

## Data description for Case II

The dataset *EmoPairCompete* [1] comprises of physiological measurements from the Empatica E4 wristband [2, 3, 4]. Few biosignal observations are shown in Fig. 1. The dataset was collected over multiple acquisition rounds ( $D_{11}$ ,  $D_{12}$ ,  $D_{13}$ ) with identical experimental design, conducted at different times of the year. **For this case  $D_{13}$  has been divided into  $D_{13}$ ,  $D_{14}$ ,  $D_{15}$ ,  $D_{16}$ .** The dataset is described below.

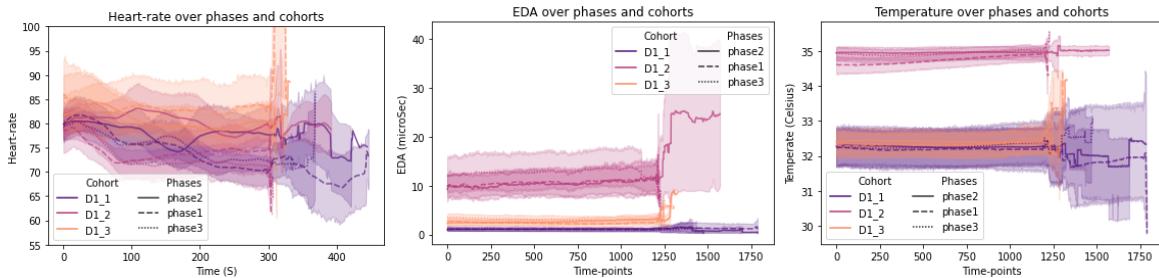


Figure 1: Illustration of some biosignals

### Experimental design

All participants were given an E4 biosensor and instructed on how to wear it, turn it on and off and mark events. Biosignals were collected under a resting period, an emotion-eliciting condition (puzzling-phase), and a recovery period as depicted in Figure 2. After each phase, participants completed a questionnaire. This procedure was repeated four times successively (rounds). We instructed participants to turn off the biosensor after each recovery period. Each time the biosensor was turned on, it was allowed to calibrate for one minute. For the resting and recovery periods, we asked participants to find a comfortable position, sitting still and quietly and rest to the best of their ability. To elicit emotions, we used a game-elicited-emotion paradigm. The tangram task was originally designed to study parent and social behaviour between parents and children [5]. The phases are:

1. Pre-/Post-puzzle phase: Before and after the stress-inducing task participants will be asked to find a comfortable sitting position and sit quietly for 5 minutes.
2. Stress/frustration condition: Participants are divided in multiple teams of 2 and compete in a puzzle competition.

The puzzle phase is designed as a competition, wherein a team of two compete against other teams. The competition consists of 4 rounds and is in total 72 minutes. The task of the competition is to solve as many Tangram puzzles as possible in total during the 4 rounds. The task is designed to be difficult. A solved puzzle yields 1 point and the team with the most points at the end of the competition wins. Hence, a round consists of 1+3 phases: wristband calibration (1 minute), pre-task resting (5 minutes), task solving (5 minutes), post-task resting (5 minutes). 60 seconds at the end of the pre-task and task phases are reserved to answer the emotion evaluation questionnaire.

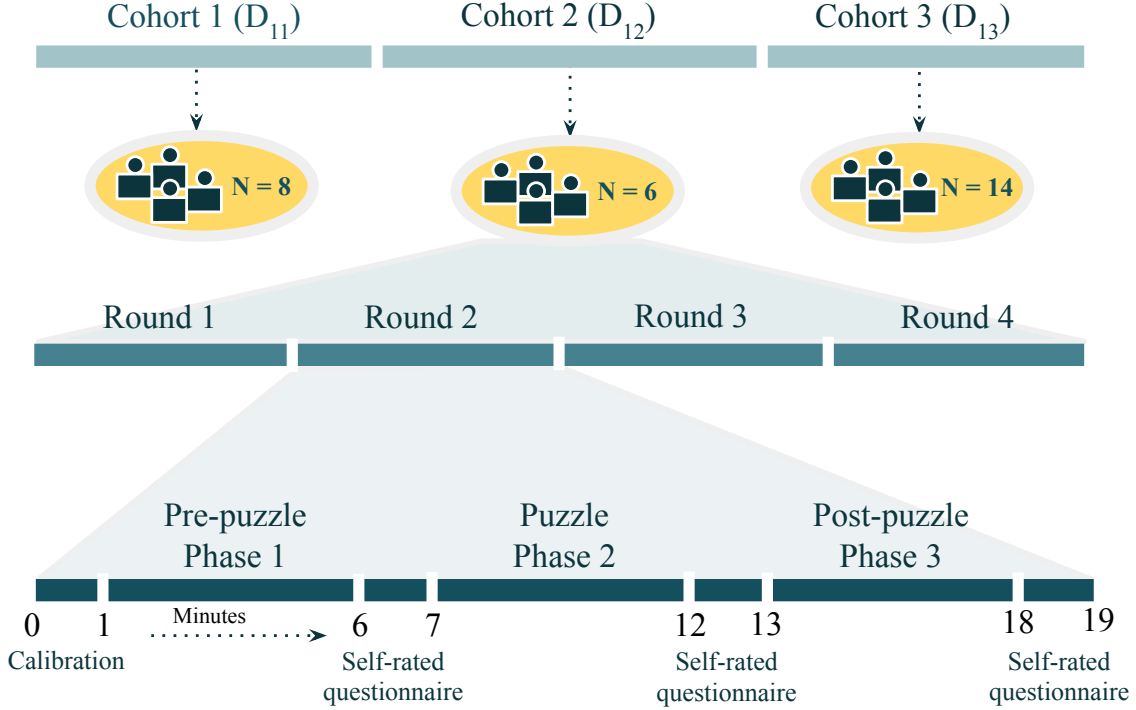


Figure 2: Illustration of the dataset structure

A team is handed a 7-piece tangram puzzle and a set of sketches of various puzzles. The participants of the teams have one role each: an instructor and a puzzle-solver. The puzzle-solver is the only one who is allowed to touch and assemble the puzzle pieces. The instructor is the only one who is allowed to look at the puzzle solutions. A round consisted of seven phases, calibration, pre-puzzle and post-puzzle rest, puzzling, and recording emotions (self-rated questionnaire), organised as shown in Fig. 2.

## Participation

Participants (total  $N = 28$ ) were recruited from our interdisciplinary research group and students and employees within a university department. The ages of the participating women and men ranged from approximately 20 to 42 years. The experiment was run on three separate occasions.  $D_{11}$  were completed in the winter and  $D_{12}$  and  $D_{13}$  were completed in the fall.  $D_{13}$  were conducted in four separate sessions;  $D_{13_1}$  and  $D_{13_3}$  were completed in the morning and  $D_{13_2}$  and  $D_{13_4}$  were collected in the evening. Data from two participants were removed due to device malfunction. Therefore, there is data from  $N = 26$  individuals in the dataset.

## E4 biosensor

The dataset comprises of the electrodermal activity (EDA), heart-rate (HR) and temperature (TEMP) and the blood-volume pulse (BVP) data from the E4 biosensors. The participants were instructed to turn on the E4 at the beginning of each round, tag the start and stop times of the rest and puzzle phases within each round, and to turn off the E4 at the end of each round. Thus, each participant has  $4 \times 3$  recordings corresponding to the four rounds and three phases. The data from each biosensor is provided under the phase directory for each individual as *biosignal.csv*.

## Response questionnaires:

After each phase (pre-puzzle, puzzle, post-puzzle), a self-rated questionnaire is filled by the participants to record their perception of the induced emotions. The international PANAS Short Form (I-PANAS-SF) questionnaire was used for this purpose. It includes five positive emotions (active, alert, attentive, determined, inspired) and five negative emotions (afraid, ashamed, hostile, nervous, upset) that participants rate how much they were experiencing each emotion on a five-point scale from 1 (not at all) to 5 (a lot) [6]. We also assessed frustration level using a visual analogue scale ranging from 0 (not at all frustrated) to 10 (extremely frustrated) after each condition. Finally, participants rated how difficult they found the task on a visual analogue scale ranging from 0 (not at all difficult) to 10 (extremely difficult) after the stress condition. The questions are reiterated below. The self-rated responses are provided under each phase directory for each participant as *response.csv*

1. On a scale from 0-10, where 0 is not frustrated at all and 10 is extremely frustrated, how frustrated are you feeling right now?
2. On a scale from 0-5, where 0 is not upset at all and 5 is extremely upset, how upset are you feeling right now?
3. On a scale from 0-5, where 0 is not hostile at all and 5 is extremely hostile, how hostile are you feeling right now?
4. On a scale from 0-5, where 0 is not alert at all and 5 is extremely alert, how alert are you feeling right now?
5. On a scale from 0-5, where 0 is not ashamed at all and 5 is extremely ashamed, how ashamed are you feeling right now?
6. On a scale from 0-5, where 0 is not inspired at all and 5 is extremely inspired, how inspired are you feeling right now?
7. On a scale from 0-5, where 0 is not nervous at all and 5 is extremely nervous, how nervous are you feeling right now?
8. On a scale from 0-5, where 0 is not determined at all and 5 is extremely determined, how determined are you feeling right now?
9. On a scale from 0-5, where 0 is not attentive at all and 5 is extremely attentive, how attentive are you feeling right now?
10. On a scale from 0-5, where 0 is not afraid at all and 5 is extremely afraid, how afraid are you feeling right now?
11. On a scale from 0-5, where 0 is not active at all and 5 is extremely active, how active are you feeling right now?
12. (Only asked in the task phase) On a scale from 0-10, where 0 is not difficult at all and 10 is extremely difficult, how difficult did you find the task?

## Response questionnaires:

The raw signals have been pre-processed and a set of features extracted. The features consist of:

1. HR: Features extracted from a pre-processed heart rate signal
2. TD: Time domain signal (the raw data in the time domain)

3. Mean: The sample mean of the signal
4. Max: The maximum value of the signal
5. std: The sample standard deviation of the signal
6. Min: The minimum value of the signal
7. Kurtosis: The kurtosis of the signal (fourth moment of the signal)
8. Skew: The skewness of the signal (third moment of the signal)
9. Slope: the slope of a linear fit to the signal (this indicates the general linear trend of the signal)
10. AUC: Area under the curve
11. TEMP: Temperature
12. EDA: Electro-dermal activity ( $EDA = EDA\_P \cup EDA\_T$ )
13. EDA\_P: EDA phasic (Component of the EDA signal)
14. EDA\_T: EDA tonic (Component of the EDA signal)
15. EDA\_TD\_P.Peaks : Peaks from EDA signal
16. EDA\_TD\_P.RT : Rise time of EDA signal
17. EDA\_TD\_P.ReT : Recovery time of EDA signal

Note that the raw signals contain more information, but you are only provided with a subset of extracted features for the case.

## References

- [1] S. Das, N. L. Lund, C. R. González, N. N. Lønfeldt, and L. H. Clemmensen, “EmoPairCompete - physiological signals dataset for emotion and frustration assessment under team and competitive behaviors,” in *ICLR 2024 Workshop on Learning from Time Series For Health*, 2024.
- [2] Empatica, “E4 empatica wristband.” <https://e4.empatica.com/e4-wristband>. [Online; accessed 29-03-2023].
- [3] N. Milstein and I. Gordon, “Validating measures of electrodermal activity and heart rate variability derived from the empatica e4 utilized in research settings that involve interactive dyadic states,” *Frontiers in Behavioral Neuroscience*, vol. 14, p. 148, 2020.
- [4] Empatica, “E4 empatica wristband information on sensors.” <https://support.empatica.com/hc/en-us/categories/200023126-E4-wristband>. [Online; accessed 29-03-2023].
- [5] J. L. Hudson and R. M. Rapee, “Parent–child interactions and anxiety disorders: An observational study,” *Behaviour research and therapy*, vol. 39, no. 12, pp. 1411–1427, 2001.
- [6] E. R. Thompson, “Development and validation of an internationally reliable short-form of the positive and negative affect schedule (panas),” *Journal of cross-cultural psychology*, vol. 38, no. 2, pp. 227–242, 2007.