

**Digital Appendix TRAIN₄Positivity – Development and Pilot Evaluation of a Mobile-Based Training
of Positivity Bias at the Level of Action Tendencies**

Appendix K3 Detailed Analysis Plan

Data entry into SPSS 23 (*n* = 41)

- Data of screening, pre- and posttest assessment (paper-pencil questionnaires) of TRAIN₄Positivity
- Data of AAT assessment (AAT-CS at pre- and posttest)
- Data of weekly stressor monitoring (i.e., Sunday data; EMA MIMIS; occurred number of microstressors and perceived severity of microstressors¹)
- missing values coded with 99; filter not fulfilled with 999
- Perceived stressor exposure (MIMIS and MIMIS EMA): For microstressors, that did not occur, the rating of perceived severity of the respective microstressors was also set as 0 (in line with the procedures in LifeStress)

A. Data preparation

- **A1. Missing Values Analysis & Expectation Maximization (EM) imputation**
 - **A1.1 Missing Values Analysis & Little's MCAR test**
 - ◆ data set: *"A1.1_data set_Train4Positivity_ORIGINAL"*
 - ◆ syntax: *"A1.1_syntax_Missing Values Analysis & Little's MCAR test"*
 - ◆ analysis of missing data (before EM imputation), separated by metric/quantitative and categorical variables including Little's MCAR test
 - ◆ output: *"A1.1_output_Missing Values Analysis & Little's MCAR test"*
 - **A1.2 Expectation Maximization (EM) imputation**
 - ◆ data set: *"A1.2_data set_Train4Positivity_ORIGINAL"*
 - ◆ syntax: *"A1.2_syntax_EM imputation"*

¹ Not analyzed for pretest or weekly stressor monitoring but only for posttest.

- ◆ EM imputation to create new data set with imputed data *“A1.2_data set_Train4Positivity_IMPUTED”* → also copied into folder “Newly created data set after EM imputation” (as back-up)

- ◆ output: *“A1.2_output_EM imputation”*

Certain variables, that were not considered in the EM imputation and removed from the data set (but are relevant for the MVA analysis), are copied back into the imputed data set (e.g., sociodemographic categorical data)

- *“A1.2_data set_Train4Positivity_IMPUTED_complete”*

- **A1.3 Missing Values Analysis & Little’s MCAR test after EM imputation (as check-up)**

- ◆ data set: *“A1.3_data set_Train4Positivity_IMPUTED_complete”*
- ◆ syntax: *“A1.3_syntax_Missing Values Analysis & Little’s MCAR test after EM imputation”*
- ◆ output: *“A1.3_output_Missing Values Analysis & Little’s MCAR test after EM imputation”*
- Imputed data set is renamed as ***“data set1_Train4Positivity”*** for further analyses

- **A2. Recoding of items; calculation of total scores and difference scores of study outcomes**

- data set: copy of *“A2_data set1_Train4Positivity”* (original data set is kept)
- syntax: *“A2_syntax_Recoding of items, calculation of total scores and difference scores of study outcomes”*
- ◆ see Appendix G5
- ◆ Recoding of items for LOT-R, SDS-17, PSS-10, BRS
- ◆ Calculation of mean scores:
 - ❖ SDS-17, BRS, WHO-5, ERQ-reappraisal
- ◆ Calculation of sum scores:
 - ❖ SPF
 - ❖ LOT-R total score, LOT-R optimism score
 - ❖ ASF-E-P, ASF-E-N
 - ❖ number of macrostressors (LE Checklist)

- ❖ number of microstressors (MIMIS paper pencil), perceived severity of microstressors (MIMIS paper pencil)
- ❖ PSS-10
- ❖ GHQ-28 total score
- ❖ PANAS-PA, PANAS-NA
- ❖ STAXI-state anger, STAXI-anger in, STAXI-anger out, STAXI-anger control
- ❖ CERQ-positive reappraisal
- ◆ Calculation of difference scores for outcome variables:
 - ❖ PSS-10
 - ❖ AAT-CS
 - ❖ GHQ-28
 - ❖ BRS
 - ❖ WHO-5
 - ❖ ERQ-reappraisal
 - ❖ CERQ-positive reappraisal
 - ❖ STAXI-state anger, STAXI-anger in, STAXI-anger out, STAXI-anger control
- ◆ Adjustment of decimals (two decimals for mean scores; no decimals for sum scores) and variable labels
- output: *“A2_output_Recoding of items, calculation of total scores and difference scores of study outcomes”*
- values of several (socio-)demographic variables incorrect (probably due to EM imputation) → adjusted manually
- Further (sociodemographic) variables, that were not considered in the EM imputation and removed from the data set, are copied back into the data set
- Resulting data set including total scores and differences scores for subsequent statistical analyses (RQ1/hypothesis 1.1, RQ2/hypothesis 2.1 and 2.2, RQ3, RQ5, RQ6) is renamed as ***“data set1_Train4Positivity_FINAL”*** (see Appendix I3)

- **A3. Outliers and extreme values**

- Analysis of outliers and extreme values for baseline variables and covariates, outcomes and difference scores of study outcomes
 - ◆ data set: *"A3_data set1_Train4Positivity_FINAL"*
 - ◆ syntax: *"A3_syntax_Outliers and extreme values"*
 - ❖ Descriptive statistics, extreme values, tests for normality, histograms, Q-Q plots, Box-whisker plots
 - ◆ output: *"A3_output_Outliers and extreme values"*

B. Reliabilities (internal consistencies) of questionnaires used

- data set: *"B_data set1_Train4Positivity_FINAL"*
- syntax: *"B_syntax_reliabilities"*
- outputs:
 - *"B_output_reliabilities_baseline&covariates"*
 - *"B_output_reliabilities_outcomes"*

C. Sample description

- **C1. Descriptive statistics of baseline variables and covariates, outcome variables, and of difference scores of study outcomes for sample description**
 - data set: *"C1_data set1_Train4Positivity_FINAL"* (copy of original data set for analysis)
 - syntax: *"C1_syntax_descriptive statistics"*
 - outputs:
 - ◆ *"C1_output_Descriptive statistics sociodemographic, lifestyle, and clinical variables"*
 - ◆ *"C1_output_Descriptive statistics baseline variables and covariates"*
 - ◆ *"C1_output_Descriptive statistics outcome variables and difference scores"*
 - ◆ *"C1_output_Descriptive statistics evaluation of training (posttest)"*

- **C2. Classification of baseline variables/covariates and outcome variables according to norm values (if available) or comparison with values from other German (representative) samples**
 - data set: *“C2_data set1_Train4Positivity_FINAL”* (original data set for analysis is kept in the folder)
 - syntax: *“C2_syntax_classification (norms)_comparative values”*
 - ◆ Classification of the TRAIN₄Positivity sample in accordance with norm values (if available): LOT-R (T values), ASF-E-P (T values), ASF-E-N (T values), BRS (Stanine values), STAXI-anger in (Stanine values), STAXI-anger out (Stanine values), STAXI-anger control (Stanine values)

T values:

- ❖ Calculate new variable (e.g., based on respective pretest/posttest value of covariate or outcome variable) that indicates T values
- ❖ Classification of T values in three subgroups based on conventions (T values: M = 50, SD = 10)
- ❖ Frequency of each subgroup in the TRAIN₄Positivity sample

Stanine values:

- ❖ Calculate new variable (e.g., based on respective pretest/posttest value of covariate or outcome variable) that indicates Stanine values
- ❖ Classification of T values in three subgroups based on conventions (Stanine values: M = 5, SD = 2)
- ❖ Frequency of each subgroup in the TRAIN₄Positivity sample
- ◆ Percentiles for comparison with percentiles of other German (representative) samples: PSS-10 (percentiles), WHO-5 (percentile ranks)
- ◆ For STAXI subscales anger in, anger out, and anger control, new values representing Stanine values (based on the STAXI total scores at pre- and posttest) are created manually

- ◆ Newly created variables are moved behind respective relevant variables
- ◆ Preparation of independent t tests for some variables (e.g., by analyzing descriptive statistics for ERQ in men and women)
- ◆ Independent t tests using separate data sets for each variable (see folder “Independent t tests_data sets”) → each data set (e.g., *data set_sample description_t test_comparative values_BRS_pre*) includes x_1 , sd_1 , n_1 , x_2 , sd_2 , n_2 of TRAIN₄Positivity and comparative sample (e.g. other German representative sample); t value, df and p value are supplemented automatically
- outputs:
 - ◆ “C2_output_classification_baseline&covariates”
 - ◆ “C2_output_classification_outcomes”
 - ◆ “C2_output_preparation_t tests”

D. Testing of research questions and hypotheses

- **D. RQ1_hypothesis 1.1_Effects on perceived stress**

- **a. Paired t test**
 - ◆ data set: “D_RQ1_1.1a_data set1_Train4Positivity_FINAL”
 - ◆ syntax: “D_RQ1_1.1a_syntax_t test_PSS-10”
 - ❖ Testing of normality of difference score of PSS-10
 - ❖ (independence of observations within a sample, i.e., time point) not tested
 - ❖ Paired t test
 - ◆ output: “D_RQ1_1.1a_output_t test_PSS-10”
- **b. Multiple linear regression**
 - ◆ data set: “D_RQ1_1.1b_data set1_Train4Positivity_FINAL” (original data set for analysis is kept in the folder)
 - ◆ syntax: “D_RQ1_1.1b_syntax_multiple regression_PSS-10”

- ❖ z standardization of predictors
 - ❖ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - ❖ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - ❖ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
 - ◆ output: *“D_RQ1_1.1b_output_multiple regression_PSS-10”*
- **D. RQ1_exploratory RQ 1.2_Effects on perceived microstressor severity (matched control group design)**
 - For preparation of LifeStress data and combination of data from TRAIN₄Positivity and LifeStress, see folder F.
 - **Further preparation for PS matching**
 - ◆ data set (**generation of this data set is explained in folder F.**): *“Data set2_LifeStress & Train4Positivity_FINAL”*
 - ◆ syntax (do file in Stata): *“D_RQ1_1.2_syntax_further preparation data set PS matching”*
 - ◆ output (smcl file in Stata): *“D_RQ1_1.2_output_further preparation for PS matching”*
 - ◆ → data set with changes is saved as **“D_RQ1_1.2_data set2_FINAL”** which is used for the next steps
 - **Pre-analyses:**
 - ◆ data set: *“D_RQ1_1.2_data set2_FINAL”*
 - ◆ syntax (do file): *“D_RQ1_1.2_syntax_perceived microstressor severity”*
 - ❖ Between-group comparison of intervention group (TRAIN₄Positivity) and control group (LifeStress) in variables that are considered as potential covariates: age,

- gender, education, mental health (GHQ-28), sum microstressors (MIMIS), sum macrostressors (LE Checklist)
- ❖ independent t test for metric variables age, GHQ-28, MIMIS, LE Checklist
- ❖ χ^2 test for categorical variables gender and education
- ◆ output (smcl file): *"D_RQ1_1.2_output_PS matching_pre-analyses"*
- STEP 1 & 2: Common support assumption
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*
 - ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*
 - ❖ step 1: estimation of propensity score
 - ❖ step 2 (within step 1): common support/overlap assumption → testing of two possible methods (Minima Maxima comparison vs. trimming methods)
 - ❖ for detailed explanations, see do file
 - ❖ first matching algorithm (1-to-1 matching without replacement) is tested without and with "common" command (for Minima Maxima comparison) as well as different variants of trimming method (1%, 5%, 25%)
 - ◆ output (smcl file):
 - ❖ *"D_RQ1_1.2_output_STEP 1 & 2_Common support assumption"*
 - ❖ corresponding graphs (PS distribution, PS density plot) 1–10 are saved separately
- Generation of PS density plot before matching
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*
 - ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*
 - ◆ outputs:
 - ❖ *"D_RQ1_1.2_output_graph of PS scores before matching.smcl"*
 - ❖ *"Graph11_PS density plot for both groups before matching.gph"*
- STEP 3: Selection of final matching algorithm
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*

- ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*
 - ❖ Based on PS distribution by condition, PS density plot, statistical tests (*t* tests) concerning the distribution of covariates in the two groups, Pseudo R^2 and the *p* value of Likelihood Ratio (LR) test, 12 different matching algorithms are tested (all using "common" command; see step 1)
- ◆ outputs:
 - ❖ *"D_RQ1_1.2_output_STEP 3_selection of final algorithm"*
 - ❖ corresponding graphs 12–35 for PS distribution and PS density plot are saved separately
- STEP 4: Selection of variables
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*
 - ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*
 - ❖ For the algorithm with the best matching quality among the 12 algorithms tested (2-nearest neighbor matching with replacement), different covariates are tested:
 - ✚ three original covariates (age, gender, dummy variables for education)
 - ✚ three original covariates + mental health (GHQ-28)
 - ✚ three original covariates + number of microstressors (MIMIS)
 - ✚ three original covariates + number of macrostressors (LE Checklist)
 - ◆ outputs:
 - ❖ *"D_RQ1_1.2_output_STEP 4_selection of variables.smcl"*
 - ❖ corresponding graphs 36–43 for PS distribution and PS density plot are saved separately
- STEP 5: Sensitivity analysis for final algorithm and final model
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*
 - ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*

- ❖ sensitivity analysis for 2-nearest neighbor matching with replacement with age, gender, and education
- ◆ output: *"D_RQ1_1.2_output_STEP 5_sensitivity analysis"*
- STEP 6: Matching using teffects psmatch
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*
 - ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*
 - ❖ use of *"teffects psmatch"* instead of *"psmatch2"*
 - ❖ generation of graph for testing of overlap assumption
 - ◆ outputs:
 - ❖ *"D_RQ1_1.2_output_STEP 6_teffects psmatch"*
 - ❖ graphs 44–45 (PS density plot before and after matching, graph for overlap assumption) are saved separately
- STEP 7: Saving of matched data for further analyses and posttest comparison between matched groups (which is relevant for hypothesis testing)
 - ◆ data set: *"D_RQ1_1.2_data set2_FINAL"*
 - ◆ syntax (do file): *"D_RQ1_1.2_syntax_perceived microstressor severity"*
 - ❖ → matched data are saved in a new data set: *"D_RQ1_1.2_data set2_FINAL_matched data"*
 - ❖ testing of assumptions for independent t test (outliers, normality, homoscedasticity)
 - ❖ independent- t test
 - ◆ output: *"D_RQ1_1.2_output_STEP 7_posttest comparison"*
- **D. RQ2_hypothesis 2.1_Effects on implicit action tendencies**
 - **Paired t test**
 - ◆ data set: *"D_RQ2_2.1a_data set1_Train4Positivity_FINAL"*
 - ◆ syntax: *"D_RQ2_2.1a_syntax_t test_AAT-CS"*

- ❖ Testing of normality of difference score of AAT-CS
- ❖ (independence of observations within a sample, i.e., time point) not tested
- ❖ Paired *t* test
- ◆ output: *"D_RQ2_2.1a_output_t test_AAT-CS"*
- **b. Multiple linear regression**
 - ◆ data set: *"D_RQ2_2.1b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
 - ◆ syntax: *"D_RQ2_2.1b_syntax_multiple regression_AAT-CS"*
 - ❖ z standardization of predictors
 - ❖ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - ❖ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - ❖ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
 - ◆ output: *"D_RQ2_2.1b_output_multiple regression_AAT-CS"*
- **D. RQ2_hypothesis 2.2_Moderation of effects on implicit action tendencies by baseline attributional style**
 - **Median & median split**
 - ◆ data set: *"D_RQ2_2.2a_data set1_Train4Positivity_FINAL_median split"* (original data set for analysis is kept in the folder)
 - ◆ syntax: *"D_RQ2_2.2a_syntax_median & median split"*
 - ❖ Determination of median for ASF-E-P and ASF-E-N (later between-subjects factors in three-way repeated-measures ANOVA)
 - ❖ Calculate two new categorical variables (0: < median; 1: ≥ median) using median split

- ◆ output: *"D_RQ2_2.2a_output_median & median split"*
- **b. Three-way repeated-measures (RM) ANOVA**
 - ◆ data set: *"D_RQ2_2.2b_data set1_Train4Positivity_FINAL_median split"*
 - ◆ syntax: *"D_RQ2_2.2b_syntax_3-way RM ANOVA"*
 - ❖ (repeated-measures) within-subjects factor: time with pre- and posttest
 - ❖ between-subjects factors: low versus high ASF-E-P, low versus high ASF-E-N
 - ❖ testing of normality of dependent variable for each combination of factor levels
(using Shapiro-Wilk test)
 - ❖ further prerequisites (homogeneous variances and homogenous variance-covariance matrices between the levels of the non-repeated factor(s)) using Box test and Levene's test tested within 3-way RM ANOVA
 - ◆ output: *"D_RQ2_2.2b_output_3-way RM ANOVA"*
- **D. RQ3_Effects on resilience, ability to recover from stress and well-being**
 - **Paired *t* tests (BRS, WHO-5) and Wilcoxon signed-rank test (GHQ-28)**
 - ◆ data set: *"D_RQ3a_data set1_Train4Positivity_FINAL"*
 - ◆ syntax: *"D_RQ3a_syntax_t test_WSR test_GHQ-28, BRS, WHO-5"*
 - ❖ Testing of normality of difference score of GHQ-28, BRS and WHO-5 → Normality assumption not fulfilled for GHQ-28
 - ❖ (independence of observations within a sample, i.e., time point) not tested
 - ❖ Paired *t* tests for BRS and WHO-5
 - ❖ Wilcoxon signed-rank test (non-parametric) for GHQ-28
 - ◆ output: *"D_RQ3a_output_t test_WSR test_GHQ-28, BRS, WHO-5"*
 - **b. Multiple linear regression (GHQ-28, BRS, WHO-5)**
 - ◆ **GHQ-28 – 2 predictors (only stressor exposure):**

- ❖ data set: *"D_RQ3b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
- ❖ syntax: *"D_RQ3b_syntax_multiple regression_GHQ-28"*
 - ✚ z standardization of predictors
 - ✚ Multiple linear regression (including covariates) – parsimonious model with stressor exposure to micro- and macrostressors to represent resilience (2 predictors)
 - ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ❖ output: *"D_RQ3b_output_multiple regression_GHQ-28"*
- ◆ **BRS:**
 - ❖ data set: *"D_RQ3b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
 - ❖ syntax: *"D_RQ3b_syntax_multiple regression_BRS"*
 - ✚ z standardization of predictors
 - ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors) → however, based on partial regression diagram, no linear association between microstressors and BRS (inverse U-shaped relationship)
 - ✚ Multiple linear regression (including covariates) – parsimonious model (3 predictors, i.e., without microstressors)
 - ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis

- ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ❖ output: “D_RQ3b_output_multiple regression_BRS”
- ◆ **WHO-5:**
- ❖ data set: “D_RQ3b_data set1_Train4Positivity_FINAL” (original data set for analysis is kept in the folder)
- ❖ syntax: “D_RQ3b_syntax_multiple regression_WHO-5”
- ✚ z standardization of predictors
- ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
- ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
- ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ❖ output: “D_RQ3b_output_multiple regression_WHO-5”
- **D. RQ4_Effects on psychological EMA outcomes (current mood with positive valence, emotional arousal, well-being; three end-of-day measures well-being, ability to get affected by positive stimuli, ability to distance from negative stimuli)**
 - **a. Current mood**
 - ◆ preparation of data set:
 - ❖ data entry into SPSS using Excel file that was generated by the “Breezly” app and contains the participants’ data ($n = 41$) concerning the three mood scales (see Appendix K8) → “D_RQ4a_data set current mood raw”

- ❖ Recoding of items of Multidimensional Mood Questionnaire (MDMQ) and calculation of total scores for positive valence, emotional arousal, and well-being
 - ✚ data set: copy of *"D_RQ4a_data set_current mood_raw"* → after recoding & calculation of total scores renamed to ***"D_RQ4a_data set3_current mood_FINAL"***
 - ***file is saved as .dta format (for further analyses in Stata)***
 - ✚ syntax: *"D_RQ4a_syntax_current mood_recoding & total scores"*
 - ✚ output: *"D_RQ4a_output_current mood_recoding & total scores"*
- ◆ Missing Values Analysis for current mood variables:
 - ❖ data set: *"D_RQ4a_data set3_current mood_FINAL"*
 - ❖ syntax (do file in Stata): *"D_RQ4a_syntax_current mood_Missing Values Analysis"*
 - ✚ number and pattern of missing data for the three mood outcomes
 - ✚ checking of MCAR assumption
 - ❖ output (smcl format): *"D_RQ4a_output_current mood_Missing Values Analysis"*
- ◆ **Positive valence:**
 - ❖ data set: *"D_RQ4a_data set3_current mood_FINAL"*
 - ❖ syntax (do file in Stata): *"D_RQ4a_syntax_current mood_valence"*
 - ✚ a. Means by time for positive valence
 - ✚ b. Create filter variable to count the number of missing values for the maximum set of variables in order to estimate the multilevel models with the same number of participants
 - ✚ c. Scatter plot by individual participants
 - ✚ d. Exclusion of six cases with ≤ 3 observations (userid 23, 24, 28, 29, 30, 48)
 - ✚ e. Scatter plots to explore the data structure and test the linear change over time
 - linear trend for change in positive valence over time
 - ✚ f. Calculation of different linear mixed effects models
 - ◇ 1) Null model without predictors

- ◇ 2) (2-level) Random-intercept model (fixed linear time/level-1 predictor with random intercepts)
- ◇ 3a) (2-level) Random-coefficients model (random linear time/level-1 predictor with random intercepts and random slopes); no covariance structure pre-specified
- ◇ 3b) (2-level) Random-coefficients model; exchangeable or Compound Symmetry (CS) covariance matrix
- ◇ 3c) (2-level) Random-coefficients model, unstructured covariance matrix

- ✚ g. Fitting plots for best-fitting model; for positive valence: (2-level) Random-coefficients model with unstructured covariance matrix
- ✚ h. Testing of assumptions for multilevel modeling (linearity, normal distribution of residuals, homoscedasticity)
- ✚ i. Multiple Imputation (MI) with accounting for clustering in data; **for three mood outcomes simultaneously**
- ✚ j. Multilevel modeling after MI ("mi estimate") → (2-level) Random-coefficients model with unstructured covariance matrix
- ✚ Graphs 1–10 are saved in separate folder
- ❖ output (smcl format): *"D_RQ4a_output_current mood_valence"*

◆ **Energetic arousal:**

- ❖ data set: *"D_RQ4a_data set3_current mood_FINAL"*
- ❖ syntax (do file in Stata): *"D_RQ4a_syntax_current mood_arousal"*
- ✚ see under "Positive valence"
- ✚ Graphs 1–10 are saved in separate folder
- ❖ output (smcl format): *"D_RQ4a_output_current mood_arousal"*

◆ **Calmness:**

- ❖ data set: *"D_RQ4a_data set3_current mood_FINAL"*

- ❖ syntax (do file in Stata): *"D_RQ4a_syntax_current mood_calmness"*
 - ✚ see under "Positive valence"
 - ✚ Graphs 1–10 are saved in separate folder
- ❖ output (smcl format): *"D_RQ4a_output_current mood_calmness"*
- **b. End-of-day measures**
 - ◆ immediate data entry into Stata using Excel file (see Appendix K8) that was generated by the "Breezly" app and contains the participants' data ($n = 41$) concerning the three end-of-day measures (.dta format)
 - ◆ Missing Values Analysis for end-of-day measures:
 - ❖ data set: *"D_RQ4b_data set4_end-of-day_FINAL"*
 - ❖ syntax (do file in Stata): *"D_RQ4b_syntax_end-of-day_Missing Values Analysis"*
 - ✚ number and pattern of missing data for the three end-of-day outcomes
 - ✚ checking of MCAR assumption
 - ❖ output (smcl format): *"D_RQ4b_output_end-of-day_Missing Values Analysis"*
 - ◆ **Well-being:**
 - ❖ data set: *"D_RQ4b_data set4_end-of-day_FINAL"*
 - ❖ syntax (do file in Stata): *"D_RQ4b_syntax_end-of-day_well-being"*
 - ✚ a. Means by time for well-being
 - ✚ b. Create filter variable to count the number of missing values for the maximum set of variables in order to estimate the multilevel models with the same number of participants
 - ✚ c. Scatter plot by individual participants
 - ✚ d. Exclusion of eight cases with ≤ 3 observations (userid 23, 24, 26, 27, 28, 36, 40, 48)
 - ✚ e. Scatter plots to explore the data structure and test the linear change over time
→ linear trend for change in well-being over time
 - ✚ f. Calculation of different linear mixed effects models

- ◇ 1) Null model without predictors
- ◇ 2) (2-level) Random-intercept model (fixed linear time/level-1 predictor with random intercepts)
- ◇ 3a) (2-level) Random-coefficients model (random linear time/level-1 predictor with random intercepts and random slopes); no covariance structure pre-specified
- ◇ 3b) (2-level) Random-coefficients model; exchangeable or Compound Symmetry (CS) covariance matrix
- ◇ 3c) (2-level) Random-coefficients model, unstructured covariance matrix
- + g. Fitting plots for best-fitting model; for well-being: (2-level) Random-coefficients model with unstructured covariance matrix
- + h. Testing of assumptions for multilevel modeling (linearity, normal distribution of residuals, homoscedasticity)
- + i. Multiple Imputation (MI) with accounting for clustering in data; **due to too few observations, MI not possible for all three end-of-day measures simultaneously**
→ **separate MI for each of the end-of-day measures**
- + j. Multilevel modeling after MI ("mi estimate") → (2-level) Random-coefficients model with unstructured covariance matrix
- + Graphs 1–10 are saved in separate folder
- ❖ output (smcl format): *"D_RQ4b_output_end-of-day_well-being"*
- ◆ **Ability to get affected by positive stimuli:**
 - ❖ data set: *"D_RQ4b_data set4_end-of-day_FINAL"*
 - ❖ syntax (do file in Stata): *"D_RQ4b_syntax_end-of-day_ability to get affected by positive"*
 - + see under "Well-being"
 - + Graphs 1–10 are saved in separate folder

- ❖ output (smcl format): *"D_RQ4b_output_end-of-day_ability to get affected by positive"*
- ◆ **Ability to distance from negative stimuli:**
 - ❖ data set: *"D_RQ4b_data set4_end-of-day_FINAL"*
 - ❖ syntax (do file in Stata):
 - ✚ a. to f. identical to well-being and ability to get affected by positive stimuli
 - ✚ g. Fitting plots for best-fitting model; for ability to distance from negative stimuli:
 - (2-level) Random-intercept model**
 - ✚ Graphs 1–10 are saved in separate folder
 - ❖ output (smcl format): *"D_RQ4b_output_end-of-day_ability distance from negative"*
- **D. RQ5_Effects on emotion regulation (reappraisal, ERQ, CERQ) and emotional experience (PANAS-PA, PANAS-NA, STAXI-state anger, STAXI-anger in, STAXI-anger out, STAXI-anger control)**
 - **Paired *t* tests (ERQ, CERQ, PANAS-NA, STAXI-anger in, STAXI-anger control) and Wilcoxon signed-rank test (PANAS-PA, STAXI-state anger, STAXI-anger out)**
 - ◆ data set: *"D_RQ5a_data set1_Train4Positivity_FINAL"*
 - ◆ syntax: *"D_RQ5a_syntax_t test_WSR test_ERQ, CERQ, PANAS, STAXI"*
 - ❖ Testing of normality of difference score of ERQ, CERQ, PANAS-PA, PANAS-NA, STAXI-state anger, STAXI-anger in, STAXI-anger out, STAXI-anger control → Normality assumption not fulfilled for PANAS-PA, STAXI-state anger, STAXI-anger out
 - ❖ (independence of observations within a sample, i.e., time point) not tested
 - ❖ Paired *t* tests for ERQ, CERQ, PANAS-NA, STAXI-anger in, STAXI-anger control
 - ❖ Wilcoxon signed-rank tests (non-parametric) for PANAS-PA, STAXI-state anger, STAXI-anger out
 - ◆ output: *"D_RQ5a_t test_WSR test_ERQ, CERQ, PANAS, STAXI"*

- **b. Multiple linear regression (ERQ, CERQ, PANAS-PA, PANAS-NA, STAXI-state anger, STAXI-anger in, STAXI-anger out, STAXI-anger control)**
 - ◆ **ERQ:**
 - ❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
 - ❖ syntax: *"D_RQ5b_syntax_multiple regression_ERQ"*
 - ✚ z standardization of predictors
 - ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
 - ❖ output: *"D_RQ5b_output_multiple regression_ERQ"*
 - ◆ **CERQ:**
 - ❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
 - ❖ syntax: *"D_RQ5b_syntax_multiple regression_CERQ"*
 - ✚ z standardization of predictors
 - ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression

❖ output: *"D_RQ5b_output_multiple regression_CERQ"*

◆ **PANAS-PA:**

❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)

❖ syntax: *"D_RQ5b_syntax_multiple regression_PANAS-PA"*

✚ z standardization of predictors

✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)

✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis

✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression

❖ output: *"D_RQ5b_output_multiple regression_PANAS-PA"*

◆ **PANAS-NA:**

❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)

❖ syntax: *"D_RQ5b_syntax_multiple regression_PANAS-NA"*

✚ z standardization of predictors

✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)

✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis

✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression

❖ output: *"D_RQ5b_output_multiple regression_PANAS-NA"*

◆ **STAXI-state anger:**

- ❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
- ❖ syntax: *"D_RQ5b_syntax_multiple regression_STAXI-state anger"*
 - + z standardization of predictors
 - + Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - + Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - + Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ◆ output: *"D_RQ5b_output_multiple regression_STAXI-state anger"*

◆ **STAXI-anger in:**

- ❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
- ❖ syntax: *"D_RQ5b_syntax_multiple regression_STAXI-anger in"*
 - + z standardization of predictors
 - + Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - + Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - + Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ◆ output: *"D_RQ5b_output_multiple regression_STAXI-anger in"*

◆ **STAXI-anger out:**

- ◆ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL_n=41"* (original data set for analysis is kept in the folder)
- ◆ syntax: *"D_RQ5b_syntax_multiple regression_STAXI-anger out"*
 - ✚ z standardization of predictors
 - ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors) → with $n = 41$ participants
- ❖ output: *"D_RQ5b_output_multiple regression_STAXI-anger out_n=41"*
- ❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL_n=40"* (participant TP050TAF is removed from the data set as this participant is an outlier and the prerequisites of multiple linear regression are better fulfilled if this participant is excluded; original data set with $n = 40$ participants is kept in the folder)
- ❖ syntax: same syntax as above; syntax: *"D_RQ5b_syntax_multiple regression_STAXI-anger out"*
 - ✚ z standardization of predictors
 - ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
 - ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
 - ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ❖ output: *"D_RQ5b_output_multiple regression_STAXI-anger out_n=40"*

◆ **STAXI-anger control:**

- ❖ data set: *"D_RQ5b_data set1_Train4Positivity_FINAL"* (original data set for analysis is kept in the folder)
- ❖ syntax: *"D_RQ5b_syntax_multiple regression_STAXI-anger control"*

- ✚ z standardization of predictors
- ✚ Multiple linear regression (including covariates) – parsimonious model (4 predictors)
- ✚ Multiple linear regression (including covariates) – comprehensive model (6 predictors) → conducted, but NOT analyzed for thesis
- ✚ Testing of assumptions (linearity, no outliers or influential data points, no multicollinearity, homoscedasticity, normality of residuals, independence of residuals) is included in syntax for multiple linear regression
- ◆ output: *“D_RQ5b_output_multiple regression_STAXI-anger control”*

- **D. RQ6_Validation of AAT**

- data set: *“D_RQ6_data set1_Train4Positivity_FINAL”*
- syntax: *“D_RQ6_syntax_validation of AAT”*
 - ◆ a. testing of linearity (scatter plots) → regression line and background of scatterplot adapted
 - ◆ b. test of univariate normality for all variables
 - ◆ c. calculation of bivariate Pearson correlations
- output: *“D_RQ6_output_validation of AAT”*

E. Additional data sets including AAT data for calculation of AAT-CS at pre- and posttest (which is incorporated into data set 1)

- **Pretest**

- Data entry into SPSS using Millisecond inquisit data sets (in iqdat format) including the reaction times of participants in the AAT assessment at pretest
- ◆ false reactions in the AAT are immediately deleted during data entry

- ◆ using syntax: *"00_E_syntax_calculation minima, maxima for exclusion 1% RTs"*, the slowest and fastest 1% are identified and immediately deleted → *"00_E_data set7_AAT data_pretest_false reactions deleted_n=41participants"*
- ◆ Exclusion of $n = 2$ participants (TP021SOF, TP003JEF) who correctly performed the AAT during practice trials but not during the assessment trials at pretest → *"01_E_data set7_AAT data_pretest_n=39 participants"*
- ◆ This data set (01) is duplicated and used to calculate the median RTs and the AAT Compatibility Score (AAT-CS) at pretest using the syntax *"01_E_syntax_calculation median RTs and AAT-CS"*
- ◆ → Final data set 7: *"02_E_data set7_AAT data_pretest_FINAL"*
- ◆ AAT-CS at pretest calculated here is pasted into data set 1 (see above)
- **Posttest**
 - Data entry into SPSS using Millisecond inquisit data sets (in iqdat format) including the reaction times of participants in the AAT assessment at posttest
 - ◆ false reactions in the AAT are immediately deleted during data entry
→ *"00_E_data set8_AAT data_posttest_false reactions deleted_n=41participants"*
 - ◆ Consistent with pretest: exclusion of $n = 2$ participants (TP021SOF, TP003JEF) who correctly performed the AAT during practice trials but not during the assessment trials at pretest → *"01_E_data set8_AAT data_posttest_n=39participants"*
 - ◆ This data set (01) is duplicated and used to exclude the slowest and fastest 1% RTs using the syntax *"01_E_syntax_calculation minima, maxima for exclusion 1% RTs"* →
"02_E_data set8_AAT data_posttest_n=39 participants"
 - ◆ This data set (02) is duplicated and used to calculate the median RTs and the AAT Compatibility Score (AAT-CS) at posttest using the syntax *"02_E_syntax_calculation median RTs and AAT-CS"*
 - ◆ → Final data set 8: *"03_E_data set8_AAT data_posttest_FINAL"*

- ◆ AAT-CS at posttest calculated here is pasted into data set 1 (see above)

F. Preparation of data from weekly stressor monitoring (sundays) in TRAIN₄Positivity using

“movisensXS”

- Download of Excel files including stressor monitoring data for single participants from movisens homepage → total data set and separate Excel files for single participants → file renamed as *“00_Train4Positivity_2019_n=41_stressor monitoring_sundays_TOTAL (markings removed)”*
- Excel file is transformed into SPSS file → *“00_data set_Train4Positivity_stressor monitoring_sundays_TOTAL (prepared in SPSS)”*
- Variables are renamed → *“01_data set_Train4Positivity_stressor monitoring_sundays_TOTAL (variables renamed)”*
- Restructuring of data set from long into wide format
 - data set: copy of *“01_data set_Train4Positivity_stressor monitoring_sundays_TOTAL (variables renamed)”*
 - ◆ syntax: *“01_syntax_restructuring data set (into wide format)”*
 - ◆ → data set renamed as *“02_data set_Train4Positivity_stressor monitoring_sundays_TOTAL (restructured data set)”*
 - Data from this data set (02) are integrated into the original data set with pre- and posttest data of TRAIN₄Positivity
 - ◆ data set: copy of *“02_data set_Train4Positivity_stressor monitoring_sundays_TOTAL (restructured data set)”*
 - ◆ → data set renamed as *“03_data set_Train4Positivity_ORIGINAL DATA SET_complete (including weekly EMA stressor monitoring)”*
 - In this data set (03): in case a stressor did not occur (e.g., DH_mid_1.1 = 0), the severity rating of the respective stressor received the value “.” (i.e., is treated as missing) → To be

consistent with the procedures in LifeStress, severity ratings for non-occurred stressors are replaced with 0 (i.e., stressor did not occur and, thus, severity is 0)

- ◆ data set: copy of *"03_data set_Train4Positivity_ORIGINAL DATA SET_complete (including weekly EMA stressor monitoring)"*
- ◆ → data set renamed as *"04_data set_Train4Positivity_ORIGINAL DATA SET_complete (including weekly EMA stressor monitoring, severity 0 if non-occurred stressor)"*
- Calculation of sum scores for sum of microstressors at three Sundays
 - ◆ data set: copy of *"04_data set_Train4Positivity_ORIGINAL DATA SET_complete (including weekly EMA stressor monitoring, severity 0 if non-occurred stressor)"*
 - ◆ syntax: *"04_syntax_Train4Positivity_sum microstressors sundays"*
 - ◆ output: *"04_output_Train4Positivity_sum microstressors sundays"*
 - ◆ data set renamed as *"05_data set_Train4Positivity_ORIGINAL DATA SET_complete (including weekly EMA stressor monitoring)"*
- → This data set (05) is duplicated and renamed as
"Train4Positivity_2019_41participants_ORIGINAL DATA SET_complete (including weekly EMA stressor monitoring)" → which is used as basis for the data preparation of data set 1 (see A. Data preparation)

G. Preparation of combined data set with LifeStress and TRAIN₄Positivity data for RQ1, exploratory

RQ 1.2

- **G1. Using Excel file with original LifeStress data (from Secutrial, provided by Lara Mey), data entry into SPSS ($n = 70$) → *"G1_data set_only LifeStress_n=70participants"***
- **G2. Exclusion of LifeStress participants not meeting the eligibility criteria for TRAIN₄Positivity (e.g., stressor exposure to microstressors ≥ 30 stressors)**
 - data set: *"G2_data set_only LifeStress_n=70participants_GHQ-28&MIMIS total scores"*
 - syntax: *"G2_syntax_exclusion not meeting Train4Positivity criteria"*

- output: *"G2_output_exclusion not meeting Train4Positivity criteria"*
- manual deletion of $n = 24$ participants not meeting the criteria in a copy of data set *"G2_data set_only LifeStress_n=70participants_GHQ-28&MIMIS total scores"*
- → resulting data set is named *"G3_data set_LifeStress_meeting criteria"* that is used for further analyses
- **G3. Integration of stressor monitoring (MIMIS online survey) after 3 weeks using corresponding Excel file (MIMIS_weekly)**
 - Preparation of Excel file:
 - ◆ *"1_MIMIS_weekly_LifeStress_n=70participants"*
 - ◆ *"MIMIS_weekly_LifeStress_n=46participants"*
 - ◆ *"MIMIS_weekly_LifeStress_n=46_only week 3"* → that is, deletion of data of weekly MIMIS assessment at week 1, 2 and 4, that was also conducted in LifeStress
 - Integration of data from Excel file into SPSS data set → *"G3_data set_LifeStress_meeting criteria_MIMIS 3 weeks"*
 - ◆ Recoding of items regarding the frequency (number of days) and perceived severity of microstressors (e.g., "n.a." replaced with "."; severity rating from 1 to 5 is recoded to 0–4 consistent with TRAIN₄Positivity) → resulting data set is named *"G4_data set_LifeStress_meeting criteria_MIMIS 3 weeks_recoded"*
- **G4. Integration of TRAIN₄Positivity data (pretest data + posttest MIMIS data)**
 - using *"G4_data set_LifeStress_meeting criteria_MIMIS 3 weeks_recoded"* from G3
 - integration of TRAIN₄Positivity data (from original data set) → data set is renamed as *"G4_data set_LifeStress & Train4Positivity (pretest data & posttest MIMIS)"*
 - "n.a." concerning days of occurrence of microstressors and severity in copy of this data set is replaced with 0
 - ◆ syntax *"G4_syntax_MIMIS days and severity n.a. replaced with 0"*
 - ◆ output *"G4_output_MIMIS days and severity n.a. replaced with 0"*

→ data set renamed as *"G5_data set_LifeStress & Train4Positivity_MIMIS days and severity replaced with 0"*

- **G5. For combined data set: Missing Values Analysis, EM imputation, and calculation of total scores**

- Combined data set (*"G5a_data set_LifeStress & Train4Positivity_combined"*) is reduced by further variables that are not needed for PS matching
- → data set *"G5b_data set_LifeStress & Train4Positivity_combined_reduced"*
- Missing Values Analysis & Little's MCAR test:
 - ◆ data set: *"G5b_data set_LifeStress & Train4Positivity_combined_reduced"*
 - ◆ syntax: *"G5b_syntax_Missing Values Analysis & Little's MCAR test"*
 - ◆ output: *"G5b_output_Missing Values Analysis & Little's MCAR test"*
- EM imputation:
 - ◆ data set: *"G5b_data set_LifeStress & Train4Positivity_combined_reduced"*
 - ◆ syntax: *"G5c_syntax_EM imputation"*
 - ◆ output: *"G5c_output_EM imputation"*
 - ◆ → newly created data set is named as *"G5c_data set_LifeStress & Train4Positivity_combined_imputed"*
- Re-integration of (sociodemographic) variables that were removed during EM imputation (e.g., education):
 - ◆ data set: copy of *"G5c_data set_LifeStress & Train4Positivity_combined_imputed"*
 - ◆ → renamed as *"G5d_data set_LifeStress & Train4Positivity_combined_imputed_complete"*
- Variable education is modified from string variable to numeric variable:
 - ◆ data set: copy of *"G5d_data set_LifeStress & Train4Positivity_combined_imputed_complete"*

- ◆ → renamed a *"G5e_data set_LifeStress & Train4Positivity_combined_imputed_complete_educationnum"*
 - Recoding of items and calculation of total scores:
 - ◆ data set: copy of *"G5e_data set_LifeStress & Train4Positivity_combined_imputed_complete_educationnum"*
 - ◆ syntax: *"G5f_syntax_recoding & calculation of total scores"*
 - ◆ output: *"G5f_output_recoding & calculation of total scores"*
 - ◆ data set renamed as *"G5f_data set_LifeStress & Train4Positivity_combined_imputed_with total scores"*
 - Variable labels (changed during EM imputation) are slightly modified:
 - ◆ data set: copy of *"G5f_data set_LifeStress & Train4Positivity_combined_imputed_with total scores"*
 - ◆ → renamed as *"G5g_data set_LifeStress & Train4Positivity_combined_imputed_with total scores & correct labels"*
 - ◆ Data set is duplicated and renamed as ***"Data set2_LifeStress & Train4Positivity_FINAL"***
 - saved in .dta format for analyses in Stata → used for analysis of RQ1 – exploratory RQ
- 1.2 (see D)