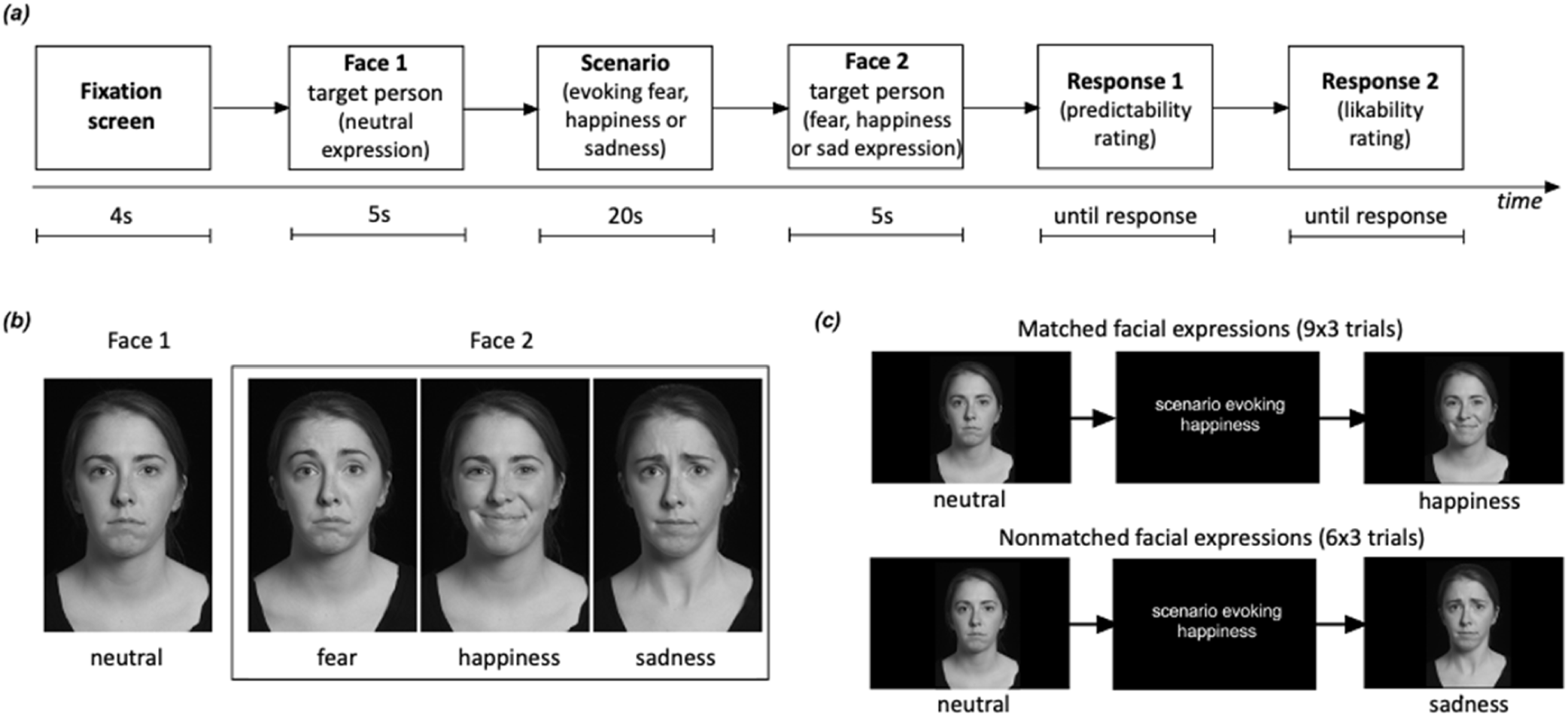
This is a read me file for folder **Datasets (CASLab\_AI)**

This folder contains the datasets that gather different projects. In all projects the aim was to assess the relationship between predictive coding (assessed through a social perception task) and psychological wellbeing.

All participants performed a computer task in which predictive coding and social perception was assessed (Figure 1).

**Figure 1:** Scheme of social perception task (taken from Ramos-Grille et al., 2022).

In addition, all participants answered a series of psychological scales. The exact scales filled depend on the exact project.

We have provided 2 excel files per project:

1. [*Project name*]\_task.xlsx: Computer task data, which contains the information of the 45 experimental trials performed. [This Excel file has 1 sheet in PREDWELL\_RETOS and BEBRASK projects, and 2 sheets in PREDWELL\_PSYCH\_ASD\_POSTP and PREDWELL\_MD\_PD\_BPD projects (one for patients and one for controls). PREDWELL\_RETOS and BEBRASK contain data of healthy participants.]->Second sheet to be added soon.
2. [*Project name*]\_scales.xlsx: Scales, gender and age, which contains the scores of the scales filled by the participants of that project plus gender and age. [This file also has 1 tab in PREDWELL\_RETOS and BEBRASK projects, and 2 tabs in PREDWELL\_PSYCH\_ASD\_POSTP and PREDWELL\_MD\_PD\_BPD projects, following the same structure as the previous Excel file.]-> Second sheet to be added soon.

The following is a brief description of each Excel file:

**Excel 1: Computer task ([Project name]\_task.xlsx)**

Each participant has 2 IDs:

* **Subject:** a code used for the computer task that we are using in the experiment. This code is repeated in some cases.
* **DataFile.Basename**: ID for each participant. It is different for every participant. This variable will appear in the other Excel file as ***EPRIME.CODE***. **We recommend to use this one to identify each participant**.

Response variables:

* **Rating0.RESP**: predictability rating (from 1 to 4).
* **Rating.RESP**: likability rating (from 1 to 4)
* **Rating.RT**: reaction time in Rating.RESP (ikability rating)
* **Rating0.RT**: reaction time in Rating0.RESP (predictabity rating)

Condition variables:

* **EvokedEmotion**: it indicates the emotion category evoked by the scenario presented during each trial. This variable has 3 values: “happiness”, “sadness” and “fear”.
* **Expression**: it indicates the emotion category of the facial expression showed after the scenario during each trial. It has 3 values: “Happy”, “Sad” and “Fear”.
* **Fulfilled**: it has 2 values. 0 for trials in which EvokedEmotion is different from Expression (nonmatched trial); and 1 for trials in which EvokedEmotion and Expression are the same (matched trial).
* **Gender** (of the identity shown in the screen): 1 is for a woman, 0 for a man.
* **Model**: 2 photographs of a person (a model) are shown in each trial. This variable indicates which model is shown in the trial. M for male models and F for female models.
* **ScenarioPick**: this variable indicates which scenario was presented in each trial.
* **TrialCount**: this variable indicates the number of trials. TrialCount has values from 1 to 45 (corresponding to each experimental trial), the rest of the column has NULL values. It indicates the number of trial. NULL values are for practice trials or screen with text.

Each excel file has the same structure, because they are different versions from the same task.

PREDWELL\_RETOS\_task.xlsx: corresponds to PREDWELL RETOS experiment, in which responses (ratings) were limited to 3 seconds. This experiment was run in the lab at UAB and participants were young adults (from 18 to 36 years old). This project has three different versions of the same task (ExperimentVersion could be 1.0.0.824, 1.0.0.826 or 1.0.0.828).

PREDWELL\_PSYCH\_ASD\_POSTP\_task.xlsx and PREDWELL\_MD\_PD\_BPD\_task.xlsx correspond to patients and healthy participants that served as controls for patients of different conditions. This experiment was run (by a researcher) in the clinic for the patients and either in the lab at UAB or at participants’ home for controls. Responses had no time limit.

BEBRASK\_task.xlsx corresponds to 55 healthy participants who did the same task as PREDWELL\_RETOS tab. This experiment was also run in the lab at UAB and participants were young adults (from 18 to 31 years old).

**Excel 2: Scales ([Project name]\_scales.xlsx)**

These Excel files contain scores of scales, age and gender. 2 columns as ID:

* **SUBJECT.CODE** is the main ID for each participant.
* **EPRIME.CODE** in RETOS\_controls\_task.xlsx is DataFile.BaseName. We recommend to use this one to identify each participant.
* Next 2 columns correspond to **Gender** and **Age** of each participant.
* The rest of the columns of this dataset are the scales filled by participants. The table below shows how many participants are in each project and which scales did they fill. Additionally, missing data for each project was added at the end of the document.

|  |  |  |
| --- | --- | --- |
| **Project** | **N** | **Scales** |
| PREDWELL\_RETOS | 95 | PANAS, ERQ, UPPSP, BISBAS, TEPS, SHS, FS, LOT-R, RRQ, ASI-3, ZKPQ, SPQ, MSS-B |
| PREDWELL\_ASD | 38 | PANAS, ERQ, UPPSP, SPQ, MSSB, PHQ9, BSL23, AQS, IUS, BEH |
| PREDWELL\_BPD | 21 | PANAS, ERQ, UPPSP, PHQ9, BSL23 |
| PREDWELL\_EMAR | 4 | PANAS, ERQ, UPPSP, SPQ, MSSB, AQS, IUS, BEH |
| PREDWELL\_MD | 15 | PANAS, ERQ, UPPSP, PHQ9, BSL23 |
| PREDWELL\_PD | 13 | PANAS, ERQ, UPPSP, PHQ9, BSL23 |
| PREDWELL\_PSYCH | 12 | PANAS, ERQ, UPPSP, SPQ, MSSB, AQS, IUS, BEH |
| BEBRASK | 55 | PANAS, ERQ, UPPSP, BISBAS, TEPS, SHS, FS, LOT-R, RRQ, ASI-3, ZKPQ, SPQ, MSS-B, IUS, BEH, DASS-21 |

All values that could not be calculated because the participant did not answer the scale (did not answer it correctly or experimenter mistake) have been indicated as "missing". All missing values of scales that were not included in the protocol and therefore were not filled in by the participants are indicated as "NA".

The scales measure the following:

PANAS – Positive and Negative Affect Scale. It contains 2 scores, 1 for negative affect (NA.) and 1 for positive affect (PA)

ERQ – Emotional Regulation Questionnaire. It contains 2 scores, 1 for cognitive reappraisal (ERQ\_CR) and 1 for emotional suppression (ERQ\_ES).

UPPSP – Impulsive Behavior Scale. It contains 5 scores, 1 for negative urgency (UPPSP\_NU), 1 for positive urgency (UPPSP\_PU), 1 for sensation seeking (UPPSP\_SS), 1 for lack of premeditation (UPPSP\_PMD) and 1 for lack of perseverance (UPPSP\_PSV).

BIS/BAS - Behavioral Inhibition/Activation system Scale. It contains 4 scores, 1 for behavioral inhibition system (BIS), 1 for motivation to follow goals (BAS\_D), 1 for reward responsiveness (BAS\_RR) and 1 for fun seeking (BAS\_FS).

TEPS – Temporal Experience of Pleasure Scale. It contains 2 scores, 1 for anticipatory factor (TEPS\_AF), and 1 for consummatory factor (TEPS\_CF).

SHS – Subjective Happiness Scale. It contains 1 score (SHS).

FS – Flourishing Scale. It contains 1 score (FS).

LOT-R – Life Orientation Test. It contains 1 score (LOT\_R).

RRQ – Rumination-Reflection Scale. It contains 2 scores, 1 for rumination (RRQ\_Rum) and 1 for reflection (RRQ\_Ref).

ASI-3 – Anxiety Sensitivity Index. It contains 3 scores, 1 for physical anxiety (ASI\_P), 1 for cognitive anxiety (ASI\_C) and 1 for social anxiety (ASI\_S).

ZKPQ – Zuckerman-Kuhlman Personality Questionnaire. It contains 11 scores, 1 for neuroticism and anxiety personality traits (ZKPQ\_N\_ANX), 1 for need for general activity (ZKPQ\_GENACT), 1 for need for work activity (ZKPQ\_WORKACT), 1 for liking parties (ZKPQ\_PARTIES), 1 for intolerance of social isolation (ZKPQ\_SY), 1 for sensation seeking (ZKPQ\_SS), 1 for impulsivity (ZKPQ\_IMP), 1 for readiness to express verbal aggression (ZKPQ\_AGG\_HOST) and 1 as a control variable (ZKPQ\_INFREQ). ZKPQ\_INFREQ variable is used to eliminate those subjects whose responses may have been influenced by extreme social desirability (it is not a psychological scale). A score greater than 3 indicates that the validity of their responses is questionable.

SPQ – Schizotypal Personality Questionnaire. It contains 2 scores, 1 for suspiciousness (SPQ) and 1 for (SPQ\_IR).

MSS-B – Multidimmensional Schizotypy Scale Brief. It contains 3 scores, 1 for positive schizotypy (MSSB\_POS), 1 for negative schizotypy (MSSB\_NEG) and 1 for disorganized schizotypy (MSSB\_DES).

PHQ-9 – Patient Helth Questionnaire. It contains 1 score (PHQ9).

BSL23 – Borderline Symptoms List. It contains 1 score (BSL23).

AQS – Autism Spectrum Quotient – Short. It contains 1 score (AQS).

IUS – Intolerance Uncertainty Scale. It contains 1 score (IUS).

BEH – Brief Hypervigilance Scale. It contains 1 score (BEH).

DASS-21 – Depression Anxiety Stress Scale (21 items). It contains 3 scores, 1 for depressive symptoms (DASS\_D), 1 for anxiety symptoms (DASS\_A), and 1 for stress symptoms (DASS\_S).

**Missing data (should be there but is missing):**

* Project PREDWELL\_RETOS:
  + PREDWELL\_RETOS-1-1, PREDWELL\_RETOS-2-1, PREDWELL\_RETOS-3-1, PREDWELL\_RETOS-101-1, PREDWELL\_RETOS\_102-1, PREDWELL\_RETOS-201-1, PREDWELL\_RETOS-202-1, PREDWELL\_RETOS-301-1, PREDWELL\_RETOS-302-1, PREDWELL\_RETOS-401-1, PREDWELL\_RETOS-402-1, PREDWELL\_RETOS-403-1, PREDWELL\_RETOS-404-1, PREDWELL\_RETOS-405-1, PREDWELL\_RETOS-406-1, PREDWELL\_RETOS-407-1, PREDWELL\_RETOS-408-1, PREDWELL\_RETOS-409-1, PREDWELL\_RETOS-410-1, PREDWELL\_RETOS-411-1, PREDWELL\_RETOS-412-1, PREDWELL\_RETOS-413-1, PREDWELL\_RETOS-414-1, PREDWELL\_RETOS-415-1, PREDWELL\_RETOS-416-1, PREDWELL\_RETOS-417-1, PREDWELL\_RETOS-418-1, PREDWELL\_RETOS-419-1, PREDWELL\_RETOS-420-1: SPQ and MSSB are missing.
  + PREDWELL\_RETOS-25-1: ZKPQ-ISOL is missing.
  + PREDWELL\_RETOS-106-1: ZKPQ-PARTIES is missing.
  + PREDWELL\_RETOS-115-1: ASI\_C is missing.
  + PREDWELL\_RETOS-119-1: ZKPQ- ISOL, ZKPQ-PARTIES and ZKPQ-WORKACT are missing.
  + PREDWELL\_RETOS-206-1: SPQ\_IR is missing.
  + PREDWELL\_RETOS-307-1: ERQ\_ES is missing.
  + PREDWELL\_RETOS-309-1: ZKPQK\_N\_ANX, ZKPQ\_WORACT, ZKPQ\_PARTIES and ZKPQ\_SS are missing.
  + PREDWELL\_RETOS-310-1: ZKPQ\_AGG\_HOST is missing.
* Project PREDWELL\_ASD:
  + PREDWELL\_ASD\_C001, PREDWELL\_ASD\_C002, PREDWELL\_ASD\_C003, PREDWELL\_ASD\_C004: SP1\_IR, AQS, IUS and BEH are missing.
* Project PREDWELL\_BPD:
  + PREDWELL\_BPD\_C017: PHQ9 is missing.
* Project PREDWELL\_EMAR:
  + PREDWELL\_EMAR\_C001, PREDWELL\_EMAR\_C002: AQS, IUS and BEH are missing.
* Project PREDWELL\_MD:
  + PREDWELL\_MD\_C015: PHQ9 is missing.
  + PREDWELL\_MD\_C016: BSL23 is missing.
* Project PREDWELL\_PSYCH:
  + PREDWELL\_PSYCH\_C001, PREDWELL\_PSYCH\_C003, PREDWELL\_PSYCH\_C004, PREDWELL\_PSYCH\_C005, PREDWELL\_PSYCH\_C006, PREDWELL\_PSYCH\_C007, PREDWELL\_PSYCH\_C008: AQS, UIS and BEH are missing.
  + PREDWELL\_PSYCH\_C013: IUS and BEH are missing.

1 The dataset in question includes outputs from three versions of the same task. Although the outputs are the same, there are slight differences between the versions. Versions 1.0.0.826 and 1.0.0.828 are identical, but both differ from version 1.0.0.824. Version 1.0.0.824 contains variables "GetReady.DEVICE" and "GetReady.RT" which are not present in the other two versions, while it lacks variables "TakeBreak.DEVICE" and "TextDisplay1.DEVICE" which are present in the other versions. That is, "GetReady.DEVICE" and "GetReady.RT" have no values for participants who completed versions 1.0.0.826 or 1.0.0.828, whereas "TakeBreak.DEVICE" and "TextDisplay1.DEVICE" have no values for participants who completed version 1.0.0.824. However, none of these four variables are of interest, so their presence or absence should not pose a problem.