

# PREACT-digital: Feature Database Documentation

Leona Hammelrath

Tessa Meyer

2025-03-03

## Introduction

In the following sections, you find an overview of our datasets on the study server as well as a detailed description of variables and constructs.

---

### Glossary

- **beep** = Timepoint where participants receive prompt to answer questions about their current mood, situation, etc.
- **measurement burst** = Period (i.e. 2 weeks) where participants receive regular EMA questionnaires
- **epoch** = Highest granularity of passive data. Epoch length depends on data source (i.e. 30 seconds for ECG data)

## Data Structure

### Folder Structure on High Performance Cluster (HPC) [wip]

```
SP6/  
  |- processed/  
    |- passive/  
      |- epoch                                # not aggregated; most finegrained resolution  
      |- activity_epoch
```

```

    |- heart_rate_epoch
    |- ecg_epoch
    |- gps_epoch
  |- daily                                # daily aggregates
    |- activity_daily
    |- heart_rate_daily
    |- ecg_daily
    |- gps_daily
  |- ema
    |- ema_beep                          # not aggregated; most finegrained resolution
    |- ema_daily                         # daily aggregate
    |- ema_burst                         # burst aggregate
    |- ema_meta                          # technical meta data
  |- ecg
    |- ecg_raw                          # raw data (sampling rate: 300 Hz; 9000 data points)
    |- ecg_processed                     # processed, e.g. heart rate variability (hrv)
  |- meta
    |- monitoring                        # study monitoring

```

## EMA Data

This section outlines the EMA data sets ([files](#)) in detail and provides a thorough description of the eight [EMA constructs](#) and a [item-level overview](#).

### Data sets

#### Files:

- `ema_beep.pkl`
- `ema_meta.pkl`

#### Details `ema_content.pkl` file:

Show details

No.	Column name	Description	Data type	Scale level	Variable Level
1	id	Unique identifier wearable and ema data within subproject 6 (SP6)	str		
2	for_id	Unique identifier across all PREACT subprojects and redcap	str		
3	timestamp_item_completed	Timestamp at which a single item was completed	datetime64	interval	
4	timestamp_beep_completed	Timestamp at which a beep was completed	datetime64	interval	
5	timestamp_beep_expired	Timestamp at which the processing of the beep has expired (a beep expires after 30 min)	datetime64	interval	
6	measurment_burst	Measurement burst describes the measurement point in the longitudinal study [Baseline (T0), after 20 therapy sessions (T20), or after therapy end respectively 365 days after therapy start (TPost)]	int	ordinal	0 = T0 1 = T20 2 = TPost

No.	Column name	Description	Data type	Scale level	Variable Level
7	<code>schedule_chronotype</code>	Depending on their individual sleep-wake rhythm participants can choose to receive beeps between 07:30 and 21:30 (lark) or 09:30 and 22:30 (owl)	<code>int</code>	nominal	24 = T0 lark 25 = T0 owl 33 = T20 lark 34 = T20 owl 38 = TPost lark 39 = TPost owl
8	<code>response</code>	Chosen response by participant	<code>int</code>	ordinal, nominal, binary	
9	<code>item</code>	Question/item title	<code>str</code>		
10	<code>beep_per_person_id</code>	Unique beep identifier. Date and number of beep per ID	<code>str</code>		
11	<code>date</code>	Date on which the question/item was generated	<code>datetime64</code>	interval	
12	<code>study_version</code>	Study version (short version: includes Baseline (T0), long version: includes Baseline (T0), T20 and TPost)	<code>int</code>	nominal	1= long 2 = short
13	<code>ema_burst_start</code>	Absolute start EMA measurement burst (i.e. defined start according to study protocol)	<code>datetime64</code>	interval	

No.	Column name	Description	Data type	Scale level	Variable Level
14	<b>ema_burst_end</b>	Absolute end EMA measurement burst (i.e. defined end according to study protocol)	<b>datetime64</b>	interval	
15	<b>season</b>	Describes the four seasons	<b>int</b>	nominal	1 = Spring 2 = Summer 3 = Fall 4 = Winter
16	<b>time_of_day</b>	Time of day stratified into five categories (Early Morning = 00:00 - 00:00, Morning = 00:00 - 00:00, Afternoon = 00:00 - 00:00, Evening = 00:00 - 00:00, Night = 00:00 - 00:00)	<b>int</b>	nominal	1 = Early Morning 2 = Morning 3 = Afternoon 4 = Evening 5 = Night
17	<b>weekend</b>	Does the timestamp in the time series describes a day at the weekend?	<b>int</b>	nominal	0 = No 1 = Yes
18	<b>nr_beep_daily</b>	Number of questionnaire/beep within a day	<b>int</b>	ordinal	1 - 8
19	<b>n_beeps_completed</b>	Number of questionnaires/beeps completed by a person within a day	<b>int</b>	ordinal	1 - 9

No.	Column name	Description	Data type	Scale level	Variable Level
20	ema_relat_burst_start	Relative start EMA measurement burst (i.e. actual start)	datetime64	interval	
21	ema_relat_burst_end	Relative end EMA measurement burst (i.e. actual end)	datetime64	interval	
22	absolute_day_index	Day since expected (absolute) start	int	ratio	1 - 16
23	relative_day_index	Day since actual (relative) start	int	ratio	1 - 16

#### Details ema\_meta.pkl file:

Show details

No.	Column name	Description	Data type	Scale level	Variable Level
1	id	Unique identifier wearable and ema data within subproject 6 (SP6)	str		
2	for_id	Unique identifier across all PREACT subprojects and redcap	str		
3	response_text	Response displayed on device	str		
4	item_code_map	Numerical item code mapping	int	[insert]	nominal
5	beep_type		int		nominal
6	beep_type_name	Name of the questionnaire	str		

No.	Column name	Description	Data type	Scale level	Variable Level
7	item_order	Order in which the items are displayed	int	0 - 8	
8	beep_num_run	How many times a beep was opened before completion. Unique per answer. One beep can have multiple runs until completion	int		

### Methods: Hierarchical Data Structure

#### 1. Level 1: Measurements (Observations)

- Each person records data 8x/day over 14 days
- This results in 112 measurements per wave (8x14)

#### 2. Level 2: Days

- Measurements (Level 1) are nested within days (Level 2)
- Each wave has 14 days

#### 3. Level 3: Waves (Measurement points)

- Each person goes through three waves (long version)
- Days (Level 2) are nested within waves (Level 3)

#### 4. Level 4: Individuals (Participants)

- Waves (Level 3) are nested within participants (Level 4)

---

### EMA constructs and item-level overview

The EMA measurement includes the following constructs:

1. [Affect](#)
2. [Emotion regulation](#)

3. [Situational context](#)
4. [Significant events](#)
5. [Social context](#)
6. [Therapeutic agency](#)
7. [Physical fitness](#)
8. [ECG control](#)

## Affect

- Description: At each beep, participants were asked about their current affective state
- Construct: PANAS-X subscales [Haney et al. \(2023\)](#)
- 17 Items

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
	How ... do you feel right now?			
<b>anxious</b>	anxious	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>nervous</b>	nervous	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>attentive</b>	attentive	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>relaxed</b>	relaxed	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>calm</b>	calm	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>irritable</b>	irritable	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>angry</b>	angry	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>fatigue</b>	fatigue	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>cheerful</b>	cheerful	1-2-3-4-5-6- 7	not at all - very much	all beeps
<b>happy</b>	happy	1-2-3-4-5-6- 7	not at all - very much	all beeps



Variable	Item	Scale	Scale Endpoints	Measurement Time
ashamed	ashamed	1-2-3-4-5-6-7	not at all - very much	all beeps
dissatisfied	dissatisfied with myself	1-2-3-4-5-6-7	not at all - very much	all beeps
self_confident	self-confident	1-2-3-4-5-6-7	not at all - very much	all beeps
shy	shy	1-2-3-4-5-6-7	not at all - very much	all beeps
downcast	downcast	1-2-3-4-5-6-7	not at all - very much	all beeps
sad	sad	1-2-3-4-5-6-7	not at all - very much	all beeps
lonely	lonely	1-2-3-4-5-6-7	not at all - very much	all beeps

### Emotion regulation

- Description: At each beep, participants were asked to rate the intensity and controllability of their most negative thought since the last beep. Then, we assessed the use of different ER strategies since the last beep
- Construct: RESS-EMA scale [Medland et al. \(2020\)](#)
- 6 Items (covering reappraisal, rumination, suppression, distraction, relaxation, acceptance)

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
	Think about the strongest negative feeling since the last beep [since waking up].			

Variable	Item	Scale	Scale Endpoints	Measurement Time
er_intensity	How intense was this feeling?	1-2-3-4-5-6-7 (1 = neutral)	not at all - very much	all beeps (except the first of the day)
er_intensity	How intense was this feeling?	1-2-3-4-5-6-7 (1 = neutral)	not at all - very much	first beep of the day
er_control	How controllable was the situation that triggered this feeling?	1-2-3-4-5-6-7 (4 = neutral)	not at all - very much	all beeps (except the first of the day)
er_control	How controllable was the situation that triggered this feeling? As a reaction to the negative feeling ...	1-2-3-4-5-6-7 (4 = neutral)	not at all - very much	first beep of the day
er_relaxation	I tried to breathe deeply	1-2-3-4-5-6-7	not at all - very much	all beeps
er_rumination	I kept thinking about what was bothering me	1-2-3-4-5-6-7	not at all - very much	all beeps

Variable	Item	Scale	Scale Endpoints	Measurement Time
er_reappraisal	considered the situation from different perspectives	1-2-3-4-5-6-7	not at all - very much	all beeps
er_distraction	tried to distract myself	1-2-3-4-5-6-7	not at all - very much	all beeps
er_suppression	tried to hide my feelings	1-2-3-4-5-6-7	not at all - very much	all beeps
er_acceptance	tried to accept the situation	1-2-3-4-5-6-7	not at all - very much	all beeps

### Situational Context

- Description: At each beep, participants were asked to specify activities they had pursued in the preceding 2 hours from a given set of 9 common activities. Participants were able to select multiple options simultaneously. Subsequently, they were asked to evaluate how much they enjoyed the respective activities
- Construct: Self-constructed, based on the DIAMONDS scale [Rauthmann & Sherman \(2016\)](#) and the WARN-D study protocol [Fried et al. \(2022\)](#), a similar longitudinal digital phenotyping study. We aimed to find a balance between sparsity of items and high degree of situational coverage.
- 2 Items

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
	How did you spent the time since the last beep [since waking up]?			
	(Multiple answers possible)			

Variable	Item	Scale	Scale Endpoints	Measurement Time
situation_1	<input type="checkbox"/> Work or study <input type="checkbox"/> House-work or errands <input type="checkbox"/> Caring for children/relatives <input type="checkbox"/> Eating/drinking/personal hygiene <input type="checkbox"/> On the move (e.g., in the subway) <input type="checkbox"/> Smartphone/social media <input type="checkbox"/> Leisure activity, rather passive (e.g., watching a movie, reading) <input type="checkbox"/> Leisure activity, rather active (e.g., sports, outings) <input type="checkbox"/> Something else			all beeps (except the first of the day)
situation_1_morning	morning			first beep of the day

Variable	Item	Scale	Scale Endpoints	Measurement Time
situation_2	How much did you enjoy this activity?	-2, -1, 0, 1, 2	not at all - very much	all beeps (except the first of the day)
situation_2_morale	How much do you enjoy this activity?	-2, -1, 0, 1, 2	not at all - very much	first beep of the day

### Significant Events

- Description: Participants were asked to think about the most important event since the last beep and how pleasant they perceived it
- Construct: Self-constructed
- 1 Items

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
event_general	I think of the most significant moment (situation/experience) since the last survey. How did you perceive it?	-2, -1, 0, 1, 2	very unpleasant - very pleasant	all beeps (except the first of the day)

Variable	Item	Scale	Scale Endpoints	Measurement Time
event_general	Thinking the most significant moment (situation/experience) since waking up. How did you perceive it?	-2, -1, 0, 1, 2	very unpleasant - very pleasant	first beep of the day

### Social context

- Description: Participants were asked if they had social contacts since the last beep, how (online/ in person/ phone) and how agreeable the contact was.
- Self-constructed
- 3 Items

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
event_social	Have you had social contacts since the last survey?	binary: yes/no		all beeps (except the first of the day)
event_social	Have you had social contacts since waking up?	binary: yes/no		first beep of the day

Variable	Item	Scale	Scale Endpoints	Measurement Time
event_social_1	How did the social contact take place?	multiple choice: [ ] online [ ] by phone [ ] in person		all beeps
event_social_2	How did you experience the social contacts?	-2, -1, 0, 1, 2	very unpleasant - very pleasant	all beeps

### Therapeutic Agency (TA)

- Description: Participants were asked about Therapeutic Agency (TA) in everyday life
- Construct: Self-constructed based on the Therapeutic Agency Inventory (TAI) [Huber et al. \(2019\)](#). The original TAI contains 3 subscales, covering in-session activities, passivity towards the therapist and out-of-session activities. As we were interested in assessing therapeutic agency in everyday life, our TAI-EMA items are based on the “out-of-session activities” subscales and cover cognitive and behavioral aspects of TA
- 4 Items

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
ta_behavioral_1	Prompted by my therapy today, I have ... / Today I have ...	1-2-3-4-5-6-7	not at all - very much	1x/day, 8th beep



Variable	Item	Scale	Scale Endpoints	Measurement Time
ta_behavioral_2	Tried to think differently about things	1-2-3-4-5-6-7	not at all - very much	1x/day, 8th beep
ta_cognitive_1	thought about something that was discussed in therapy	1-2-3-4-5-6-7	not at all - very much	1x/day, 8th beep
ta_cognitive_2	done something to improve my situation	1-2-3-4-5-6-7	not at all - very much	1x/day, 8th beep

### Physical Fitness

- Description: Participants were asked how physically healthy they had felt today on the last beep of the day
- Construct: Self-constructed
- 1 Item

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
physical_health_1	How physically healthy did you feel today?	-2, -1, 0, 1, 2	worse than usual / normal / better than usual	1x/day, 8th beep

## ECG Control

- Description: During measurement bursts, patients were asked twice per day to conduct a resting-state ECG on their Scanwatch. To control for potential confounders influencing the signal, we asked if they had consumed nicotine, caffeine or alcohol or had a heavy meal in the last 30 minutes
- Construct: Self-constructed
- 1 Item

Show Items

Variable	Item	Scale	Scale Endpoints	Measurement Time
ecg_control	Within the last 30 minutes, did you ... - drink coffee or alcohol? - smoke? - eat a heavy meal?	binary: yes/no		2x/day, 1th and 5th beep

## Passive Sensor Data

This section outlines the passive sensor data set ([files](#)) in detail and provides a thorough description of the different wearable modalities (heartrate, activity, sleep, GPS).

### Data sets

#### Files:

- `passive_data.feather`

Details `passive_data.feather` file:

Show details

No.	Column name	Description	Data type	Scale level	Variable Level
1	id	Unique identifier wearable and ema data within subproject 6 (SP6)	str		
2	for_id	Unique identifier across all PREACT subprojects and redcap	str		
3	modality	Type of modality	str	categorical	
4	timestamp_start	Timestamp at which the specific modality recording starts	datetime64	interval	
5	timestamp_end	Timestamp at which the specific modality recording ends	datetime64	interval	
6	time_interval	Duration recording	str		
7	float_value	Variable level of the modality	float		
8	boolean_value	Variable level of the modality	boolean		
9	start_date	Start date of recording	datetime64		
10	start_hour	Start hour of recording	datetime64		
11	study_version	Study version (short version: includes Baseline (T0), long version: includes Baseline (T0), T20 and TPost)	int	nominal	1= long 2 = short

## Heartrate

Show details

No.	Modality	Device	Data type	Sampling Rate	Scale level	Features
1	heartrate_PPG	Withings Scanwatch				
2	rmssd	Withings Scanwatch				

## Activity

Show details

No.	Modality	Device	Data type	Scale level	Features
1	Steps	Withings Scanwatch			
2	ActivityType	Withings Scanwatch			
3	ActivityBinary	Withings Scanwatch			
4	RunBinary	Withings Scanwatch			
5	BikeBinary	Withings Scanwatch			
6	WalkBinary	Withings Scanwatch			
7	FloorsClimed	Withings Scanwatch			
8	ElevationGain	Withings Scanwatch			
9	ElevationGain	Withings Scanwatch			
10	ActiveBurnedCalories	Withings Scanwatch			
11	ActiveTypeDetail1	Withings Scanwatch			
12	ActiveTypeDetail2	Withings Scanwatch			

## Steps: Daily Aggregates

Daily Features inspired by [insert RADAR study reference]

File name: `steps_daily`

Show details

No.	Column Name	Description
1	<code>id</code>	Unique identifier wearable and ema data within subproject 6 (SP6)
2	<code>for_id</code>	Unique identifier across all PREACT subprojects and redcap
3	<code>date</code>	Day timestamp (floor to day) UTC

No.	Column Name	Description
4	n_steps_day	Total number of walked steps within the day
5	spm_25_steps	25th percentile of daily steps per minute distribution
6	spm_50_steps	50th percentile of daily steps per minute distribution
7	spm_75_steps	75th percentile of daily steps per minute distribution
8	spm_max_steps	Maximum steps per minute along all day
9	spm_count_steps	Number of minutes with step data available
10	spm_mean_steps	Mean steps per minute along all day (among available records)
11	spm_std_steps	Standard deviation of steps per minute along all day
12	spm_skew_steps	Skewness of steps per minute along all day
13	spm_kurtosis_steps	Kurtosis of steps per minute along all day
14	night_sum_steps	Sum of steps per minute during nighttime (00:00-05:59)
15	night_mean_steps	Mean steps per minute during nighttime (00:00-05:59)
16	n_hour_steps	Mean of hourly step sums (sum of steps per minute, averaged by hour)
17	spm_max_avghr_steps	Maximum steps per minute, averaged by hour
18	spm_mean_avghr_steps	Mean steps per minute, averaged by hour
19	spm_std_avghr_steps	Standard deviation of steps per minute, averaged by hour
20	spm_skew_avghr_steps	Skewness of steps per minute, averaged by hour
21	spm_kurtosis_avghr_steps	Kurtosis of steps per minute, averaged by hour
22	n_steps_activehr_steps	Maximum of the hourly sum of steps along all day
23	timestamp_max_activehr_steps	Most active hour (hour with maximum hourly sum of steps)

No.	Column Name	Description
24	max_spm_activehr_steps	Maximum step cadence during the most active hour
25	mean_spm_activehr_steps	Average step cadence during the most active hour
26	dailysteps_25perc_steps	Hour at which 25th percentile of daily steps occurred (cumulative)
27	dailysteps_50perc_steps	Hour at which 50th percentile of daily steps occurred (cumulative)
28	dailysteps_75perc_steps	Hour at which 75th percentile of daily steps occurred

**Sleep**

**GPS**

**ECG Data**

**Data sets**