**Organization and structure of the data**

There are 4 main datasets including one raw file and 3 processed files.

Furthermore, there are these script files:

Code:

Cross-correlation R code: this code reads in raw data and creates cross-correlations

Analysis Procedure: Includes all SPSS and R code description of the analysis order

Machine\_Learning\_Code: Python scripts

Cor\_Table\_FDR: R code comparing true couples and mismatched correction table with FDR

Code\_Spss\_Express\_Attraction uses multilevel models to predict attraction

Processed data:

1.behavior\_long\_format.sav: behavioural data including only behavioural ratings collected from all subjects (N = 69)

2.Multilevel\_GLM.sav (N = 54): Multilevel GLM data including participant’s physiology and eye fixations (N=54).

3.Centered (N = 54): Same as Multilevel GLM but with centred variables used for machine learning and Multilevel modelling.

All relevant columns in the 1.behavior\_long\_format.sav include:

Accuracy\_Binominal: 0 = correct, 1 = error

Partner\_likes\_me: 1 = yes, 2 = no

Partner\_really\_likes: 1 = yes, 2 = no

All relevant columns in the 2.Multilevel\_GLM.sav (N = 54) include:

Attraction: scale: 0 to 9

Eyes: mean duration looking at partner’s eyes

Face: mean duration looking at partner’s face

Head: mean duration looking at partner’s head

Body: mean duration looking at partner’s body

Smile: mean duration smiling

Laugh: mean duration laughing

Head\_nod: mean duration nodding

Hand\_gesture: mean duration gesturing

Touch\_face: mean duration touching face

Zscore\_IHR: baseline corrected z-scored heart rate response

Zscore\_EDA: baseline corrected z-scored skin conductance response

All relevant columns in the Centered (N = 54) include:

Dummynonverbal: 1 = nonverbal interaction, 0 = verbal interaction

dummyVerbFirst\_first= 1 = verbal interaction happened first, 0 = verbal interaction happened second

WCCmean\_SCL: synchrony in skin conductance

WCCmean\_IHR: synchrony in heart rate

Both\_Smile\_mean: mean duration of smiling mimicry

Both\_Laugh\_mean= mean duration of laughing mimicry

Both\_Head\_Shake\_mean= mean duration of head nodding mimicry

Both\_Hand\_Shake\_mean= mean duration of hand gesture mimicry

Both\_Head\_Fix\_mean= mean duration of head-to-head fixation

Both\_Face\_Fix\_mean= mean duration of face-to-face fixation

Both\_Eye\_Fix\_mean = mean duration of eye-to-eye fixation

1. Names of the individual who collected the data: Eliska Prochazkova, Danile Lindh, Elio Sjak\_Shie, Friederike Behrens
2. Time period during which data were collected: August 2016 to October 2018
3. Addresses of the field locations where data were collected: Lowlands festival: Walibi Holland at Biddinghuizen, The Netherlands, Nacht van Ontdekkingen : Herengracht 48 2312 LE Leiden, InScince Festival: Arsenaalpoort 8, 6511 PN Nijmegen
4. Names of the authors who analysed the data: Eliska Prochazkova, Daniel Lindh
5. A list of excluded subjects, corrupted data files, and any other exceptions to the standard analysis pipeline, to enhance understanding of the data and the data analysis: In the 2.Multilevel\_GLM.sav there are columns called: EDA\_Outlier (Skin conductance), HFHR\_Outlier (heart rate): 0 = no outlier, 1 = outlier.