Leg je hersenen te luister! Online meting van auditieve aandacht via elektroencephalografie (EEG) bij individuen en groepen van lerenden

A Data Management Plan created using DMPonline.be

Creators: Alexander Bertrand, n.n. n.n., Bert De Smedt https://orcid.org/0000-0002-3313-3278, Elien Bellon https://orcid.org/0000-0002-0607-6634, n.n. n.n.

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Principal Investigator: Bert De Smedt https://orcid.org/0000-0002-3313-3278, Alexander Bertrand, n.n. n.n., n.n. n.n., Elien Bellon https://orcid.org/0000-0002-0607-6634

Project Administrator: n.n. n.n.

Grant number / URL: IDN/23/006

ID: 202030

Start date: 01-10-2023

End date: 30-09-2027

Project abstract:

Attention plays a vital role in understanding how students learn and get distracted in classroom situations. Educational scientists study attention via post-hoc assessments (self- reports or cognitive tests), yet these are inaccurate and do not capture fluctuations in attention over time. On the other hand, biomedical engineers have recently made progress in objectively tracking attention to auditory stimuli via brain recordings such as electroencephalography (EEG), but these methods are not yet adapted to the complexity of the classroom context. We propose a strong interdisciplinary collaboration between biomedical engineers and educational neuroscientists, to develop novel EEG-based markers of auditory attention during learning. This will lead to new data-driven EEG-based algorithmic tools that quantify attention over time, both in groups and individuals. These tools will be used to identify student profiles and to investigate the effects of (neurofeedback) interventions.

Last modified: 18-10-2023

Leg je hersenen te luister! Online meting van auditieve aandacht via elektroencephalografie (EEG) bij individuen en groepen van lerenden

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
EEG Lecture	EEG data from subjects listening/watching a video of a lecture	*New*	Digital	EEG data	EDF, BIDS, mat, binary (depending on processing stage)	200GB	NA
EEG neurofeedback	EEG data from subjects listening/watching a video of a lecture while getting feedback on attention levels	*New*	Digital	EEG data	EDF, BIDS, mat, binary (depending on processing stage)	200GB	NA
EEG classroom	EEG data from subjects in a classroom attending a lecture	*New*	Digital	EEG data	EDF, BIDS, mat, binary (depending on processing stage)	100GB	NA
Behavioral data Lecture - Dataset 1	Behavioral data (on-topic performance test, self-report questionnaires of attention and cognitive tests) collected from the same participants as in dataset EEG lecture	New	Digital and physical	Observational data	.csv, .pdf	< 100 GB	Data of around 80 participants
Behavioral data profiles - Dataset 2	Behavioral and cognitive data will be collected to identify attention profiles	New	Digital and physical	Observational data	.csv, .pdf	< 100 GB	Data of around 250 participants
Behavioral data neurofeedback - Dataset 3	Behavioral and cognitive data from the same participants as in dataset EEG neurofeedback	New	Digital and physical	Observational data	.csv, .pdf	< 100 GB	Data of around 90 participants
Behavioral data classroom - Dataset 4	Behavioral and cognitive data from the same participants as in dataset EEG classroom	New	Digital and physical	Observational data	.csv, .pdf	< 100 GB	Data of around 80 participants

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

at the moment, we do not plan to reuse data sets

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- Yes, human subject data (Provide SMEC or EC approval number below)

SMEC application in progress
We will pseudonymize the collected data and will follow KU Leuven's GDPR code of using and processing personal data.

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- Yes (Provide PRET G-number or EC S-number below)

PRET application in progress
Personal data of the participants are name and date-of-birth. These are collected for ID purposes during data collection. It also includes contact information (e.g. email, phone number, ...) and signed informed consents. This information is only available to researchers involved in recruitment and data collection. The file linking the code and personal identifiers age/dob is only accessible to these researchers. It is stored in the personal OneDrive folder of the researchers. For the remainder of the study, all derivative data will be coded, and thus pseudonymized and stored in a different shared OneDrive folder. We will follow KU Leuven's GDPR code of using and processing personal data. Furthermore, an ethics application for SMEC will be submitted and registered via the PRET tool.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

No obvious path for commercialization

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

The protocol for the EEG recordings will be explained in a document.
There will be a readme file explaining the content and structure of the recorded EEG files, and how to run the scripts to prepare the files in Python and/or Matlab
For the behavioral data we will make a codebook documenting the study design, measures and variables that allows a secondary data analyst to use the data accurately and effectively. All test materials will be made available on the open science framework (OSF) account of Bert De Smedt ((https://osf.io/cmvdh/). This documentation includes per measure, how it was constructed and how performance indices were calculated. We will pre-register our data-analysis plan on OSF for each work package. It will be made available after publication, along with the respective dataset.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

Data Storage & Back-up during the Research Project

Where will the data be stored?

- OneDrive (KU Leuven)
- Personal network drive (I-drive)
- Large Volume Storage
- Other (specify below)

For processing, the data will be temporarily stored on a personal drive of the person executing the processing
To share the data between different researchers involved, we will use Onedrive (secured via KU Leuven login)
The data will also be stored on the database server of the research group STADIUS.
Some parts of the EEG data may be made publicly available, after consent from the participants. In this case, we will post it on a research data platform such as, e.g., Zenodo.org. The behavioral data will be made available on OSF when a study is published after completion of the pseudonymized data.
Questionnaires and informed consents will be separately archived in a locked room in the office of Bert De Smedt.

How will the data be backed up?

- Other (specify below)

The EEG data will also be stored on the database server of the research group STADIUS, which is backed up daily. The backup process is managed by the IT division of the department.
The behavioral data will be stored on the KU Leuven OneDrive. This data storage location has daily automatic back-up procedures.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data stored in the ESAT servers has access regulated by an access control list (ACL) that grants: read-write access to the project owner read-only access to specific users The ACL is managed by the project owner. Client computers can access the data using: SMB2 (or higher) from specific IP ranges NFSv4 from specific (IT managed) systems.
OneDrive is a secure cloud platform, with access control via the KU Leuven login.
The behavioral data will be pseudonymized by removing personal data and by storing this data separately form the research data on the personal OneDrive of Bert De Smedt and Davina Van den Broek. Multi-factor authentication is activated for the KU Leuven login of all researchers having access to the data.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

The (maintenance) costs are covered by the research divisions of the respective PIs

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- Large Volume Storage (longterm for large volumes)

After the period of 10 years if the data generated during the project is not longer in use locally, it will be removed from the servers. Note that some parts of the dataset may remain publicly available via public online repositories (see above)
Offline copies (questionnaires) and informed consents will be separately archived in a locked room for the expected 10 year period after the end of the project and will be destroyed after this 10 year.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The (maintenance) costs are covered by the research divisions of the respective PIs

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- Other (specify below)

The (anonymized) EEG data might be made available to the public for re-use, after obtaining consent from the participants in the study.
For the behavioral data the pseudonymized dataset of each study will be uploaded on the OSF account of Bert De Smedt (https://osf.io/cmvdh/) in a csv format upon publication of a study.

If access is restricted, please specify who will be able to access the data and under what conditions.

Only researchers that are officially involved in the project will have access to the EEG data (until the moment where we publish the data in a public platform).
The behavioral data will be pseudonymized and shared in a csv format on the OSF platform. It will be available to anyone provided that they give appropriate credit.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other data repository (specify below)

Unknown, probably Zenodo for the EEG data.
The pseudonumized behavioral dataset per study will be uploaded in a csv format at the OSF on the account of Bert De Smedt (https://osf.io/cmvdh/).

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- Other (specify below)
- CC-BY 4.0 (data)

Not decided yet for the EEG data.
For the behavioral data the pseudonymized dataset will be made available via the OSF account of Bert De Smedt under a Creative Commons Attribution license (CC-BY 4.0), so that users have to give credit to the original data creators.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

What are the expected costs for data sharing? How will these costs be covered?

no costs

Responsibilities

Who will manage data documentation and metadata during the research project?

The PIs, and the researchers working on the project

Who will manage data storage and backup during the research project?

The data storage and back up is the responsibility of the IT support of ESAT, and ICTS of KU Leuven

Who will manage data preservation and sharing?

The PIs of the project

Who will update and implement this DMP?

The PI, Alexander Bertrand, bears the end responsibility of updating & implementing this DMP.