# COG-BCI database: A multi-session and multi-task EEG cognitive dataset for passive brain-computer interfaces

M. F. Hinss, E. S. Jahanpour, B. Somon, L. Pluchon, F. Dehais, R. N. Roy


## Abstract

Brain-Computer Interfaces, and especially passive Brain-Computer Interfaces (pBCI), with their ability to estimate and detect mental states, are receiving increasing attention from both the scientific and the research and development communities. Many pBCIs aim to increase the safety of complex work environments such as in the aeronautical domain. Therefore, mental workload, vigilance and decision-making are some of the most commonly examined aspects of cognition within this field of research. A large proportion of pBCIs involve a component of machine learning and signal processing as the data that are collected need to be transformed into a reliable estimate of the users' current mental state (e.g. mental workload). Improving this component is a major challenge for researchers, requiring large quantities of data. While data sharing is common for the active BCI community, open pBCI datasets are scarcer and generally incomplete with regards to the information they report. This is particularly true for datasets encompassing several tasks or sessions, which are of importance for tackling the challenges of transfer learning. Testing new pipelines, feature extraction algorithms and classifiers are central issues for future advances in research within this domain, as well as for algorithm benchmark and research reproducibility.The COG-BCI database presented here is comprised of the recordings of 29 participants over 3 individual sessions with 4 different tasks designed to elicit different cognitive states. This results in a total of over 100 hours of open electrophysiological (EEG) and electrocardiogram (ECG) data. The project was validated by the local ethical committee of the University of Toulouse (CER number 2021-342). The dataset was validated on a subjective, behavioral and physiological level (i.e. cardiac and cerebral activity), to ensure its usefulness to the pBCI community. This body of work represents a large effort to promote the use of pBCIs, as well as the use of open science.

# Tasks

The COG-BCI database comprises four different tasks independently assessed over three sessions (spaced exactly one week apart). Furthermore, a resting state (one minute of eyes open and one minute of eyes closed) was collected at the beginning and the end of each session. The four different tasks were chosen in order to assess several aspects of executive functioning. The MATB and N-back task were used to assess different levels of mental workload. The Psychomotor Vigilance Task (PVT) was used to study vigilance decrement and the Flanker task for decision making and conflict. Each task lasted from five to ten minutes. For each participant and session, the order of tasks was pseudorandomized and is reported in the Database Notebook.

All the tasks were coded with home-made scripts in MATLAB, with the PsychToolBox-3 (http://psychtoolbox.org/) software. For each session, participants were comfortably seated approximately 50 cm away from a 120 Hz refresh rate computer screen. During each task, their responses were recorded either through a keyboard (N-back task, PVT and Flanker task) or through both a joystick (Extreme 3D Pro Logitec) and keyboard (MATB). The responses, reaction times, electrocardiogram (ECG) and EEG signals were collected from each participant throughout the entire session.

## 1. Psychomotor Vigilance Task

The PVT is a 10-minute measure of Vigilance (Dinges & Powell, 1985; Lamond et al., 2005). Participants receive instructions on the computer screen stating that they have to press the spacebar of the keyboard as soon as a timer appears on the screen. Each trial starts with an interstimulus interval (ISI) with a duration jittered between 2-10 seconds. After the ISI, the stimulus/timer appears. The timer continues to run until the participant reacts by hitting the spacebar. After the button press is registered, the timer stops and displays the final reaction time on the computer screen for another 500 ms. The task was designed to closely resemble the PC-PVT 2.0, an established computer version of the PVT (Reifman et al., 2018). Participants completed a total of 90 trials for this task during each session.

## 2. N-Back Task

The N-Back Task is a widely used measure of both working memory and mental workload. On a computer screen participants are presented with single numbers appearing for a short period of time (Brouwer et al., 2012; Jaeggi et al., 2010). Participants are instructed to remember the order in which the numbers appear and to react with a button press if the presented number is the same as the N-th number presented before. The N here is a variable that determines the difficulty of a particular block. With the increasing size of N, the difficulty of the task increases, as more numbers need to be retained. Trials begin with the presentation of a number for 500 ms, followed by a blank screen for 1500 ms. If the number presented is a hit number (e.g. in the 1-back condition, the same as the previous number), participants are instructed to respond by hitting the spacebar of the keyboard. In the 0-back condition, participants are instructed to respond whenever a previously determined target number (e.g.

“3”) appears. The frequency of hit numbers appearing in all three conditions is fixed at 1/3 (16 hit trials per block). In the 2-back condition, five conflict trials are added per block. The conflict trials are characterized by a number being followed immediately by the same number again. This should not prompt any reaction from the participant but has been suggested to result in eliciting conflict (Oberauer et al., 2005). As these trials would also occur during full randomization of the numbers, they have no adverse effect on the participant’s performance. Participants completed three blocks of each  a 0-back, a 1-back and a 2-back condition ( for a total of 9 blocks). Each block has a duration of around two minutes and consists of 48 trials. The total time to complete 3 blocks of one condition was 6 minutes. Before the onset of each block, the participant was informed about the condition of the block as well as given a short instruction on what to do.

## 3. MATB

NASA developed the MATB-II task[1] to assess task-switching and mental workload capacities (Santiago-Espada et al., 2011). Here participants are presented with up to 6 different tasks that they have to complete simultaneously. This provides a highly realistic environment of operational systems that the researcher can control to create different degrees of difficulty. An adapted version[2], coded in Matlab but providing the same measures as the original MATB-II task, was used (Verdière et al., 2020). For a full description of the original task and subtasks refer to Santiago-Espada et al. (2011). For this study, combinations of four of the available subtasks of the MATB were used. In the tracking task (TRACK), participants are presented with a moving target inside of a window. The goal is to keep, using a joystick, the target within the window. The degree of difficulty can be adapted by modifying the degree and the speed at which the target moves. For the System monitoring task (SYSMON), participants have to monitor gauges and warning lights. Action is required in the absence of a green light, the presence of a red light and deviations of four moving pointers dials from a midpoint. The system monitoring task is controlled by input into specific keyboard commands. The degree of difficulty can be adapted by increasing the number of events to which the participant has to react to. In the communications task (COMM), participants are required to listen to radio messages and determine if they are of importance to the operator (calling his/her call sign) or not (calling another call sign). If the message is relevant, the operator is required to change the frequency of a radio channel to a frequency specified in the message. The last task that is used is the Resource Management Task (RESMAN). Participants are presented with an interface displaying two main tanks and four subsidiary tanks interconnected via eight pumps bearing various fuel flows. The goal is to maintain a specific level of fluid in both of the main tanks. Participants can do this by activating or deactivating the pumps. In order to increase the difficulty of the task, events such as pump failures can be introduced.

In the current study, participants performed three independent 5-minute runs of three degrees of difficulty (i.e. see scenarios 2, 3 and 4 in Cabon et al., 2006). For the easy condition, participants only engage in the system monitoring and the tracking task. For the medium condition, participants engage in both tasks as well as the fuel management task. For the difficult condition, the communication task is added as well as the tracking task is made more difficult. Before the start of each run, participants also received a short instruction.

[1] https://ntrs.nasa.gov/citations/20110014456
[2] https://github.com/VrdrKv/MATB/blob/master/README.md

# 4. Flanker

The Flanker task is a simple choice reaction task that elicits conflict during a binary decision (Eriksen & Eriksen, 1974). Participants are presented with stimuli composed of 5 arrows in the center of a computer screen. They are instructed to react to the middle arrow and ignore the distracting (flanker) arrows on either side. These so-called flanker stimuli can either point in the same direction (congruent condition) or in the opposite direction (incongruent condition) as the central target. A typical stimulus may therefore look like ‘< < > < <’ or ‘< < < < <’. The correct response to the first stimulus is ‘>’ while the response to the second stimulus is ‘<’. Each trial begins with an ISI of 2000 ms. Following the ISI, the stimulus is presented. Each of the four possible stimuli ('> > < > >' '< < < < <' '< < > < <' '> > > > >') are presented equally frequently (0.25) in a pseudorandom order. The stimulus is presented for a fixed time of 16 ms which was determined in a pilot study based on changes in error rates. After the stimulus presentation, the same blank screen with a fixation cross is shown for 2250-2750 ms. In this time period, the participant is required to respond. At the end of the trial, the participant receives feedback about the outcome (correct, incorrect, miss) of her/his trial displayed for 500 ms. In total, 120 trials are performed, with a complete run taking around 10 minutes. Before the onset of the task, the participant received instructions on what to do for all trials.

# Data collection

The measuring equipment used in this experimental campaign was an EEG system (electroencephalography) with 64 active Ag-AgCl electrodes (ActiCap, Brain Products Gmbh) and an ActiCHamp amplifier (Brain Products, Gmbh) positioned according to the extended 10-20 system (Oostenveld and Praamstra, 2001). Data were sampled at 500 Hz. For participants 1-9, electrode Cz could not be recorded and therefore is not present in the dataset. In addition to the brain activity, electrode 10 (named ECG in the dataset) was dedicated to recording peripheral electrocardiographic activity and was placed on the left fifth intercostal. The LabStreamingLayer software[3] was used to accurately synchronize the stimuli display and physiological data (EEG, ECG) with the responses of the participants to the task.

To obtain the precise location of the electrodes on the scalp, a 3D scanning camera by STRUCTURE® and the get_chanlocs plug-in[4] developed specifically for electrode localisation purposes was used (Homölle & Oostenveld, 2019). The specific channel locations are documented in the chanlocs folder for each participant.

---

[3] https://labstreaminglayer.readthedocs.io/info/intro.html
[4] https://github.com/sccn/get_chanlocs/wiki

# Data format

The data are formatted following the BIDS standard (https://bids.neuroimaging.io/). An exemplary folder tree can be seen below.

```
+---sub-01
|  +---ses-S1
|  |  +---behavioral
|  |  |      0-Back.mat
|  |  |      ….
|  |  +---chanlocs
|  |  |      get_chanlocs.txt
|  |  \---eeg
|  |         Flanker.fdt
|  |         Flanker.set
|  |         MATBdiff.fdt
|  |         MATBdiff.set
|  |  |      ….
|  +---ses-S2
|  |     ….
+---sub-02
…
```

Participants are numbered from 1 to 29. For each session, the behavioral results, as well as the exact electrode locations and individual datasets for each task, are provided. The EEG data is saved in the .set/ .fdt file format (two files per dataset). The resting state is divided into 4 different datasets: RS_Beg refers to the resting state at the beginning of the session and RS_End to the resting state at the end of the session; EC is the abbreviation for eyes closed and EO refers to eyes open.

Furthermore, a Notebook file is available which details the order in which the tasks were acquired, if the recording had to be interrupted at any point as well as other comments if applicable. The RSME and the KSS scores are included in the database in form of a table in .txt format.

| Number in the Notebook | Task |
|---|---|
| 1 | PVT |
| 2 | Flanker |
| 3 | Two-Back |
| 4 | One-Back |
| 5 | Zero-Back |
| 6 | MATB-Easy |
| 7 | MATB-Medium |
| 8 | MATB-Difficult |

The trigger list file contains all LSL triggers and what they refer to.

# Other

## MATB Behavioral Data

The behavioral output of the MATB is saved in a MATLAB structure, with individual substructures for each of the tasks.

### TRACK

2 columns with the X and Y coordinates of the tracking tasks (2 Hz sampling rate).

### SYSMON

2 columns referring to the onset of a specific alarm (column 1) and reaction time (column 2).

### RESMAN

2 columns with the amount of fuel in the relevant reservoirs (1 Hz sampling rate).

### COMM

A table with 5 categories:

- **Target**: Was the radio message a target (1 = yes)
- **TargetRadio:** Which radio needed to be changed (1-4)
- **TargetFrequency:** What was the target frequency
- **Reacted:** Did the participant react to the radio message and change some radio frequency (1 = yes)
- **Correct** Was the change in the radio frequency correct (1 = yes)

# References


Brouwer, A.-M., Hogervorst, M. A., Erp, J. B. F. van, Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering*, *9*(4), 045008. https://doi.org/10.1088/1741-2560/9/4/045008

Cabon, P., Rome, F., Mollard, R., Vollard, C., & Reuzeau, F. (2007). Calibration of mental workload measures during the design process of future Airbus military transport Aircraft: the Workload Assessment Method (WAM). *Aviation space and environmental medicine*, *78*(3), 335.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652-655. https://doi.org/10.3758/BF03200977

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149. https://doi.org/10.3758/BF03203267

Homölle, S., & Oostenveld, R. (2019). Using a structured-light 3D scanner to improve EEG source modeling with more accurate electrode positions. *Journal of Neuroscience Methods*, *326*, 108378. https://doi.org/10.1016/j.jneumeth.2019.108378

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394-412. https://doi.org/10.1080/09658211003702171

Lamond, N., Dawson, D., & Roach, G. D. (2005). Fatigue Assessment in the Field : Validation of a Hand-Held Electronic Psychomotor Vigilance Task. *Aviation, Space, and Environmental Medicine*, *76*(5), 486-489.

Oberauer, K., Wilhelm, O., Schulze, R., & Süss, H.-M. (2005). Working memory and intelligence - their correlation and their relation : Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61-65.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical neurophysiology*, *112*(4), 713-719. https://doi.org/10.1016/S1388-2457(00)00527-7

Reifman, J., Kumar, K., Khitrov, M. Y., Liu, J., & Ramakrishnan, S. (2018). PC-PVT 2.0 : An updated platform for psychomotor vigilance task testing, analysis, prediction, and visualization. *Journal of Neuroscience Methods*, *304*, 39-45. https://doi.org/10.1016/j.jneumeth.2018.04.007

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock, J. R. (2011). *The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research : A User's Guide* (L-20031). https://ntrs.nasa.gov/citations/20110014456

Verdière, K. J., Albert, M., Dehais, F., & Roy, R. N. (2020). Physiological synchrony revealed by delayed coincidence count: application to a cooperative complex environment. *IEEE Transactions on Human-Machine Systems*, *50*(5), 395-404. https://doi.org/10.1109/THMS.2020.2986417