

Contents

1	Category	1
1.1	Category Data - Naming Conventions	1
1.2	Category Data - Key Naming Field Definitions	2
1.3	Category Data - Contents	2
2	Individuation	3
2.1	Individuation - Naming Convention	3
2.2	Individuation Data - Key Naming Field Definitions	3
2.3	Folder Structure	4

1 Category

- In the **category task**, images from different categories are shown to the subjects
- The neural activities from different electrodes are recorded and saved as csv files.
- Each file contains the **2-dimensional array** of all the data recorded from **one** single electrodes.
- In each session of the category dataset, there are 480 trials from 6 different conditions (some cases 7 conditions, see ‘Localizer__category__key.xlsx’ for details).
- Each condition is a type category of images, e.g. bodies, faces, words, houses, scrambles, etc.

1.1 Category Data - Naming Conventions

- The **broadband data** files are named as:
 - `${SUBJECT_ID}_Category_Localizer_Session${SESSION_INDEX}_Broadband_Ch${CHANNEL_INDEX}.dat`
- The **ERP data** files are named as:
 - `Session${SESSION_INDEX}_${TASK}_data_${FILTER}_all_channels_ch${CHANNEL_INDEX}.dat`

1.2 Category Data - Key Naming Field Definitions

The key fields in the naming convention are defined as follows:

- **SESSION_INDEX** is the **session number** (1 or 2 in most cases)
 - This is sometimes omitted if there is only one session
- **TASK** is the task name, i.e. **category**
- **FILTER** is the way the data is filtered, i.e. **bandpass** or **raw**
- **CHANNEL_INDEX** is the **electrode index** 1, 2, 3, ...

1.3 Category Data - Contents

- For the category task, each data file contains the $N \times T$ data matrix from channel X ($X = 1, 2, \dots$)
- Each column is a feature (time point) and each row is a trial
- ERP:
 - Typically $N = 480$ trials and $T = 1500$ time points
 - The data is sampled at **1000hz**
- Broadband:
 - Typically $N = 480$ trials and $T = 150$ time points
 - The data is sampled at **100hz**
- All the trials are aligned to the onset time of the image stimulus
- The 1500 time points start from 500 ms before stimulus onset and ends at 999 ms after stimulus onset.
- The 150 time points start from 500 ms before stimulus onset and ends at 990 ms after stimulus onset.

The `Category_label.dat` file contains the 480-by-1 label vector for each trial condition, the value of the i -th component in the vector represents the image stimulus type for the i -th trial, 1: bodies, 2: faces, 3: words(or shoes), 4: hammers, 5: houses, 6: non-objects. (see ‘`Localizer_category_key.xlsx`’ for details).

In addition to the single trial data, the continuous timecourse data for the category task is also included. You'll find the file `Category_fulltimecourse_signal.dat` for each session. This is the raw time sequence for the entire session. The file is a N by L matrix where N is the number of electrodes and L is the total time length of the entire session (typically around $1e6$ time points sampled at 1000 Hz). In the file `Category_timestamps.dat`, the timestamps for the stimulus onset time of each trial are saved in a N -by-1 vector, where N is the total number of trials. For example, if the i -th elements in the vector is x , then the stimulus onset of the i -th trial happens at the x -th timepoint in `Category_fulltimecourse_signal.dat`. The condition for each trial is the same as `Category_label.dat`.

2 Individuation

- In the experiment the patient was passively viewing a series of faces and the task was to respond whether the face being presented was a male or female face. - There were:
 - 14 different face identities in total,
 - each face identity had 5 different emotions
 - with 3 different gazes.
 - Each individual picture was presented **4 times**
- This yields 840 trials in total in one session i.e. $840 = 14 \times 5 \times 3 \times 4$.

2.1 Individuation - Naming Convention

The files are named as:

`Session${SESSION_INDEX}_${TASK}_data_${FILTER}_all_channels_ch${CHANNEL_INDEX}.dat`

2.2 Individuation Data - Key Naming Field Definitions

- `SESSION_INDEX` is the session number (1 or 2 in most cases), sometimes this is omitted if there is only one session
- `TASK` is the task name, i.e. **individuation**
- `FILTER` is the way the data is filtered, i.e. **bandpass** or **raw**
- `CHANNEL_INDEX` is the index of the electrode, 1, 2, 3, ...

2.3 Folder Structure

- The files of each session are sorted into the corresponding folder.
- For each session, the labels for the trials are written into the following files:
- `${SESSIONID}_${CONDITION}_label.csv`
 - `SESSIONID` in {Session1, Session2}
 - `CONDITION` in {Face, Gender, Expression, Gaze}
 - * **Face** - the index of the face identity, value range 1 - 14
 - * **Gender** - the index of the face gender, 1 - male, 2 - female
 - * **Expression** - the index of the face expression, 1 - angry, 2 - afraid, 3 - happy, 4 - neutral, 5 - sad
 - * **Gaze** - the index of the eye gaze, 1 - front, 2 - left, 3 - right