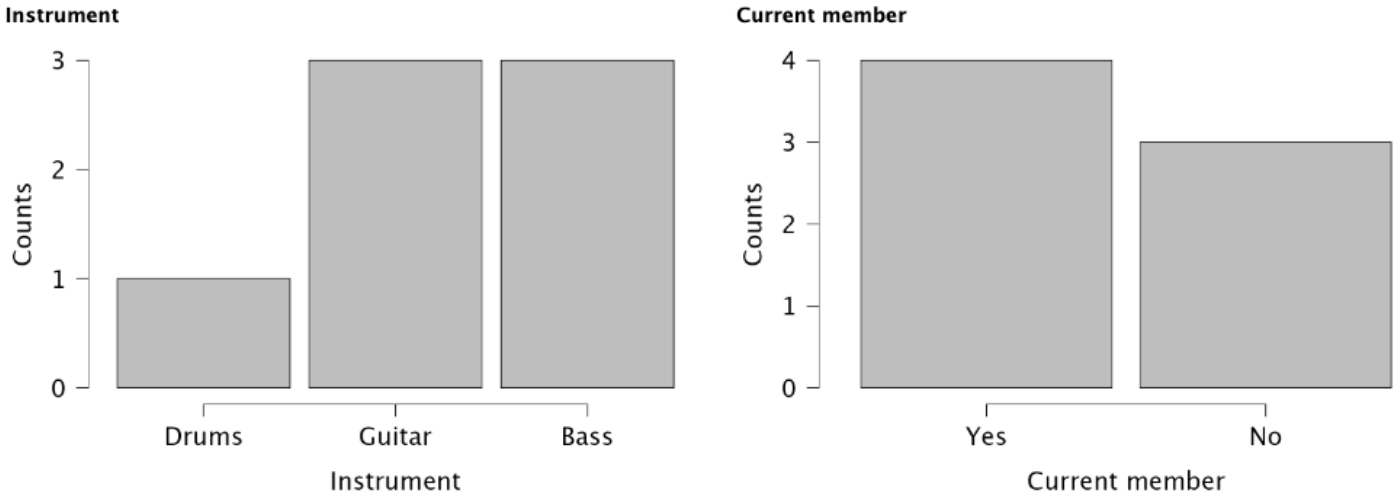
Figure 4.11 Example of annotated output

Distribution Plots

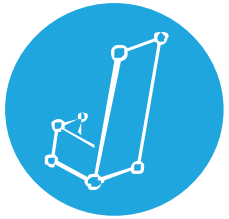
Normal Normal

We can have fancy $\LaTeX$ formula's in here, link cute cat video's, or insert a drumkit 

Below are two distribution plots outlining the members of Metallica. On the left, we see the various instruments being played and their frequencies, and on the right we see how many members are still active, and how many left the band.

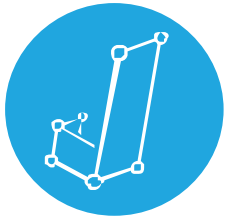


Basic Flow of Data Analysis in JASP



- Describe/visualize data
- Specify the analysis in JASP
- Assess the assumptions (tip: see the help-files)
- Interpret the main analysis table
- Consider follow-up analyses

Regression



Regression with One Predictor

A record company boss was interested in predicting album sales from advertising.

Data

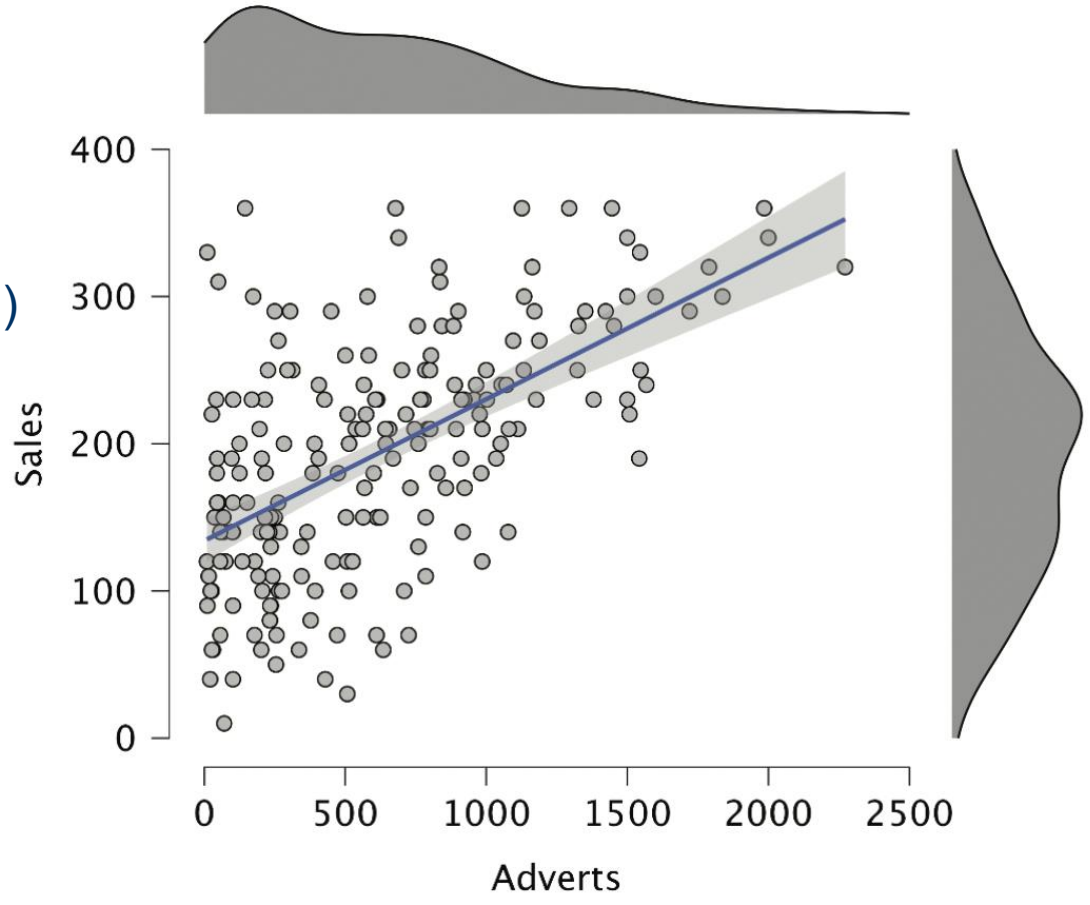
- 200 different album releases

Outcome variable:

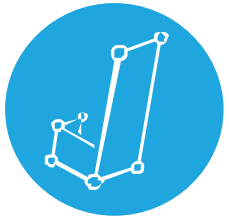
- Album sales in the week after release (x1000)

Predictor variables

- Advertisement budget (in £1000)
- Number of plays on the radio
- Image of the band.



The Model as an Equation



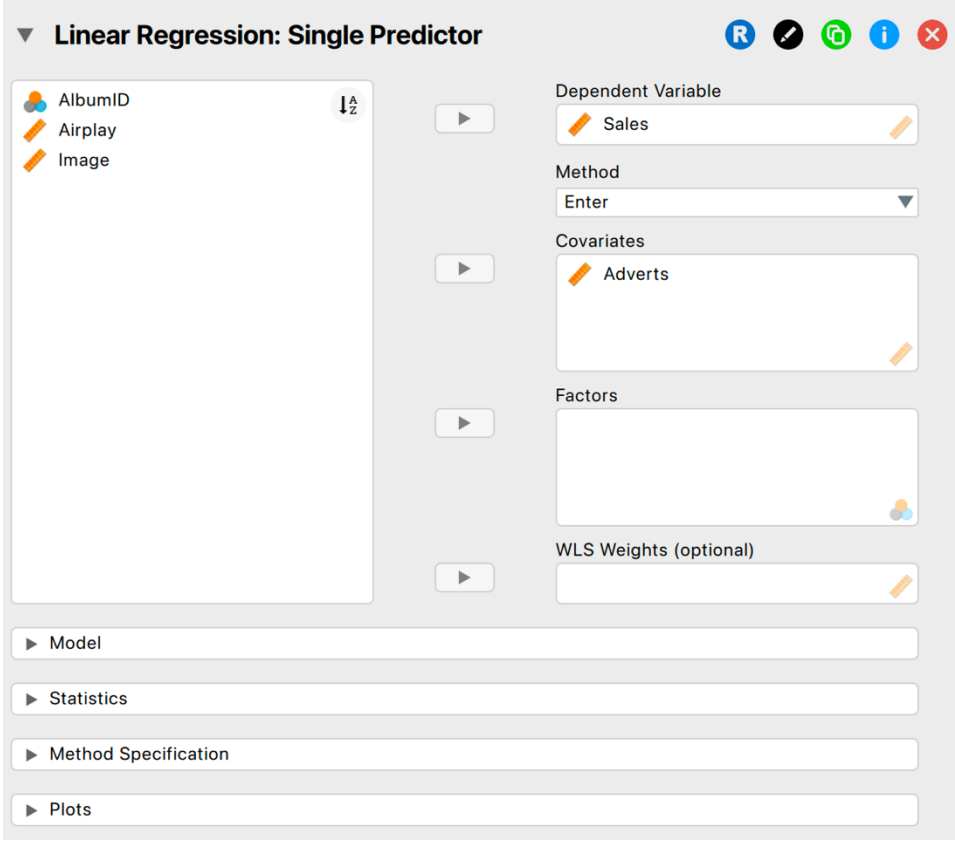
- The model contains two regression weights:

$$Y_i = (b_0 + b_1 X_{1i}) + \varepsilon_i$$

- b_0 is the intercept
  - The intercept is the value of the *Y* variable when all *Xs* = 0
  - E.g., how many albums are sold for 0£ advertisement budget
- b_1 is the coefficient for Adverts.

Regression with One Predictor

Figure 8.10 Main menu for regression



Model Summary

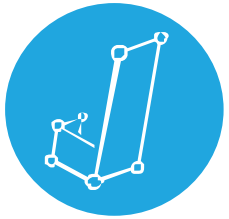
Output 8.2

Model Summary - Sales

Model	R	R^2	Adjusted R^2	RMSE
M_0	0.000	0.000	0.000	80.699
M_1	0.578	0.335	0.331	65.991

Note. M_1 includes Adverts

Multiple Regression



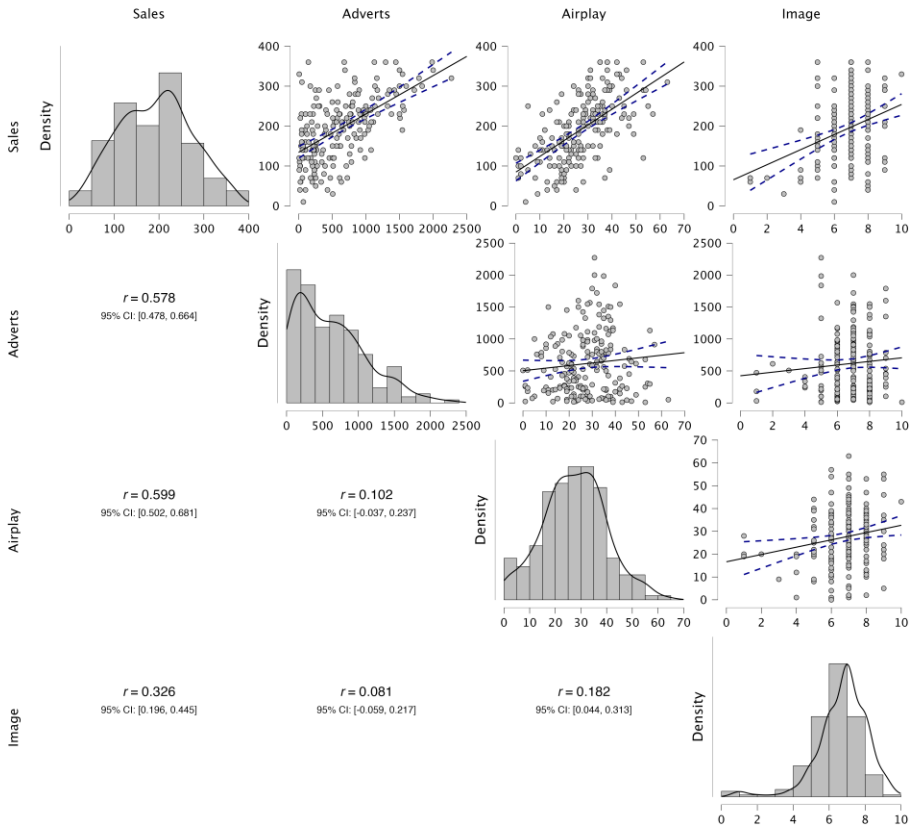
- With several predictors the model now contains multiple regression weights:

$$Y_i = (b_0 + b_1 X_{1i} + b_2 X_{2i} + \cdots b_n X_{ni}) + \varepsilon_i$$

- b_0 is the intercept.
  - The intercept is the value of the *Y* variable when all *Xs* = 0
- b_1 is the coefficient for Adverts
- b_2 is the coefficient for Airplay
- b_n is the coefficient for n^{th} variable.

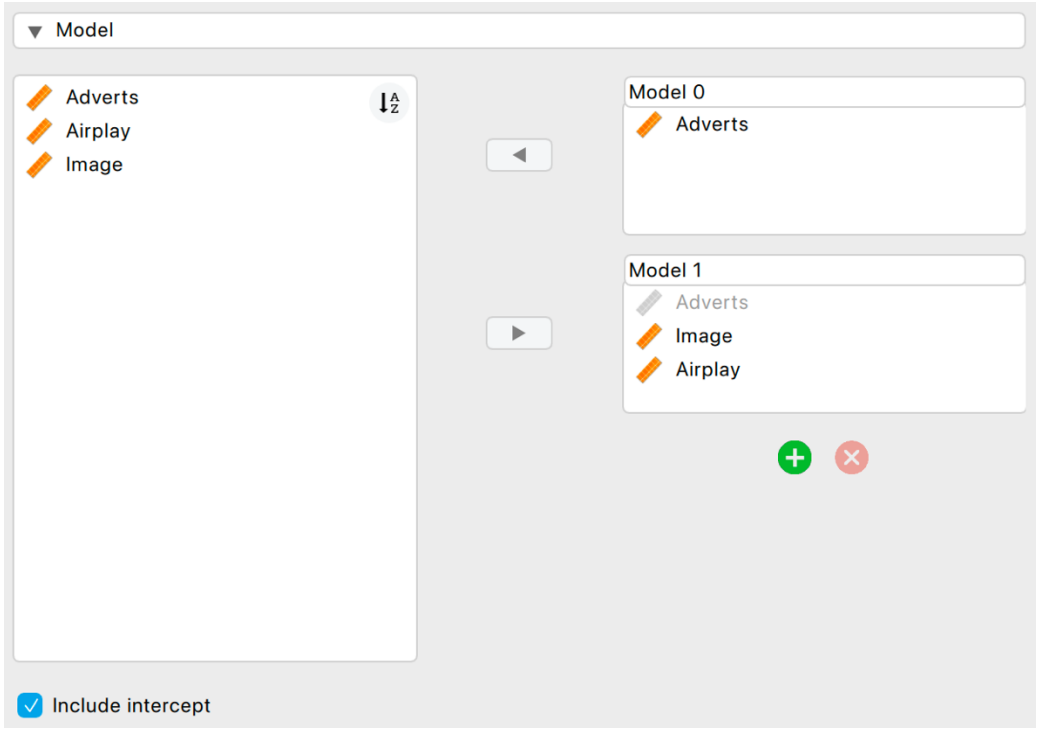
A model with Several Predictors

Figure 8.11 Matrix scatterplot of the relationships between advertising budget, airplay, image rating and album sales



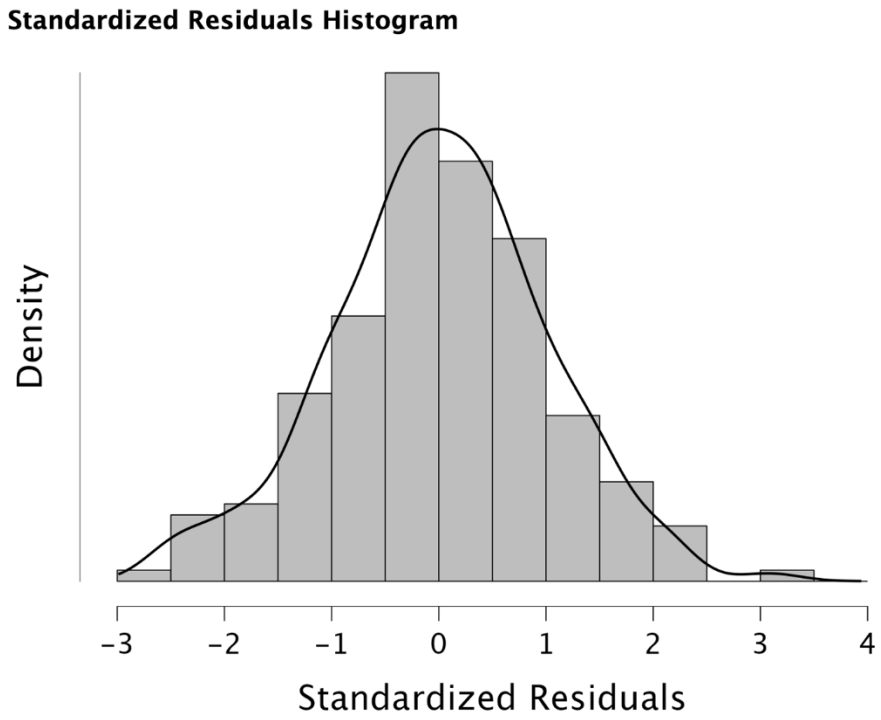
Multiple Regression

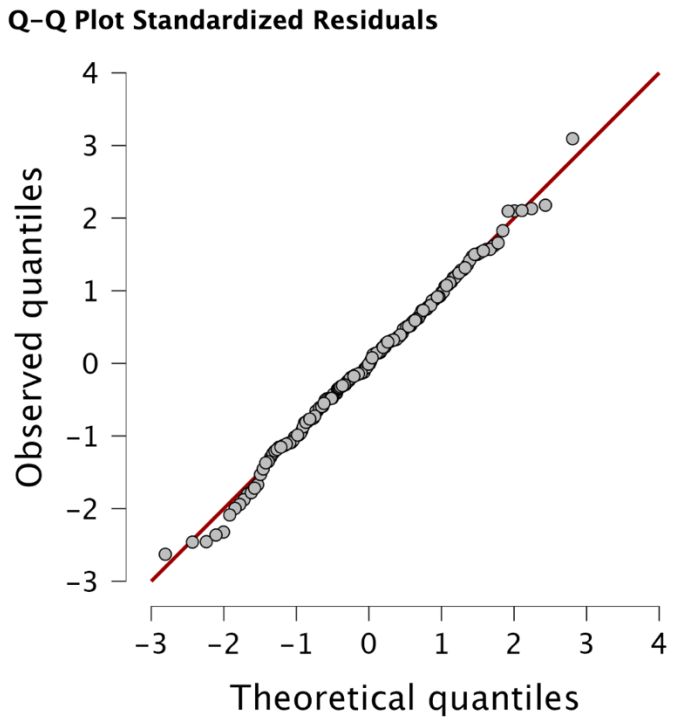
Figure 8.12 Main menu for block 2 of the multiple regression



Normality of Residuals: Histograms and Q-Q Plots

Figure 8.17 Histogram and Q-Q plot for the residuals from our model





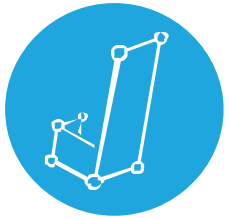
Model Parameters

Output 8.7

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p	95% CI		Collinearity Statistics	
							Lower	Upper	Tolerance	VIF
M_0	(Intercept)	134.140	7.537		17.799	< .001	119.278	149.002		
	Adverts	0.096	0.010	0.578	9.979	< .001	0.077	0.115	1.000	1.000
M_1	(Intercept)	−26.613	17.350		−1.534	0.127	−60.830	7.604		
	Adverts	0.085	0.007	0.511	12.261	< .001	0.071	0.099	0.986	1.015
	Image	11.086	2.438	0.192	4.548	< .001	6.279	15.894	0.963	1.038
	Airplay	3.367	0.278	0.512	12.123	< .001	2.820	3.915	0.959	1.043

Interpreting Model Parameters



b-values:

- The change in the outcome associated with a unit change in the predictor.
- E.g., **Advertising budget**: $b = 0.085$
  - As advertising budget increases by one unit, album sales increase by 0.085 units. Both variables were measured in thousands; therefore, for every £1000 more spent on advertising, an extra 0.085 thousand albums (85 albums) are sold. This interpretation is true only if the effects of band image and airplay are held constant.

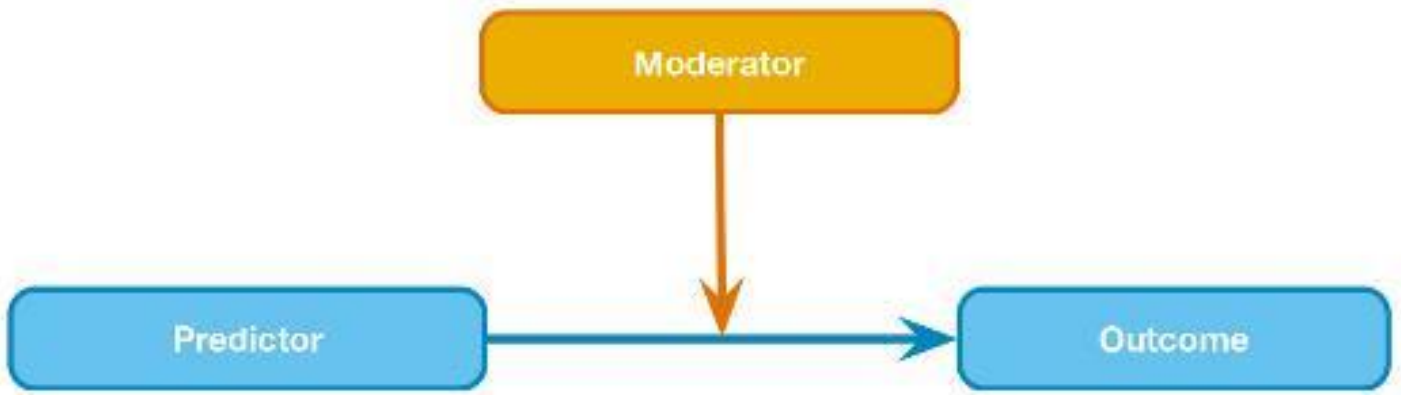
Regression Exercises

Alex Examples

Leni Examples

Moderation

Figure 10.2 Diagram of the *conceptual* moderation model

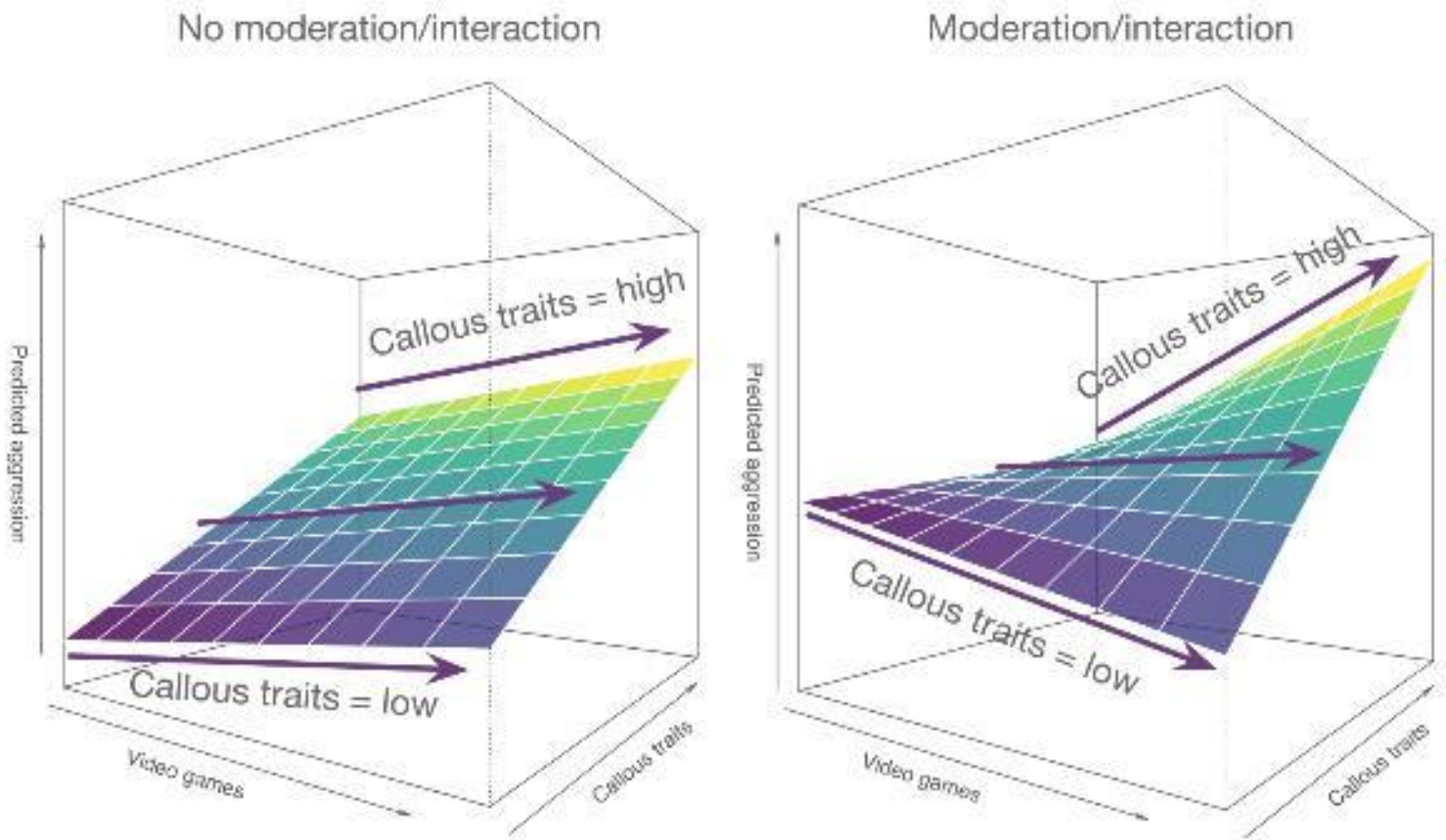


Example

- Do violent video games make people antisocial?
- Participants
  - 442 youths
- Variables
  - Aggression
  - Callous unemotional traits (CaUnTs)
  - Number of hours spent playing video games per week
- Is ‘CaUnTs’ a moderator?
- Warning
  - That’s a Lot to Process! Pitfalls of Popular Path Models

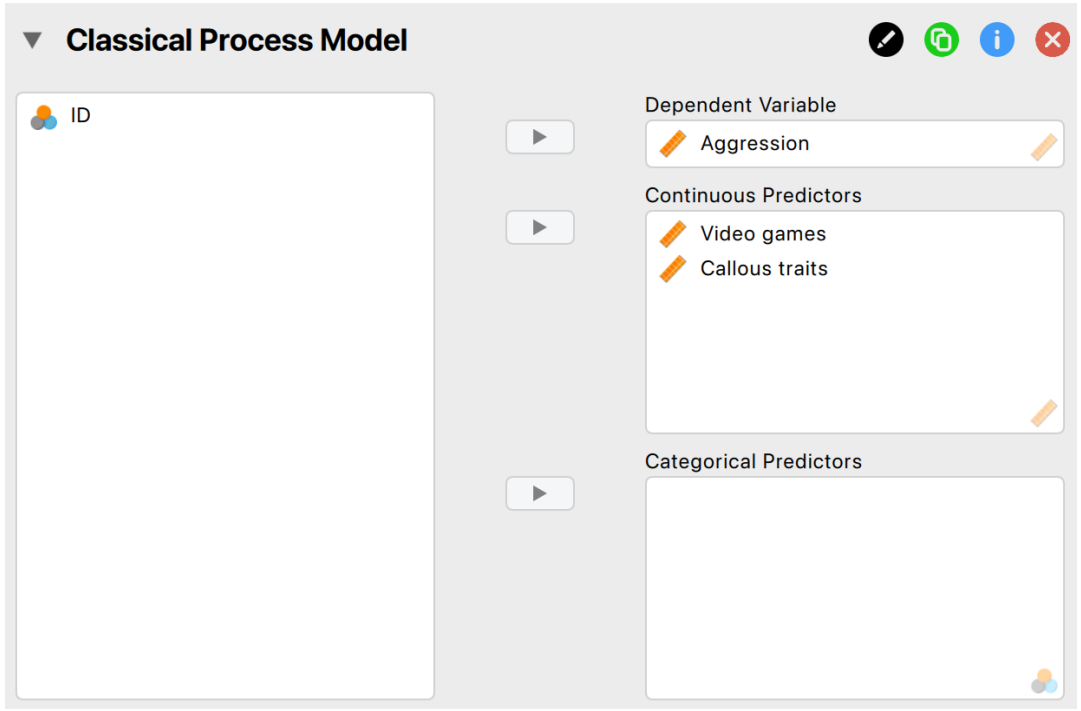
Moderation

Figure 10.4 Callousness as a moderator



Moderation Analysis in JASP

Figure 10.6 The main menu for running moderation analysis in the Process module



Moderation Analysis in JASP

Figure 10.7 Menu for the model builder for a moderation analysis

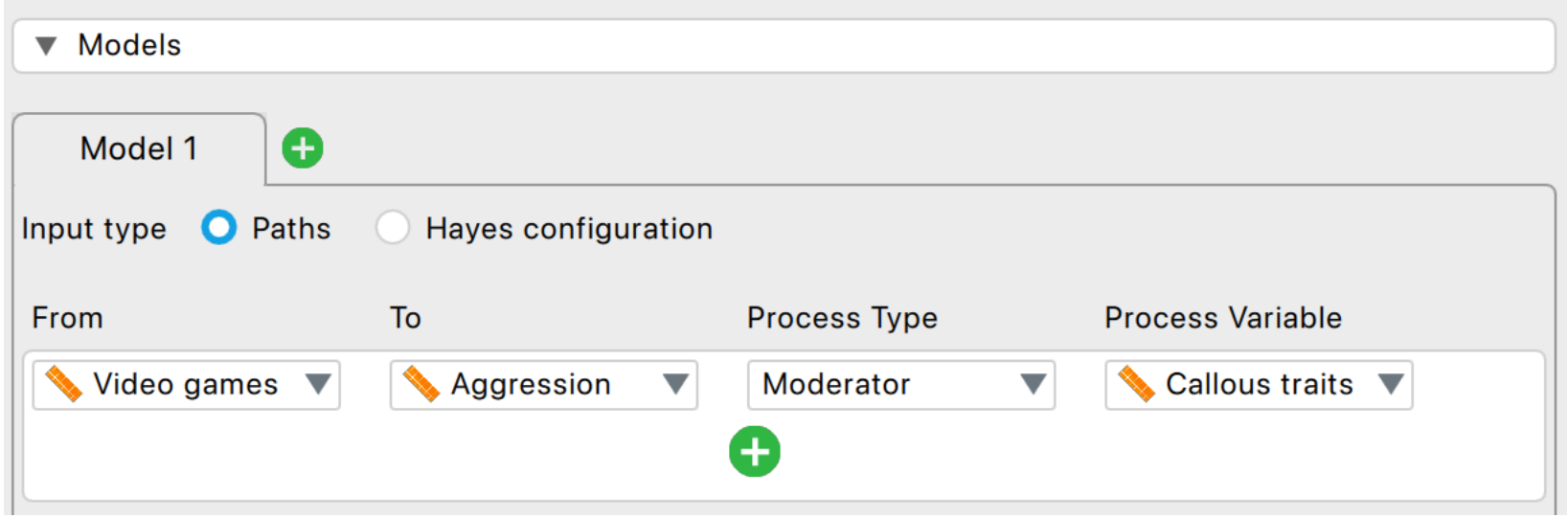
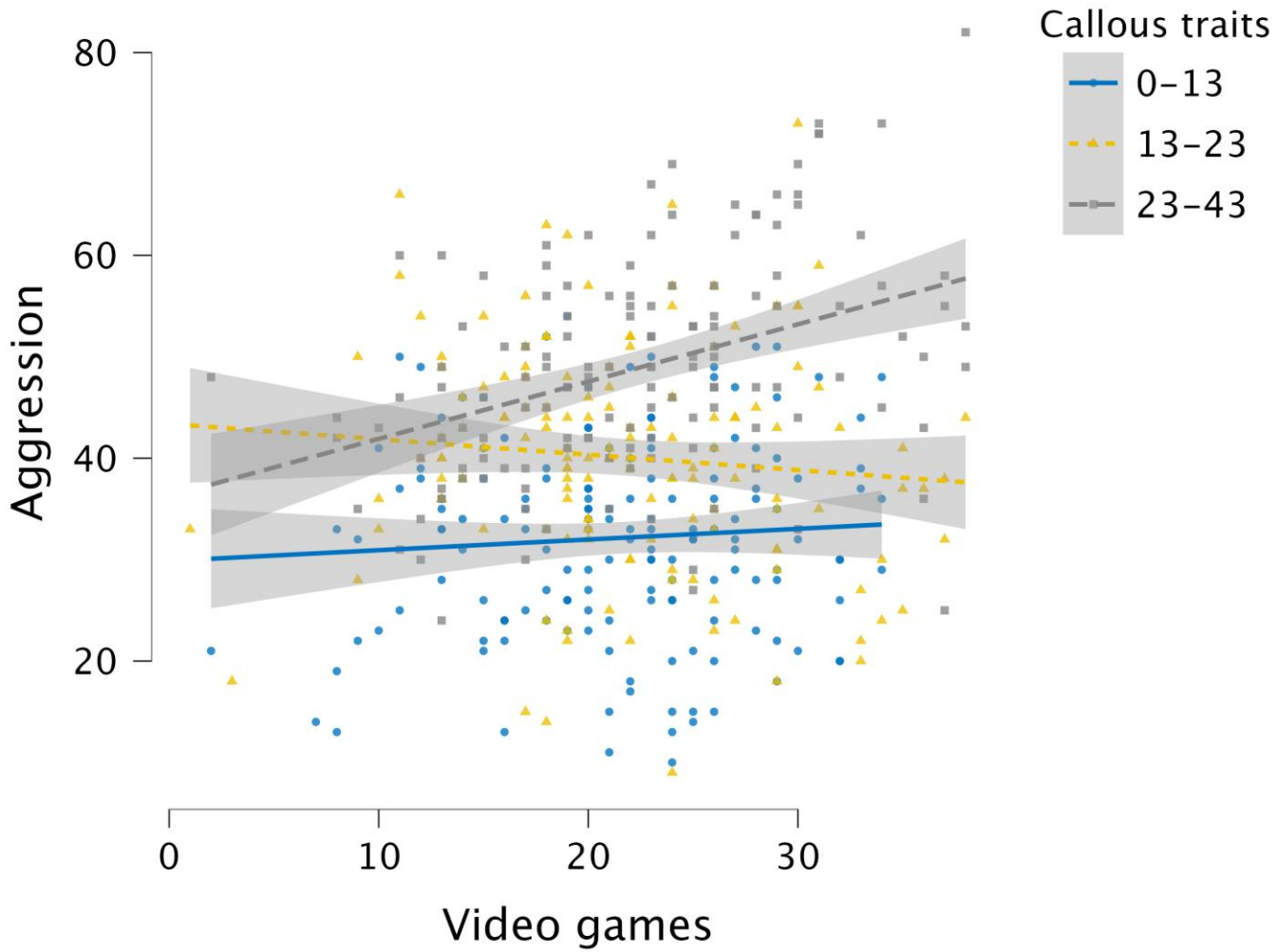
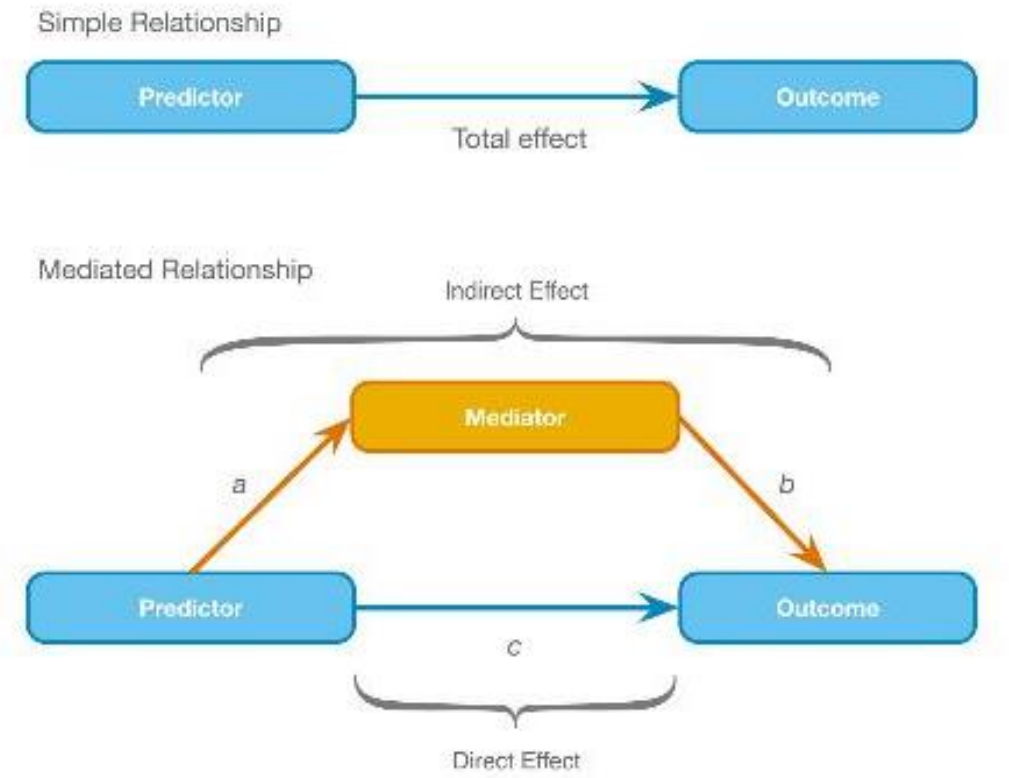


Figure 10.9 Plotting the interaction effect using Flexplot, where Callous traits is binned



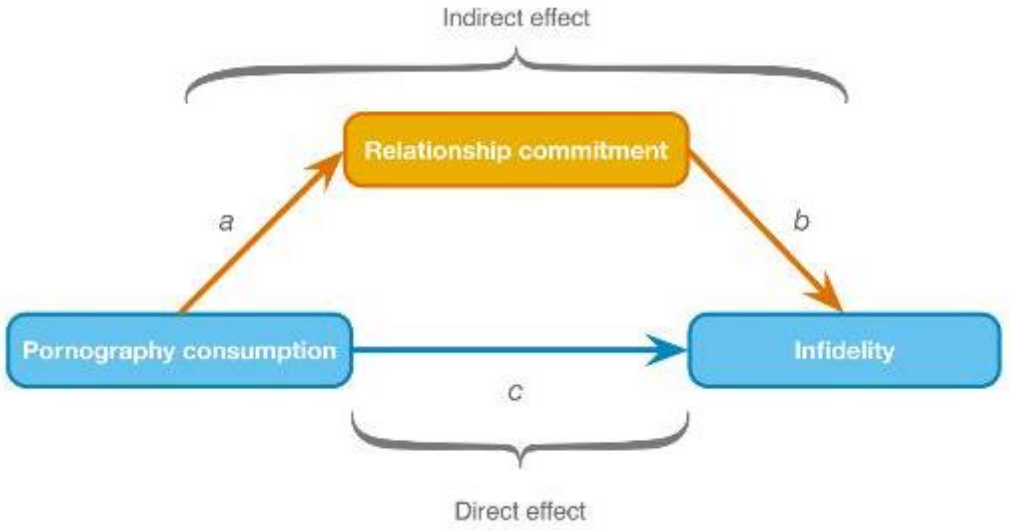
Mediation

Figure 11.9 Diagram of a mediation model



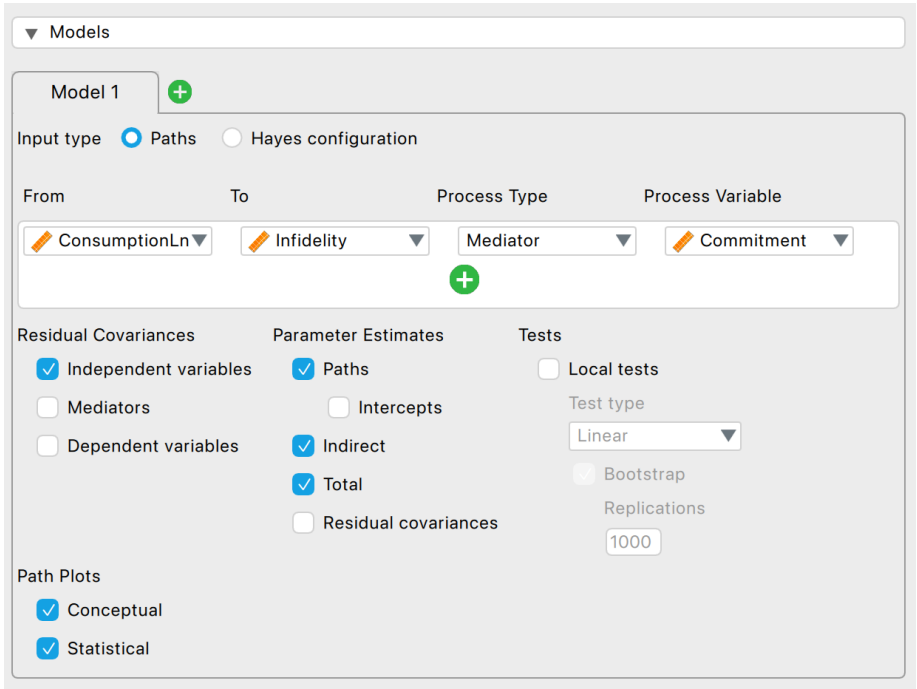
Mediation Example

Figure 10.12 Diagram of a mediation model from Lambert et al. (2012)



Mediation Analysis in JASP

Figure 10.14 The menu for specifying a mediation path



Mediation Model with Two Mediators

Figure 10.16 A mediation model with two mediators (Bronstein, 2019)

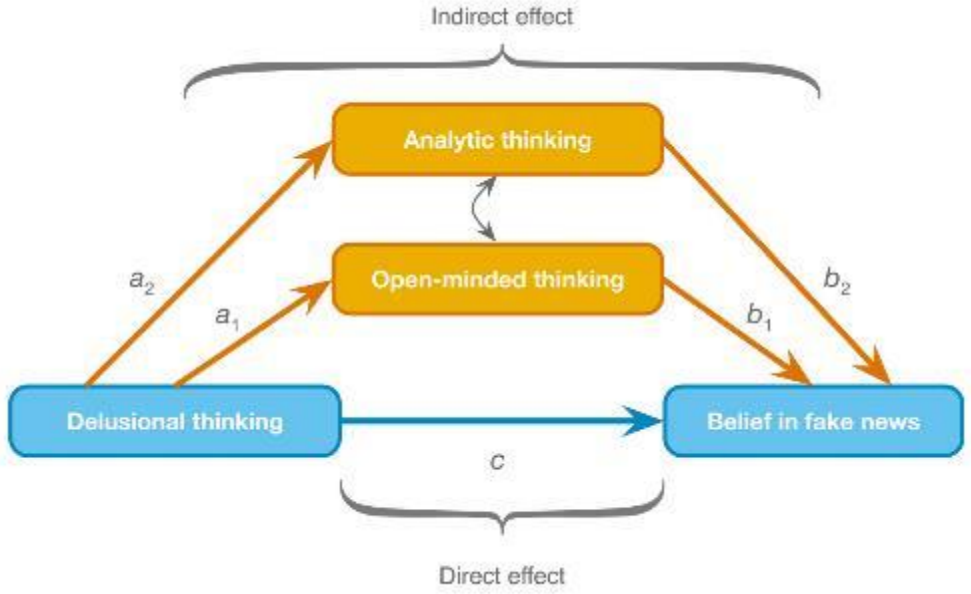
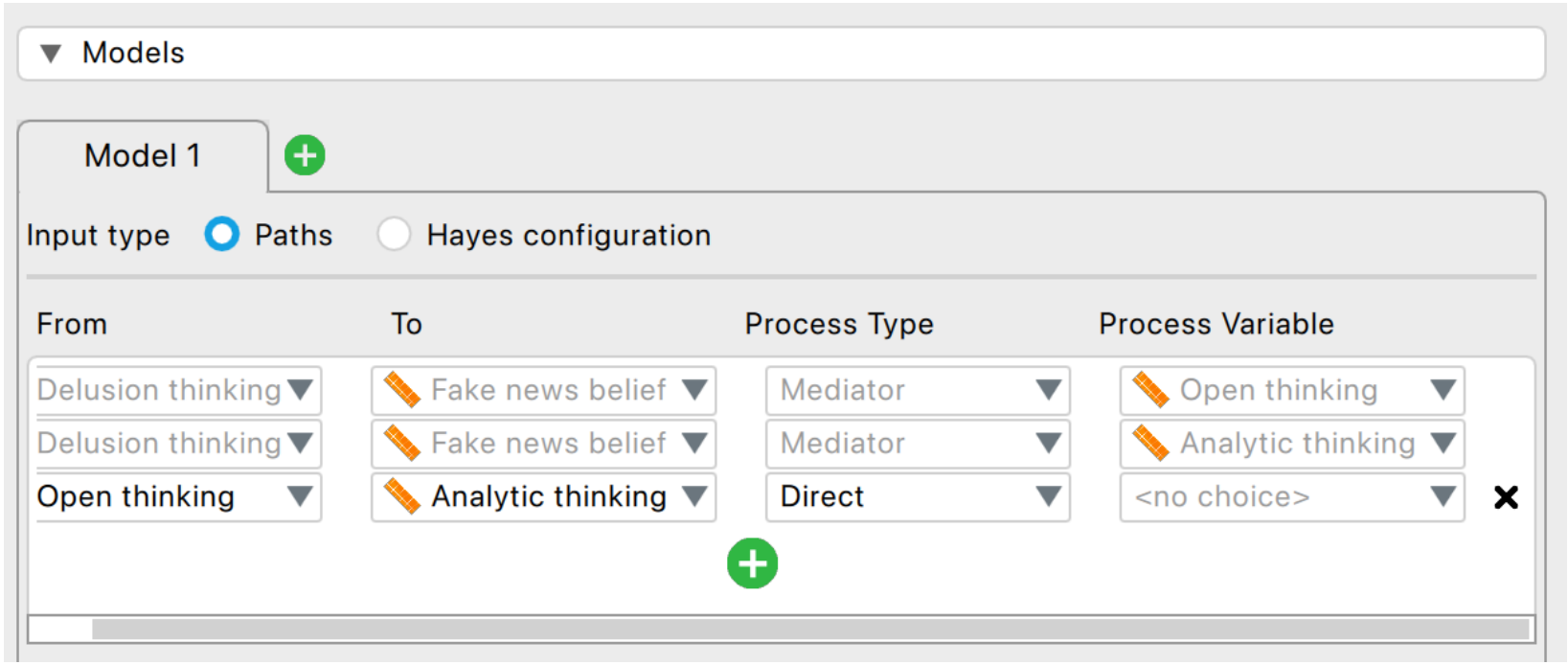


Figure 10.17 The dialogue boxes for running mediation analysis with two mediators

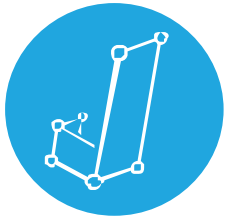


Moderation & Mediation

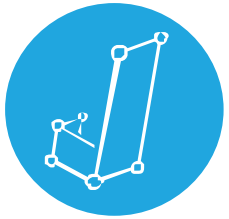
Alex Examples

Leni Examples

Comparing Means

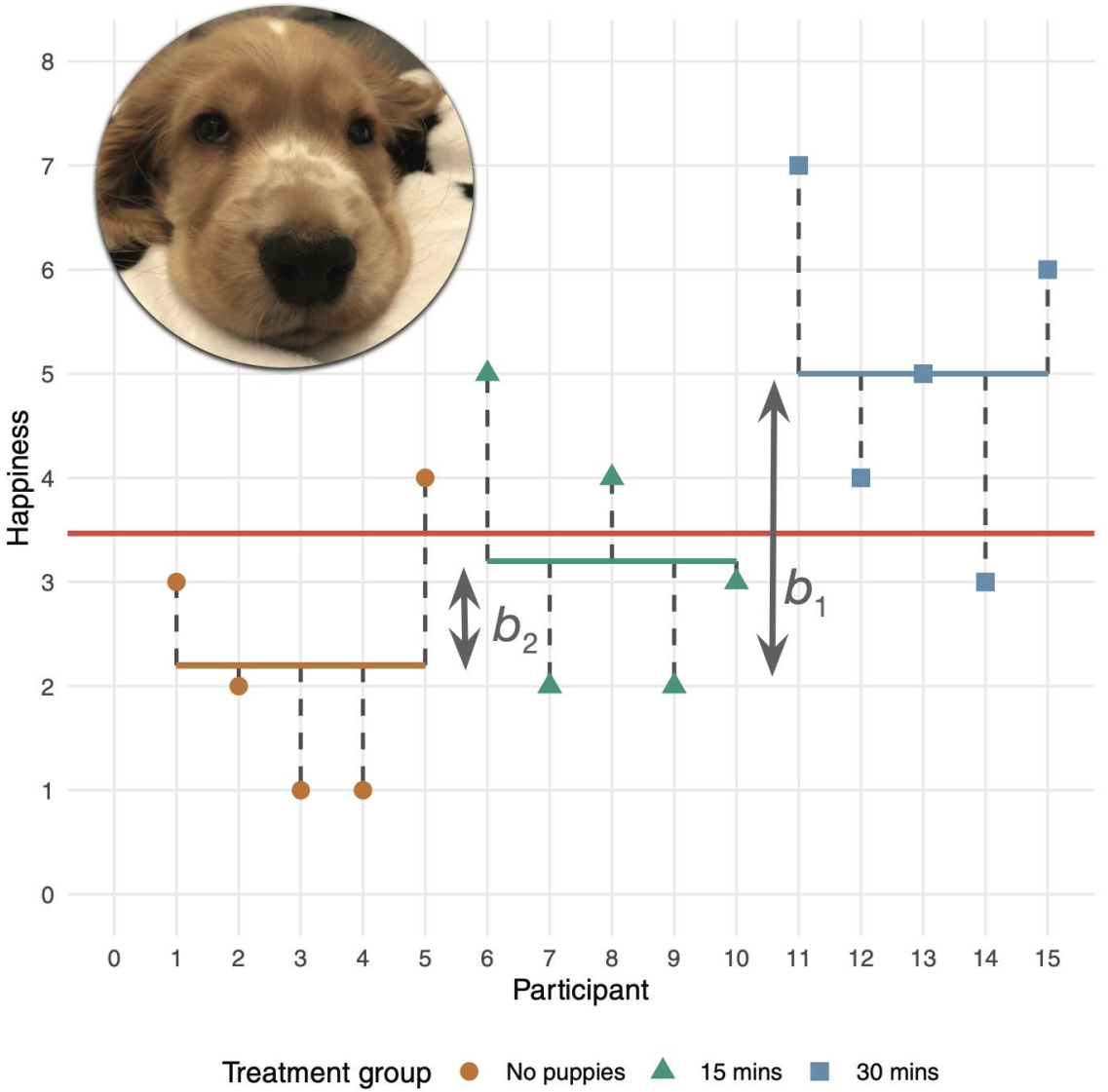


ANOVA: Puppy Example

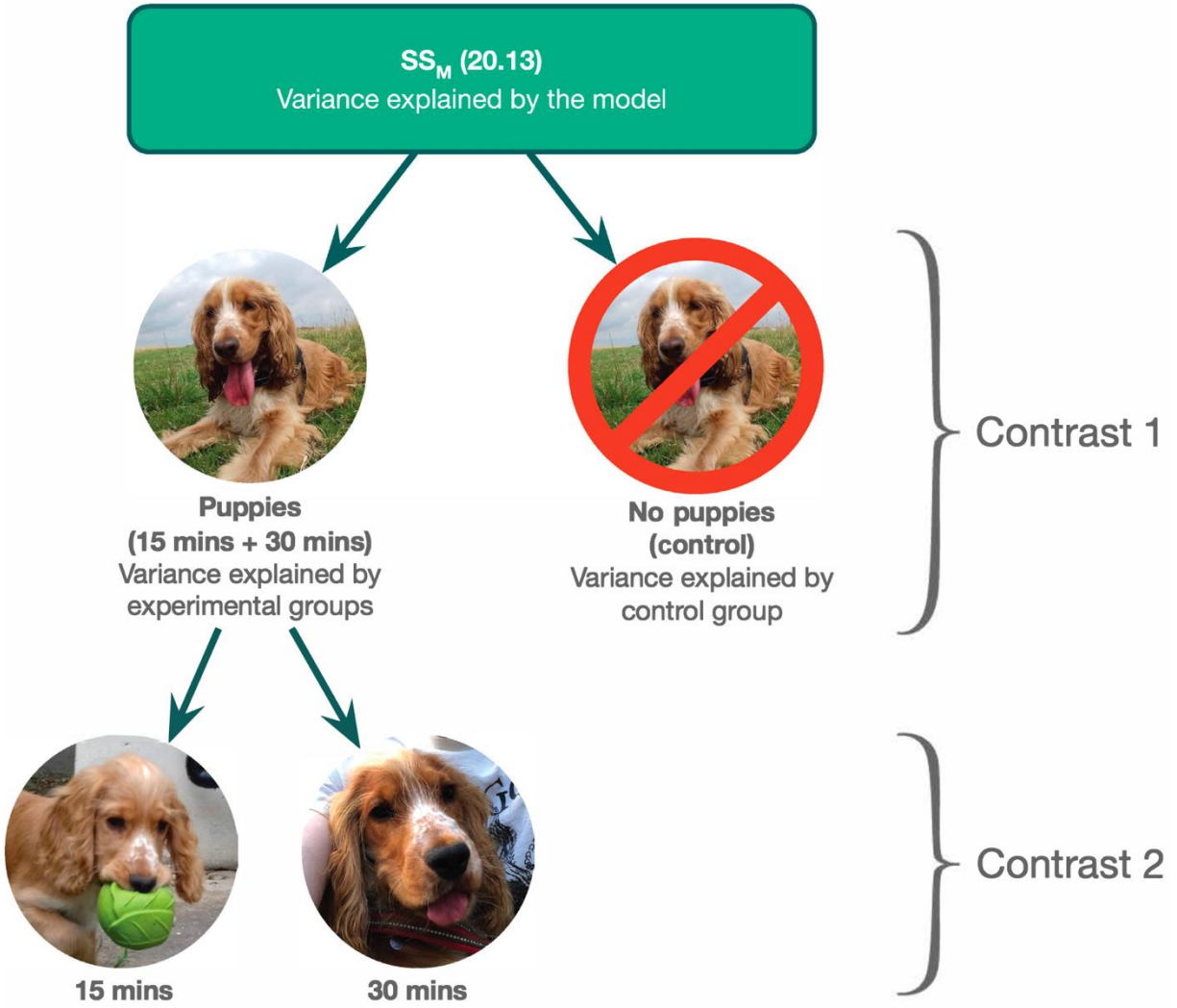


- A puppy therapy RCT in which we randomized people into three groups:
  1. A control group
  2. 15 minutes of puppy therapy
  3. 30 minutes of puppy contact
- The DV is happiness (0 = unhappy) to 10 (happy)
- Predictions:
  1. Any form of puppy therapy should be better than the control (i.e. higher happiness scores).
  2. A dose-response hypothesis that as exposure time increases (from 15 to 30 minutes), happiness will increase too

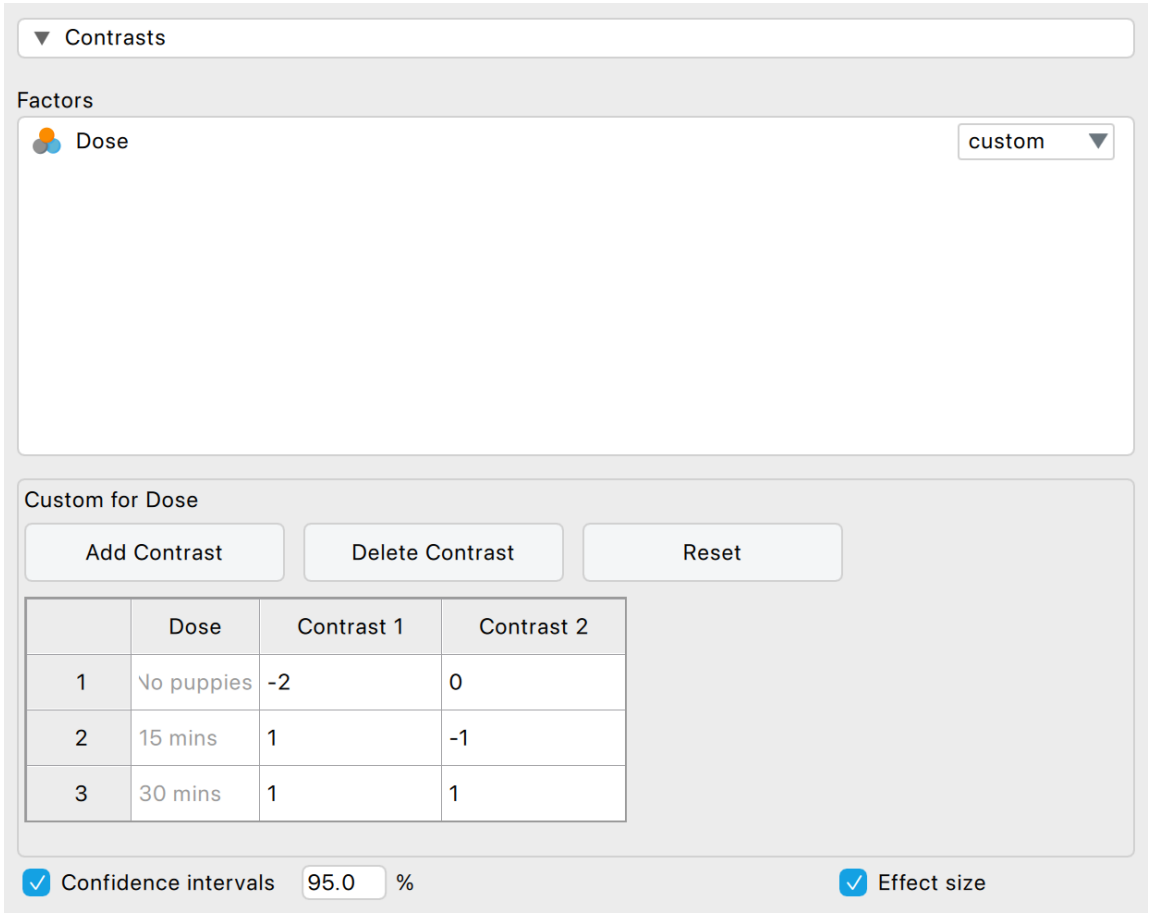
ANOVA


Happiness
8
7
6
5
4
3
2
1
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Participant
b_2
b_1
Treatment group
No puppies
15 mins
30 mins

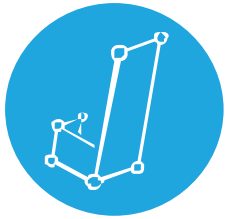
Contrasts


SS_M (20.13)
Variance explained by the model
Puppies
(15 mins + 30 mins)
Variance explained by
experimental groups
No puppies
(control)
Variance explained by
control group
15 mins
30 mins
Contrast 1
Contrast 2

Contrasts in JASP



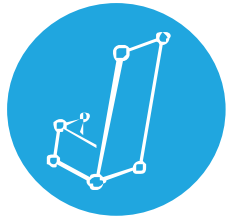
Post Hoc Tests



- Compare each mean against all others.
- In general terms, they use a stricter criterion to accept an effect as significant.
  - Hence, control the family-wise error rate.
  - Simplest example is the Bonferroni method:

$$P_{Crit} = \frac{\alpha}{K}$$

Post Hoc Tests



- Assumptions met:
  - Tukey HSD
- Safe Option:
  - Bonferroni
- Unequal variances:
  - Games-Howell

ANCOVA

- Reduces error variance
  - By explaining some of the unexplained variance (SSR) the error variance in the model can be reduced
- Greater insight
  - By including more variables, we gain deeper insight into their interplay (e.g., interactions, shared variance)
- Warning
  - Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies

ANCOVA

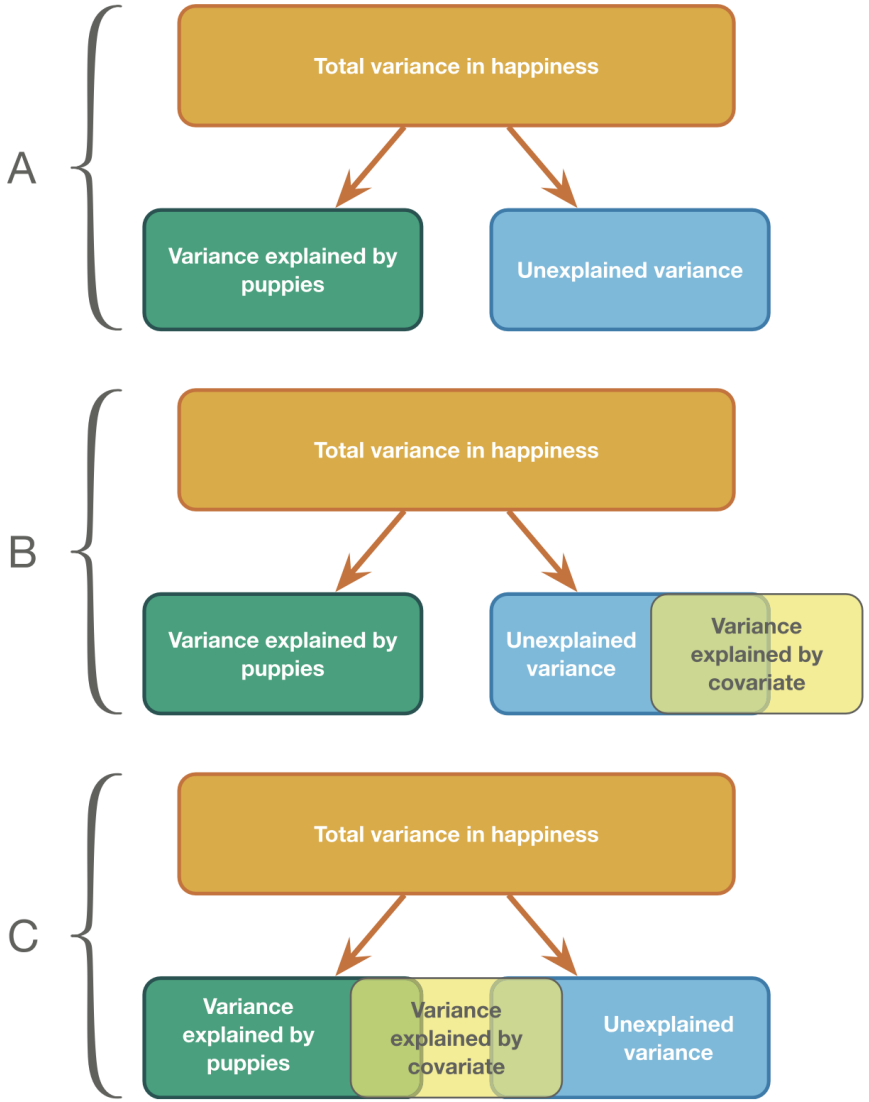
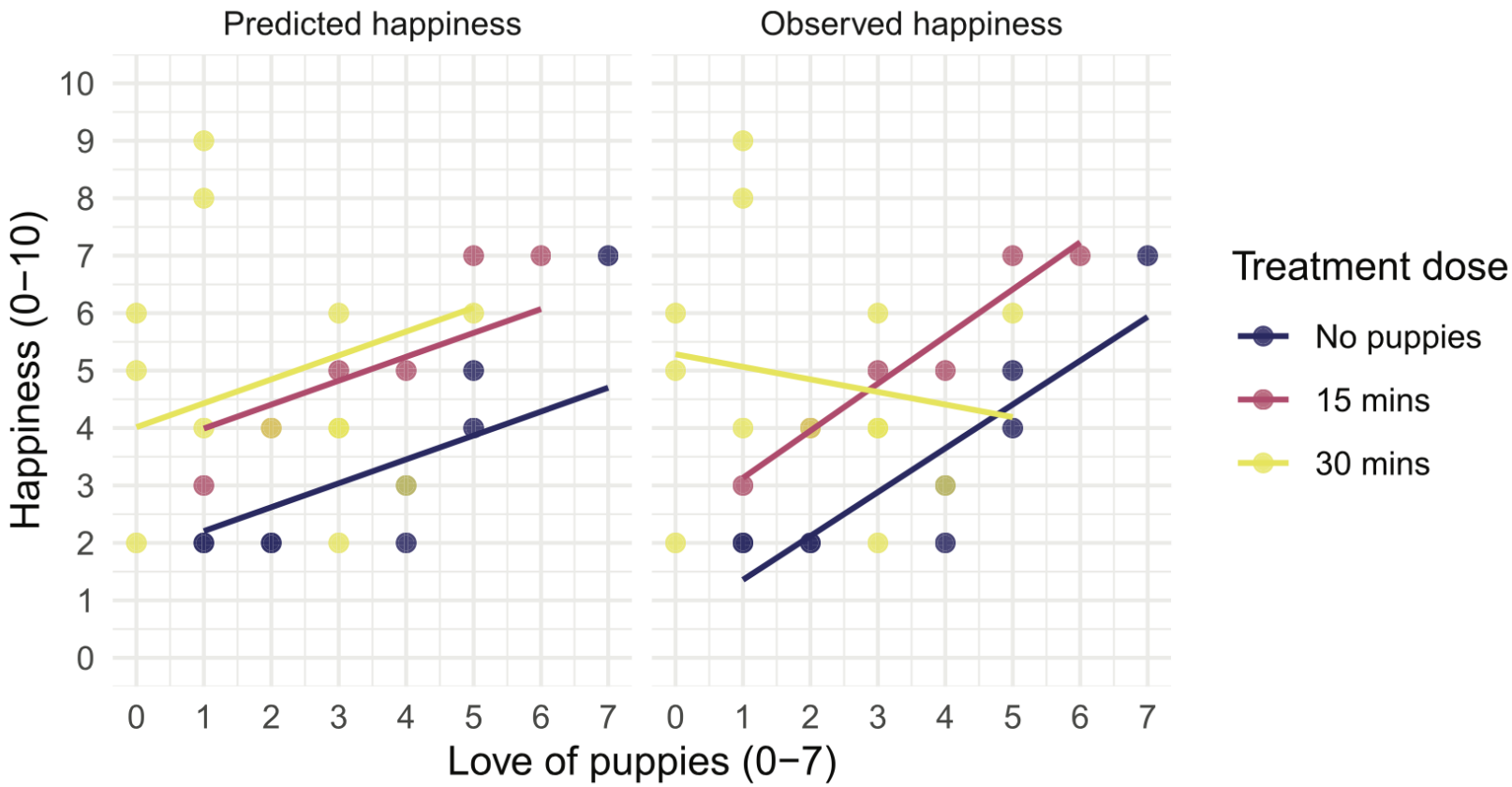


Figure 12.2 The role of the covariate in ANCOVA

Homogeneity of Slopes

Figure 12.3 Scatterplot and linear models of happiness against love of puppies for each therapy condition



Assessing Homogeneity of Slopes

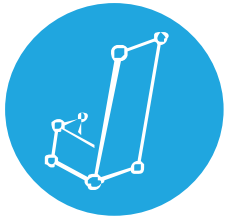
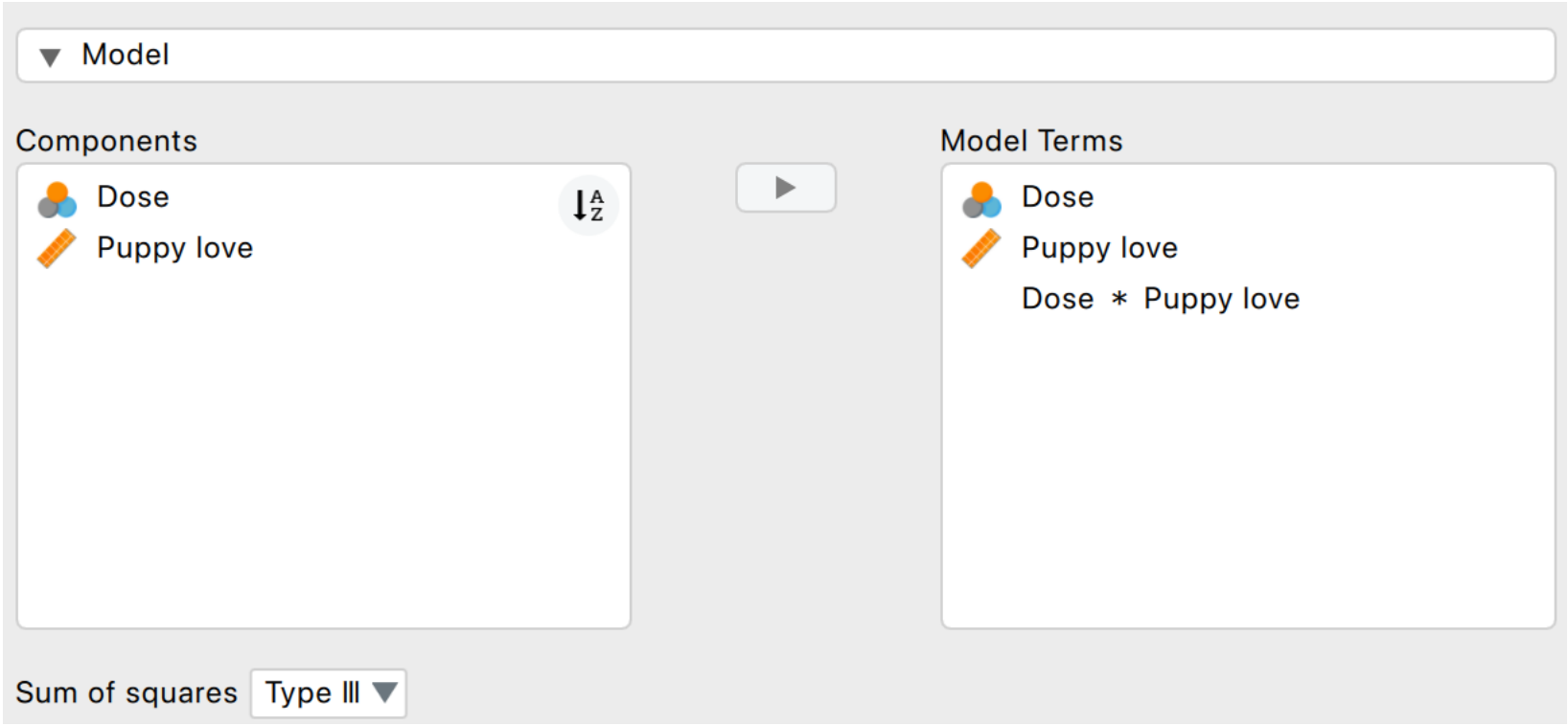


Figure 12.8 *Model* tab for ANCOVA



ANCOVA

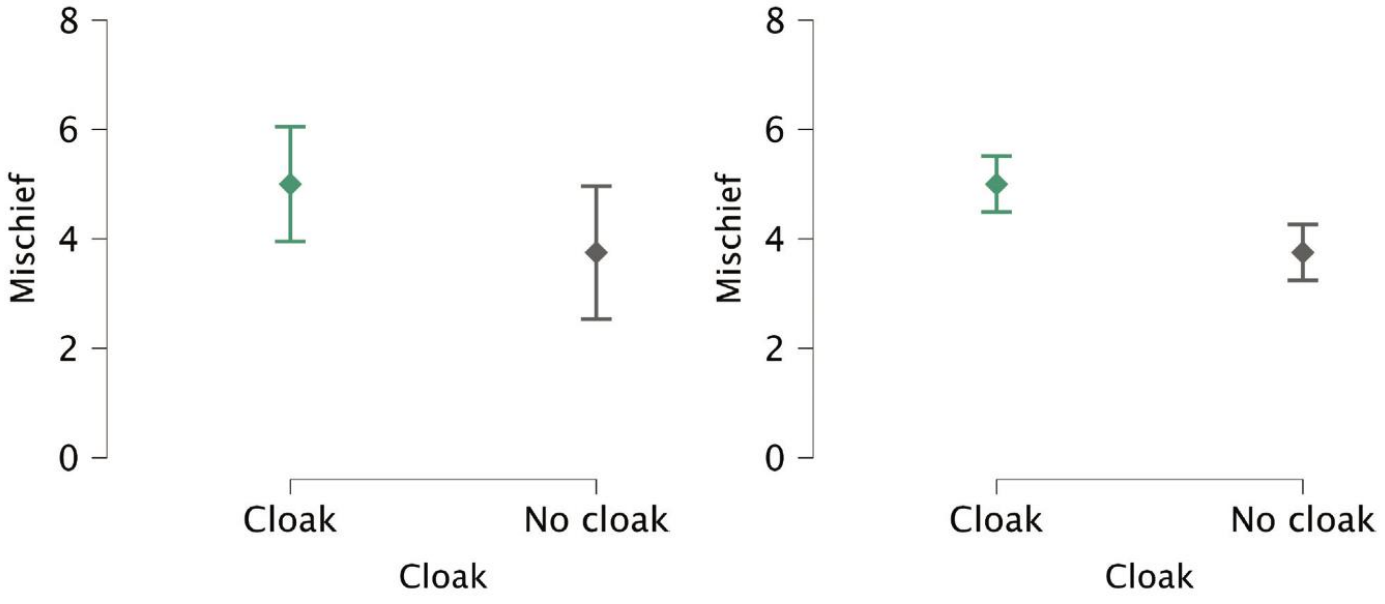
[Alex Examples]

[Leni Examples]

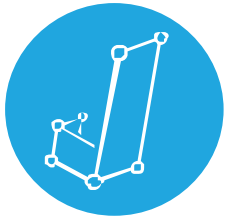
RM ANOVA

- Advantages
  - Unsystematic variance is reduced
  - More sensitive to experimental effects

Figure 9.7 Same data, between-subjects (left) and within-subjects (right)



RM ANOVA Example



- Training sniffer dogs to detect aliens
- After rigorous training, eight dogs sniffed each of four entities for 1 minute:
  - Alien space lizard in its natural form
  - Alien space lizard who had shapeshifted into humanoid form
  - Human
  - Human mannequin
- DV: Number of vocalizations made during each 1-minute sniffing session

Data for Sniffer Dog Example

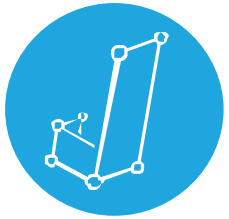
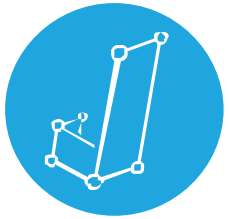


Table 14.1 Data for the sniffer-dog example

Dog	Alien	Human	Mannequin	Shapeshifter	Mean	s^2
Milton	8	7	1	6	5.50	9.67
Woofy	9	5	2	5	5.25	8.25
Ramsey	6	2	3	8	4.75	7.58
Mr. Snifficus III	5	3	1	9	4.50	11.67
Willock	8	4	5	8	6.25	4.25
The Venerable Dr. Waggy	7	5	6	7	6.25	0.92
Lord Scenticle	10	2	7	2	5.25	15.58
Professor Nose	12	6	8	1	6.75	20.92
Mean	**8.13**	**4.25**	**4.13**	**5.75**		

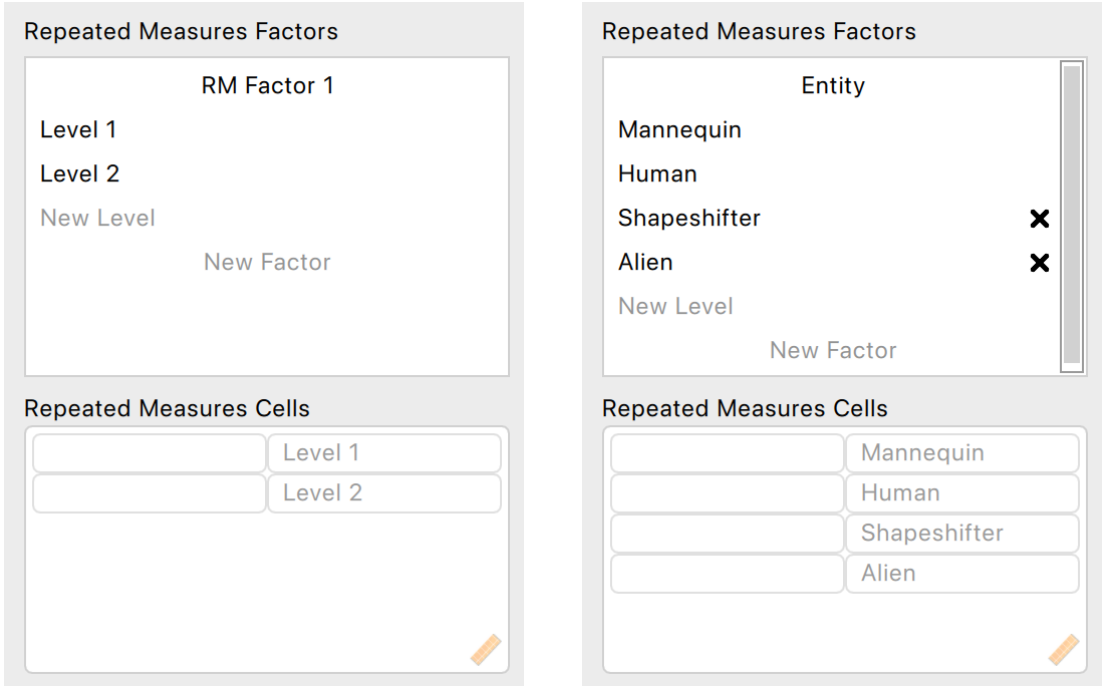
The Assumption of Sphericity



- Assumes that the variances of *differences* between conditions are equal
- Estimated and adjusted *df* using:
  - Greenhouse-Geisser estimate
  - Huynh-Feldt estimate
- Tested using Mauchly's test (not recommended)
  - $P < .05$, sphericity is violated
  - $P > .05$, sphericity is met
- Rule of thumb: G-G is conservative and H-F liberal

Defining the Repeated Factors

Figure 14.6 The *Repeated Measures Factors menu* for repeated-measures ANOVA



Factorial: Post hoc comparisons

Output 14.15

Post Hoc Comparisons - Entity ✻ Scent - Conditional on Entity

				95% CI for Mean Difference					95% CI for Cohen's d		
Entity			Mean Difference	Lower	Upper	SE	t	Cohen's d	Lower	Upper	p_{holm}
Human	None	Human	−1.180	−1.669	−0.691	0.197	−5.980	−0.504	−0.837	−0.170	< .001
		Fox	−4.340	−4.939	−3.741	0.242	−17.950	−1.852	−2.577	−1.128	< .001
	Human	Fox	−3.160	−3.877	−2.443	0.289	−10.932	−1.349	−1.972	−0.726	< .001
Shapeshifter	None	Human	1.640	0.690	2.590	0.383	4.281	0.700	0.096	1.304	< .001
		Fox	1.580	0.611	2.549	0.391	4.043	0.674	0.064	1.285	< .001
	Human	Fox	−0.060	−0.937	0.817	0.354	−0.170	−0.026	−0.538	0.486	0.866
Alien	None	Human	2.080	1.143	3.017	0.378	5.506	0.888	0.262	1.513	< .001
		Fox	2.880	1.835	3.925	0.422	6.833	1.229	0.488	1.970	< .001
	Human	Fox	0.800	−0.099	1.699	0.363	2.207	0.341	−0.196	0.879	0.032

RM ANOVA

[Alex Examples]

[Leni Examples]

JASP

Getting in Touch

- JASP Mattermost channel
- GitHub page
- Priority feature requests