BrainHack.School.Taiwan

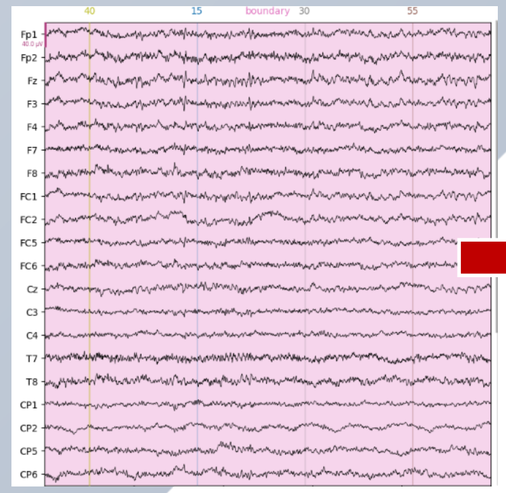
The Journey of Finding N400

Kuo Mei-Chun

Graduate Institute of Linguistics,
National Taiwan University

About
the Project...

Kuo Mei-Chun



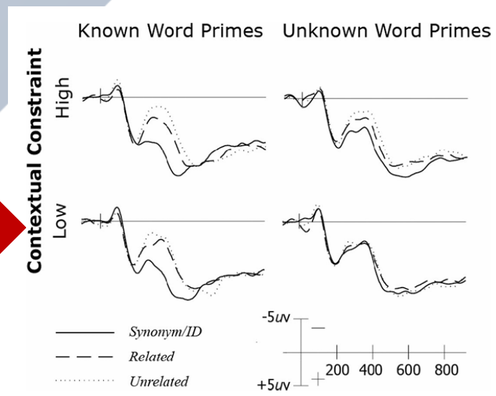


FIGURE 7 Grand average ERPs to target words in priming task at the vertex electrode (MiCE).

Background# 1

Main Interest:
Language processing
(semantic/syntactic)

Background# 2

Taking “Electrophysiology of language processing”
but no EEG experiment experience.

Goal

Hands on practice
Raw EEG data→ ERP component

The First Dataset

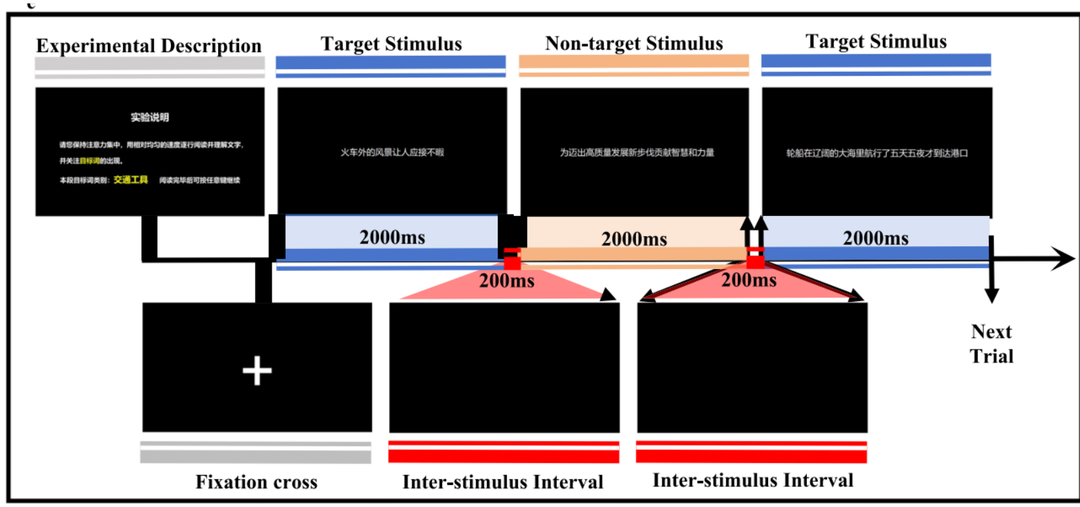
Kuo Mei-Chun

Data Descriptor | Open access | Published: 25 April 2025

TMNRED, A Chinese Language EEG Dataset for Fuzzy Semantic Target Identification in Natural Reading Environments

Yanru Bai, Qi Tang, Ran Zhao, Hongxing Liu, Shuming Zhang, Mingkun Guo, Minghan Guo, Junjie Wang, Changjian Wang, Mu Xing, Guangjian Ni & Dong Ming

Scientific Data **12**, Article number: 701 (2025) | Cite this article



The First Dataset

Kuo Mei-Chun

Target Word Category	Material Examples
Target-Names	Trump was elected the 45th President in U.S. history in 2016. 特朗普于2016年当选美国历史上第45任总统
	Charles has an almost obsessive passion for charitable causes. 查尔斯对于慈善事业有着近乎偏执的热情

Trump→ Person

vs. “Trump decided to lower the tariff.” → Government/Organization

Plan A Hypothesis:

- Names associated with more complex or ambiguous meanings may elicit a larger N400.
- When the target word appears early in the sentence, neural responses may reflect a transition from a fuzzy interpretation to a more concrete meaning as the sentence unfolds.

Difficulties:

1. The phrases weren’t time-locked.
2. The subject phrases did not always in the first position.

The First Dataset

Kuo Mei-Chun

Concrete

Number	Material statement	Labels	Context	Subject	Condition	Translation
0010	查尔斯是英国王室首个获得大学学位的继承人	24	Foreign	Person	ForeignPers	Charles was the first heir to the British throne to earn a university degree.
0011	外界对查尔斯的评价多是古板与老派	25	Foreign	Person	ForeignPers	Charles is mostly regarded by outsiders as curmudgeonly and old-fashioned.
0012	事实上，古建筑修复是查尔斯的兴趣所在	26	Foreign	Person	ForeignPers	In fact, the restoration of old buildings is Charles' passion.
0013	在艺术领域，查尔斯一直扮演一个老派的角色	27	Foreign	Person	ForeignPers	In the arts, Charles has always played an old-fashioned role.
0014	海格罗夫庄园仿佛是查尔斯人格与理念的外现	28	Foreign	Person	ForeignPers	Highgrove Estate seems to be a reflection of Charles' personality and philosophy.
0026	要推进国家治理体系和治理能力现代化	40	Fuzzy	NoSub	FuzzyNoSu	We must promote the modernisation of the national governance system and governance cap
0027	为全面建设社会主义现代化国家提供有力保障	41	Fuzzy	NoSub	FuzzyNoSu	Provide a strong guarantee for the comprehensive construction of a modern socialist countr
0028	发挥法治固根本、稳预期、利长远的保障作用	42	Fuzzy	NoSub	FuzzyNoSu	Give full play to the rule of law to consolidate the fundamentals, stabilise expectations, and
0029	只有筑牢安全屏障，才能风雨不动安如山	43	Fuzzy	NoSub	FuzzyNoSu	Only by building a solid security barrier can we be secure in the wind and rain.

Abstract

Plan B Hypothesis:

- More concrete sentences with clear subjects would elicit a smaller N400 compared to the more abstract ones.

Problem: No complete stimulus & No event code–to–condition mapping for epoch segmentation.

The Second Dataset

Kuo Mei-Chun

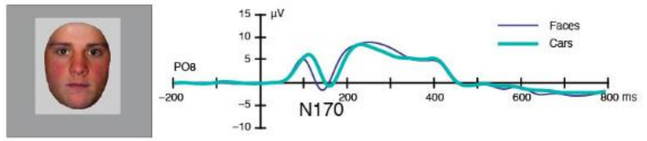
ERP CORE

The ERP CORE is a freely available online resource consisting of optimized paradigms, experiment control scripts, example data from 40 participants, data processing pipelines and analysis scripts, and a broad set of results for 7 different ERP components obtained from 6 different ERP paradigms:

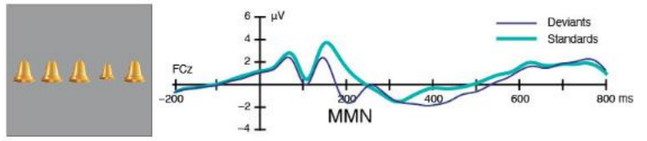
- N170 (Face Perception Paradigm)
- MMN (Passive Auditory Oddball Paradigm)
- N2pc (Simple Visual Search Paradigm)
- N400 (Word Pair Judgement Paradigm)
- P3b (Active Visual Oddball Paradigm)
- LRP and ERN (Flankers Paradigm)

The experiment control scripts, data, and data analysis scripts can be downloaded at https://doi.org/10.18115/D5JW4R

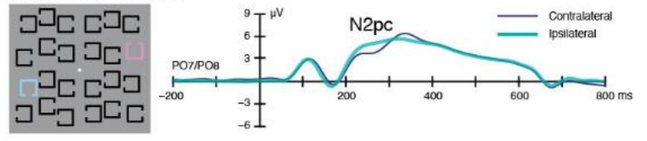
Face Perception N170



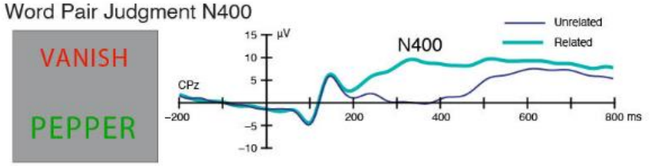
Passive Auditory Oddball MMN



Visual Search N2pc



Word Pair Judgment N400



We promote best practices in ERP research via workshops, software, books, advice, data sharing, & methods development.

Steve Luck
lucklab.ucdavis.edu
mindbrain.ucdavis.edu/people/sjluck
@stevenjluck
Google Scholar Profile

Emily Kappenman
emilykappenman.org
@emilykappenman
Google Scholar Profile

The Second Dataset

Kuo Mei-Chun

Prime

Semantically Unrelated Word

Related Word

N400: Unrelated > Related

The Second Dataset

Kuo Mei-Chun

ERP_CORE Public

forked from andrewxstewart/ERP_CORE

Pin | Watch 0 | Fork 0 | Star 0

master | 1 Branch | 0 Tags | Go to file t | Add file | Code

This branch is up to date with `andrewxstewart/ERP_CORE:master`. | Contribute | Sync fork

andrewxstewart Clean up the readme — ab507d1 · 5 years ago — 6 Commits

ERN Analysis Files	Initial add of ERP CORE scripts and info	5 years ago
LRP Analysis Files	Initial add of ERP CORE scripts and info	5 years ago
MMN Analysis Files	Initial add of ERP CORE scripts and info	5 years ago
N170 Analysis Files	Initial add of ERP CORE scripts and info	5 years ago
N2pc Analysis Files	Initial add of ERP CORE scripts and info	5 years ago
N400 Analysis Files	Initial add of ERP CORE scripts and info	5 years ago
P3 Analysis Files	Initial add of ERP CORE scripts and info	5 years ago

About

The ERP CORE is a freely available online resource consisting of optimized paradigms, experiment control scripts, example data from 40 participants, data processing pipelines and analysis scripts, and a broad set of results for 7 different ERP components obtained from 6 different ERP paradigms

Readme
Activity
0 stars
0 watching
0 forks

Releases

Methodology

Kuo Mei-Chun

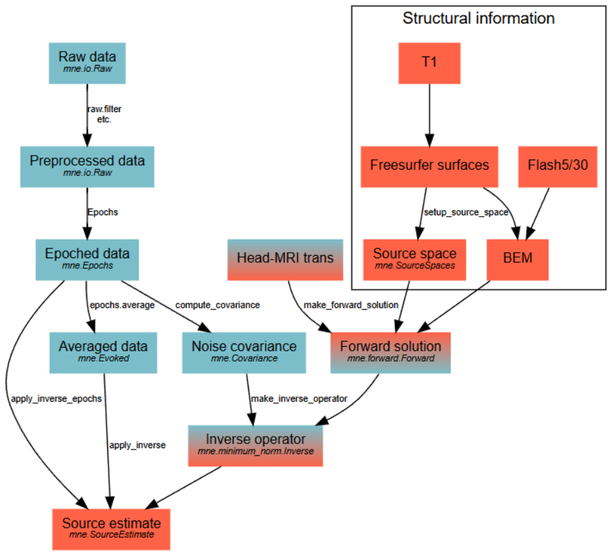
MNE Install Documentation API Reference Get help Development Choose version

Section Navigation

- Tutorials
- Examples
- Glossary
- Implementation details
- Design philosophy
- Example datasets
- Command-line tools
- Migrating from other analysis software
- The typical M/EEG workflow
- How to cite MNE-Python
- Papers citing MNE-Python

Overview

This section describes a typical MEG/EEG workflow, eventually up to source reconstruction. The workflow is summarized in Workflow of the MNE software. References below refer to Python functions and objects.



Workflow of the MNE software

On this page

- Overview
- Preprocessing
- Epoching and evoked data
- Source localization
- References

Search Ctrl + K

- MNE-Python
- EEG-ERP Preprocessing
  - Filtering EEG Data
  - Artifacts in EEG Data
  - Segmentation into ERP epochs
  - Averaging and Re-Referencing ERPs
- Group Analysis of ERP Data

MRI Data

- Introduction to MRI Data
- Learning Objectives
- Reading and Visualizing Structural MRI Data
- Working with NIfTI images

Neural Data Science in Python

Welcome! This online textbook is aimed primarily at students and researchers in neuroscience and cognitive psychology who want to learn how to work with and make sense of data using Python. It is also accessible for students with a computer science background who want to learn how to apply their skills to neuroscience. The textbook assumes no prior knowledge of Python, or any other programming language. If you do have prior knowledge of Python and want to learn how to apply it to neuroscience, you can skip the first few chapters and start with Chapter 4.

Filtering EEG Data

As described in the previous section on Time and Frequency Domains, a complex time-var
EEG can be represented as a combination of sine waves of many different frequencies. Hu
largely comprises signal power in a range of frequencies from 1–30 Hz; there is some evid
higher frequencies may also carry important neurophysiological information, however in n
studies — and certainly in ERP studies — the vast majority of research questions concerns
the 1–30 Hz frequency range.

Sources of noise also manifest as oscillating frequencies that are picked up by EEG. In part
frequency noise comes from sources such as movement of the head and electrode wires, a
on the scalp. Low frequency noise appears as slow drifts in the EEG signal over many seco
high frequency noise comes from sources including electromagnetic interference, and mus
contractions (especially facial and neck muscles). High frequency noise looks like rapid up-
(like the teeth of a saw) in the EEG.

EEG-ERP Preprocessing

- Filtering EEG Data
- Artifacts in EEG Data
- Segmentation into ERP epochs
- Averaging and Re-Referencing ERPs
- Group Analysis of ERP Data

The First

Try

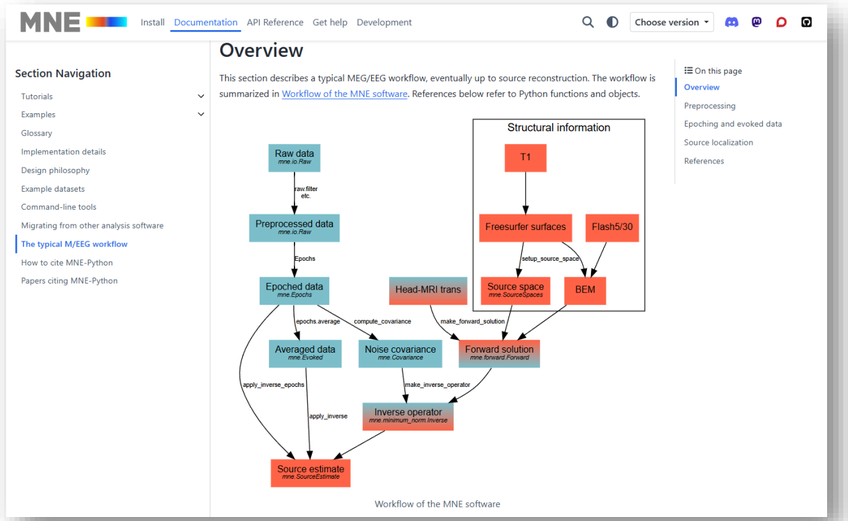
to find the N400

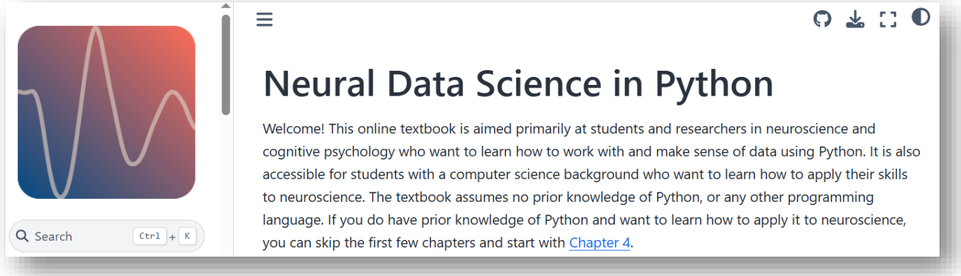
EEG Data

```
print_tree(folder_path)
```

```
├── 1
│   ├── 1_N400.fdt
│   ├── 1_N400.set
│   ├── 1_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.fdt
│   ├── 1_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.set
│   ├── Icon_
│   └── graphs
│       ├── 1_N400_ICA_Weights.pdf
│       └── Icon_
├── 10
│   ├── 10_N400.fdt
│   ├── 10_N400.set
│   ├── 10_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.fdt
│   ├── 10_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.set
│   ├── Icon_
│   └── graphs
│       ├── 10_N400_ICA_Weights.pdf
│       └── Icon_
├── 11
│   ├── 11_N400.fdt
│   ├── 11_N400.set
│   ├── 11_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.fdt
│   ├── 11_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.set
│   ├── Icon_
│   └── graphs
│       ├── 11_N400_ICA_Weights.pdf
│       └── Icon_
├── 12
│   ├── 12_N400.fdt
│   ├── 12_N400.set
│   ├── 12_N400_shifted_ds_reref_ucbip_hpfilt_ica_weighted.fdt
```

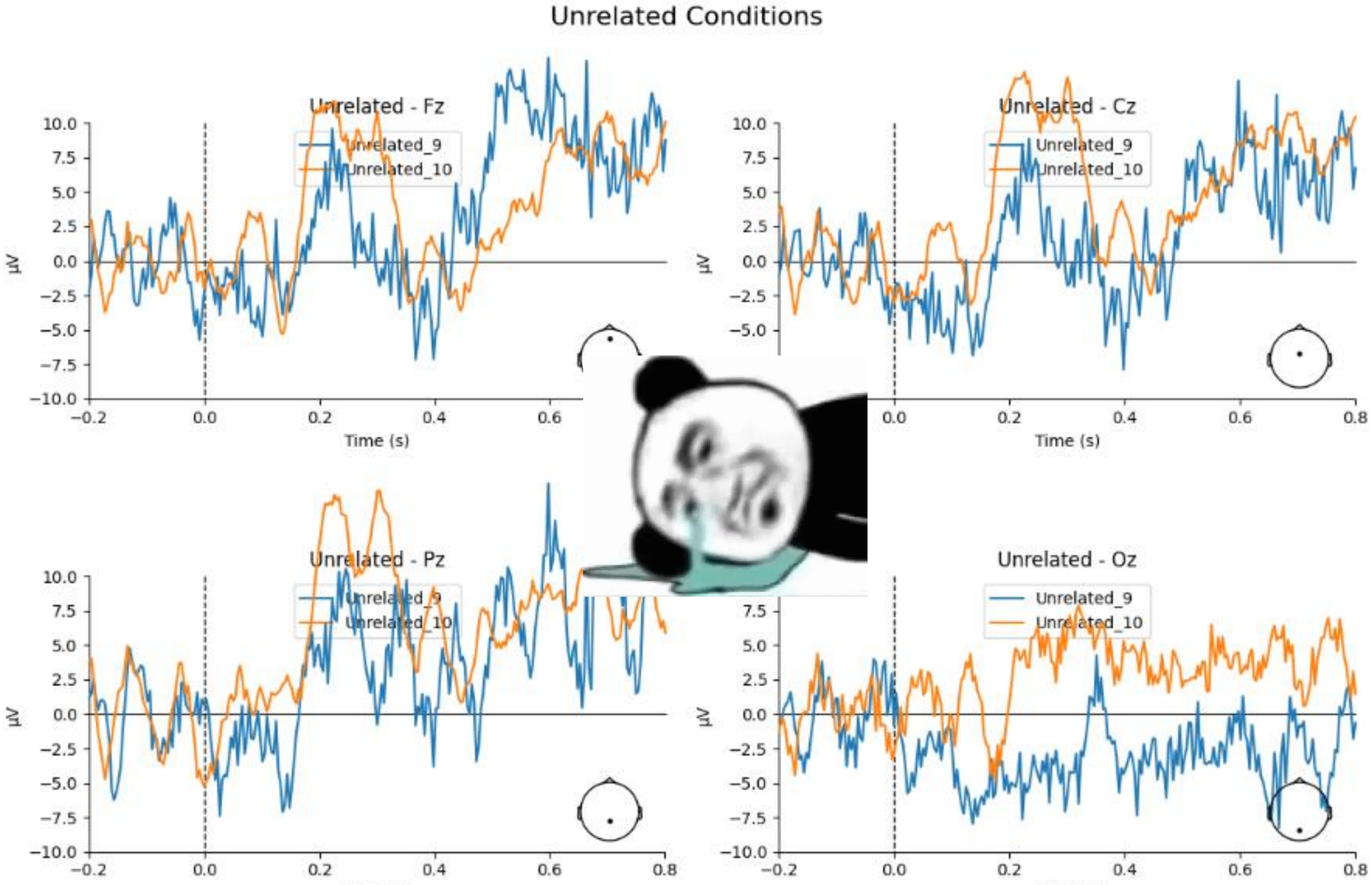
Tutorial





The First Try

to find the N400



The Second

Try

to find the N400

Individual-Subject EEG and ERP Processing Procedures

Script #1: Import_Raw_EEG_Shift_DS_Reref_Hpfilt.m

This script (located in …/N400/EEG_ERP_Processing) performs initial processing on the raw continuous EEG data file using the following operations:

1. Load the continuous raw EEG data file
   - The raw EEG data files already have been converted from their original file format to the .set EEGLAB file format. If you are applying this script to your own data, you will need to import your data into EEGLAB before running this script (see the EEGLAB documentation for more information on importing EEG files) and adjust the file names and paths to match your data.
2. Shift the stimulus event codes later in time by 26 ms to account for the LCD monitor delay (as measured with a photosensor)
   - Most LCD monitors have a delay between the time when the image is sent from the computer to the monitor and the time when the visual information actually appears on the screen. This delay, measured with a photosensor in our laboratory, was 26 ms (but this can vary quite a bit across monitors, so you'll need to measure the delay for your monitor if you are applying this script to your own data). The stimulus event codes are therefore shifted later in time by the measured amount to account for the monitor presentation delay.
3. Downsample the data from 1024 Hz to 256 Hz to speed data processing
   - This function automatically applies the appr… new sampling rate.
4. Re-reference the data to the average of P9 and P10 … bipolar HEOG channel (HEOG_left minus HEOG_… (VEOG_lower minus FP2)
   - Bipolar EOG signals are particularly helpful in identifying ocular artifacts and will be used during artifact rejection (see Luck, 2014 for more details).
5. Add 3-D channel location information corresponding to the International 10-10 System
   - A channel location file that is appropriate for our recorded data set is included in the downloaded materials (standard-10-5-cap385.elp). Additional channel locations files can be found by consulting the help page for pop_chanedit.
6. Remove the DC offsets and apply a high-pass filter (non-causal Butterworth impulse response function, half-amplitude cut-off at 0.1 Hz, 12 dB/oct roll-off)
   - These data were acquired with an EEG system that records at DC (i.e., uses no high-pass filter during acquisition).

MNE

```
montage = mne.channels.make_standard_montage('standard_1020')
for raw in raw_list:
    raw.set_montage(montage, on_missing='ignore')
```

The Second

Try

to find the N400

```
N400_01 = mne.io.read_raw_eeglab('/content/drive/MyDrive/neuro something/N400 Analysis Files/N400 Analysis Files/N400/1/1_N400.set',
N400_01.info
```

```
Reading /content/drive/MyDrive/neuro something/N400 Analysis Files/N400 Analysis Files/N400/1/1_N400.fdt
Reading 0 ... 585727  =      0.000 ...   571.999 secs...
```

General

MNE object type	Info
Measurement date	Unknown
Participant	Unknown
Experimenter	Unknown

Acquisition

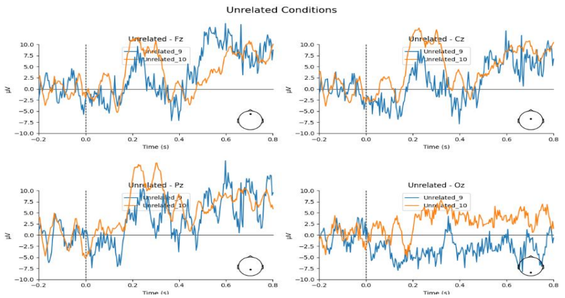
Sampling frequency	1024.00 Hz

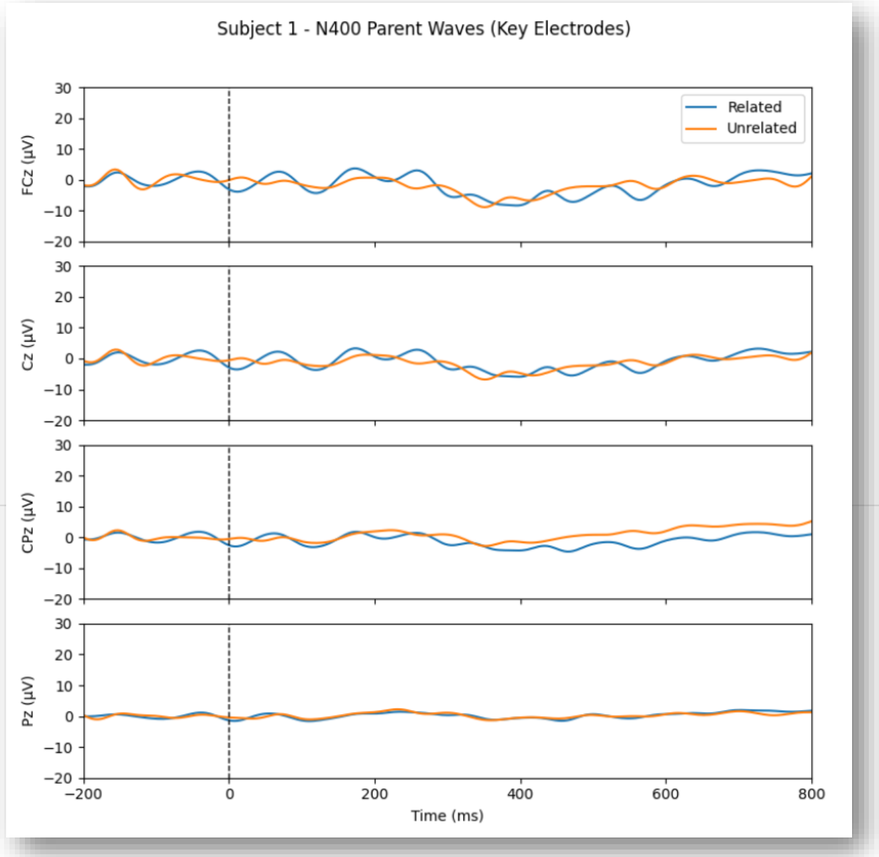
Channels

EEG	33
Head & sensor digitization	33 points

Filters

Highpass	0.00 Hz





The Final

Try

to find the N400

1_Import_Raw_EEG_Shift_DS_Reref_Hpfilt.m

2_ICA_Prep.m

4_Remove ICA

偵測 EOG（眼動/眨眼）成分

加上雙極通道（HEOG 與 VEOG）

5_Elist_Bin_Epoch.m

6_Artifact_Rejection

7_Average_ERPs

Evoked.plot()

Evoked.plot(picks=ch)

8_Plot_Individual_Subject_ERP

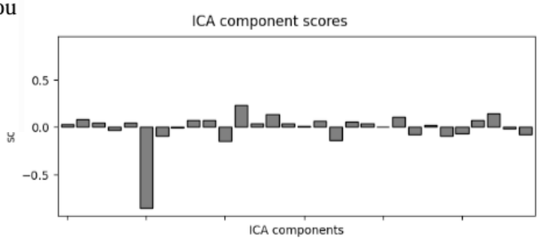
9_Grand_Average_ERP

10_Plot_Grand_Average_ERPs

Script #3: Run_ICA.m

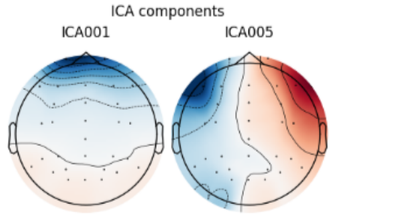
This script (located in …/N400/EEG_ERP_Processing) uses the output from Script #2 and computes the ICA weights using the following operations:

IMPORTANT: The results of ICA decomposition (i.e., the ordering of the components, the scalp topographies, and the time courses of the components) will differ slightly each time the ICA weights are computed. This is because ICA decomposition starts with a random weight matrix (and randomly shuffles the data order in each training step), so the convergence is slightly different every time it is run. As a result, the topographic maps of the ICA weights and the Excel spreadsheet (ICA_Components_N400.xlsx) containing the list of ICA component(s) to be removed for each subject included in this package will no longer be valid if your run the ICA decomposition. Consequently, to avoid confusion or accidental overwriting of relevant data files, this script has been commented out. You must skip the steps in this script and proceed directly to Script #4 if you want the results you get from these scripts to *exactly* match our results. Alternatively, this script can be un-commented and the ICA weights can be re-computed, but note that in that case the ICA component(s) will need to be re-chosen and the new values will need to be entered into the Excel file prior to running Script #4.



Identify EOG Artifacts from ICA Components

MNE has an algorithm that attempts to automatically identify ICA components. The `find_bads_eog()` function computes correlations between each IC and channels that the researcher has designated as **EOG** (electro-oculogram) channels. These are electrodes, the same as EEG electrodes, but intentionally placed close to the eyes specifically to monitor for blinks and eye movements. These are typically placed above and below one eye (to monitor blinks and vertical eye movements, as well as on the temples of the head laterally to the eyes (to monitor horizontal eye movements).

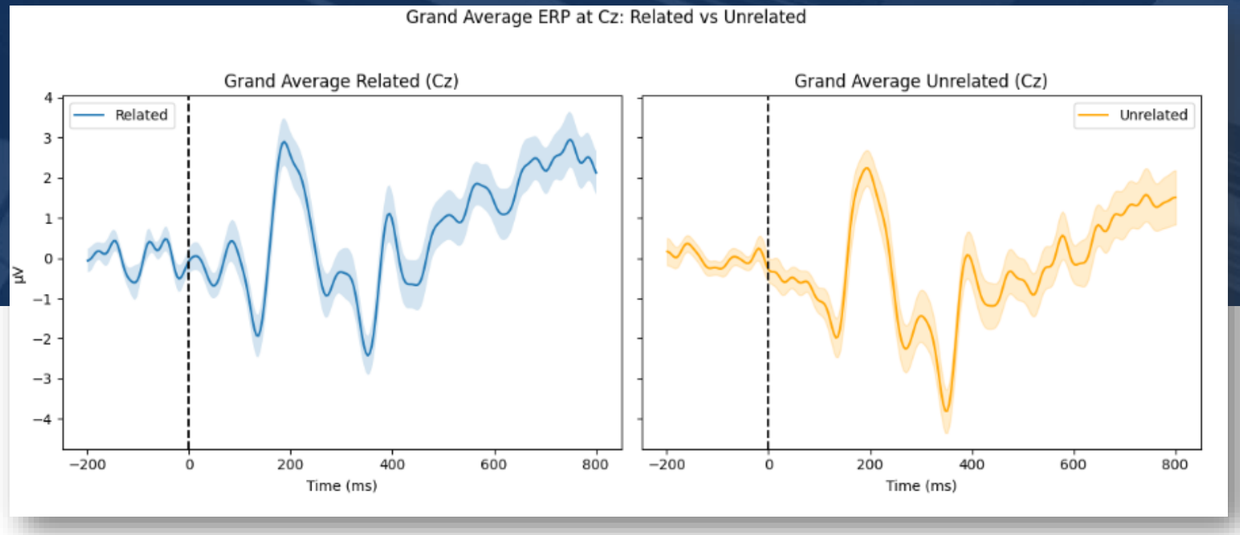


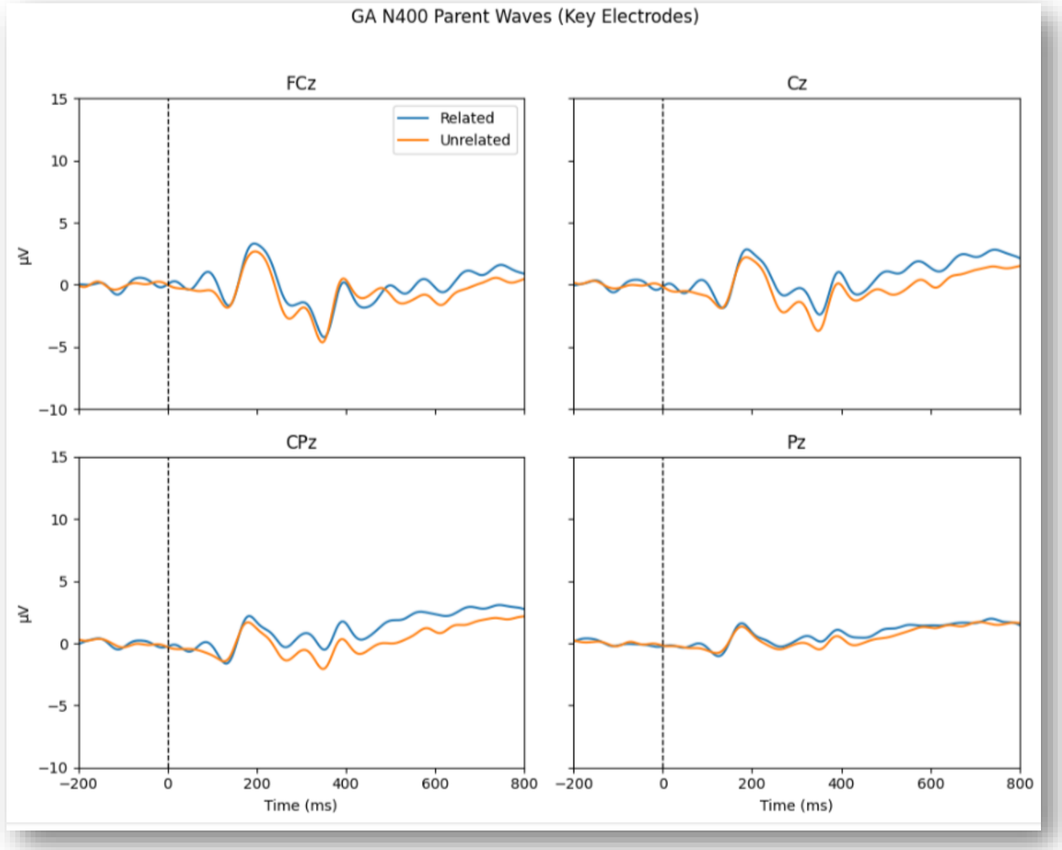

```
Subject 2 EOG-like ICA components: [np.int64(1), np.int64(5)]
```

The Final

Try

to find the N400

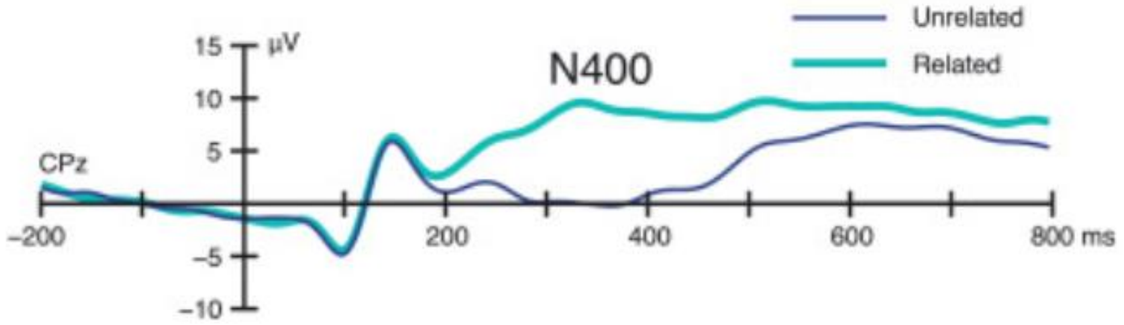




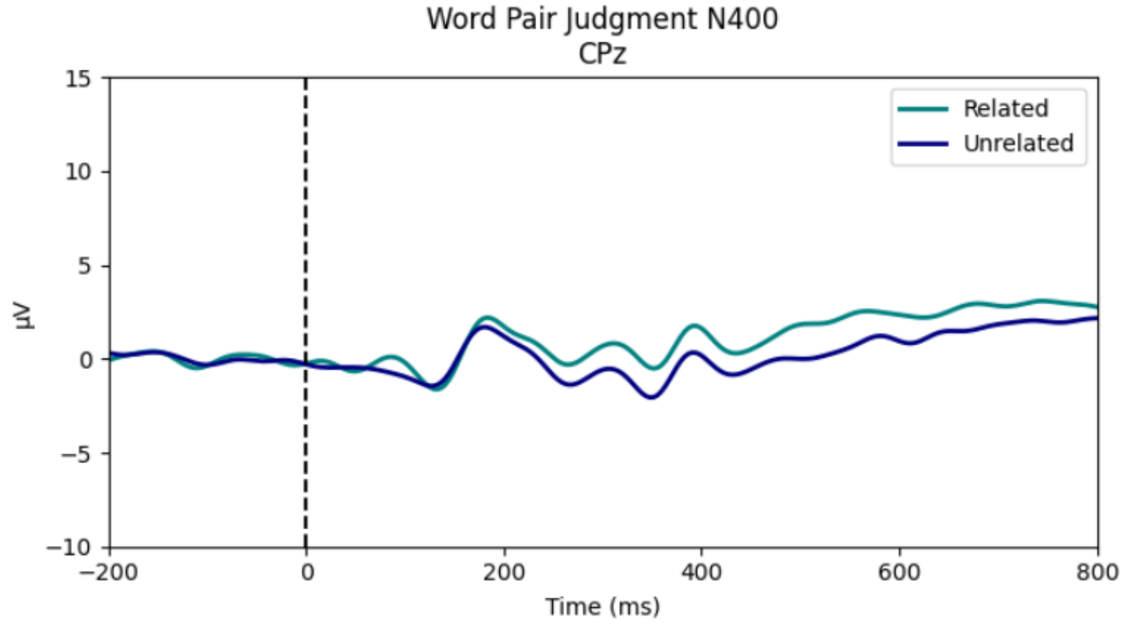
The Final

Try

to find the N400



Kappenman et al. (2021)



1. So many details…
2. Visualization
3. Understanding what you’re doing

Challenges and Difficulties

during the whole process