Mock Behavioral and Eye-Tracking Project with Human participants

A Data Management Plan created using DMP Assistant

Creator: Gabriel Pelletier

Data Manager: Gabriel Pelletier

Affiliation: Montreal Neurological Institute

Funder: Digital Research Alliance of Canada

Template: Alliance Template

Project abstract:

This is a mock project to test the DMP assistant tool.
Mock Behavioral and Eye-Tracking Project with Human participants, involving open data sharing.

Start date: 01-10-2022

End date: 31-05-2023

Last modified: 09-09-2022

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Mock Behavioral and Eye-Tracking Project with Human participants

Data Collection

What types of data will you collect, create, link to, acquire and/or record?

This is a decision making experiment where participants make decisions between hypothetical political candidates characterized by multiple attributes laid out as text on the screen.
The following data will be acquired:

- Choices: the response key that was pressed on each trial to choose either the candidate on the left or right of the screen, along with the candidates that were presented on each trial and their attributes.
- Reaction times: The time elapse between the onset of the candidates on the screen and the response.
- Eye-tracking: The position of the gaze on screen at each timepoint during each choice.

What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?

Behavioral data (choices and reaction times) will be collected in CSV text files.
Eye-tracking data will be collected in the proprietary EyeLink format (.edf) and converted to Matlab format, which is also proprietary but more widely used.

What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized?

Data structure, file naming and meta-data will follow the BIDS specifications.

- The BIDS for behavioral data are specified here: https://bids-specification.readthedocs.io/en/stable/04-modality-specific-files/07-behavioral-experiments.html
- The BIDS extension for Eye-Tracking data were is not fully adopted yet but is in an advanced stage: https://docs.google.com/document/d/1eggzTCzSHG3AEKhtnEDbcdk-2avXN6I94X8aUPEBVsw/edit#

No versioning system is planned for the raw data.

Documentation and Metadata

What documentation will be needed for the data to be read and interpreted correctly in the future?

- README file containing Contributors and roles (who designed the experiment, who collected data, or analyzed data)
- Experimental procedure including instructions to participant, code to run the experimental task and stimuli.
- Data dictionary
- Metadata associated with all data files following BIDS specifications.

How will you make sure that documentation is created or captured consistently throughout your project?

Roles will be listed in the main project README files as to assign documentation duties to specific individuals who will be tasked to keep them updated.

If you are using a metadata standard and/or tools to document and describe your data, please list here.

Metadata for behavioral and eye-tracking data will follow BIDS specifications.

- Eye-tracking BIDS specification is still not officially released but is underway(working document). The folder and file organization and general metadata are already established and will be followed. Upon release of the full specification modifications might be done to the metadata.
- Behavioral BIDS specifications foundhere

Storage and Backup

What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.) and the length of time you will be storing it?

Raw Data:

- 40 participants
- Eye-tracking data: raw .edf about 50MB and pre-processed matlab files about 100KB
- Behavioral data about 5KB
- Total space needed for the raw data is about**2.5GB**

Analysis Derivatives

- Data derived from raw data, group summary statistics, figures, etc

- Should not exceed another **2.5GB**

How and where will your data be stored and backed up during your research project?

Sensitive Personal Identifying Information

- Names, phone numbers and email addresses will be collected as part of the recruitment procedure and on the Informed Consent form. Informed consent forms will be stored in a single paper copy in a locked cabinet in the locked laboratory space and the Montreal Neurological Institute.
- Names and phone numbers will be recorded on a single sheet of paper along the participant ID number. This sheet is the only link between the participant name and their data, and will be kept in a locked cabinet in the locked laboratory space during the study.
- All study data will be collected digitally and associated with the participant ID code. No sensitive information will be stored on any computer or cloud storage space.

Primary storage - During Data acquisition

- Data will be stored primarily on the data acquisition computer during the data acquisition phase.
- Data will be deleted from this computer upon completion of data collection, once it has been copied on the Backup 3 (see bellow).

Backup 1 and 2 - During data acquisition, data analysis and long-term

- During the data acquisition stage, data will be backup up every week on the cloud-based platform Microsoft Sharepoint*, on a Sharepoint site dedicated to this project accessible only to the PI, students and staff members directly involved.
- The Sharepoint site will be synchronized to at least one local computer in the lab space where all data analyses will be carried out.
- These count as 2 locations (remote/cloud + locally synced)

Backup 3 - Long term

- After data acquisition is completed, data will be backed-up on a Network Accessible Storage space accessible only to lab members.

Data Sharing - Long term

- Upon completion of the project at the time of publication in a peer-reviewed journal at the latest, data will be made publicly available on an online repository ensuring long-term perseveration of the raw data.

* Sharepoint is a secured cloud storage solution approved by McGill IT. Nevertheless, no personally identifying information will be stored in the cloud or on any computer.

How will the research team and other collaborators access, modify, and contribute data throughout the project?

Data collection/acquisition

- Data is collected at a single location on a unique computer

During analysis

- During analysis stages, raw data and derivatives will be stored on the cloud (Sharepoint) and automatically updated when any changes are done from the main analysis computer by the student/staff leading the project. If any collaborators need to have access to those data, the Sharepoint site can be shared easily with anyone with an internet connection through a controlled access mechanisms. Collaborators can makes changes which will be automatically visible to other team members.

Preservation

Where will you deposit your data for long-term preservation and access at the end of your research project?

- Raw data will be shared openly on the cloud. They will be deposited on the Federated Research Data Registry (FRDR), which guarantees long-term preservation.
- Experiment and Analysis scripts will be made openly available on Github and a copy will be made available through Zenodo which guarantees long-term preservation and persistent identifier.
- In addition, the raw data, along with the analysis scripts, experiment scripts, and derivatives underlying all the results of the publication(s) will be stored on the Network-Attached Storage installed locally in the lab.

Indicate how you will ensure your data is preservation ready. Consider preservation-friendly file formats, ensuring file integrity, anonymization and de-identification, inclusion of supporting documentation.

- Digital data files contain no personal identifying information so no de-identification is needed.
- Data will be preservation ready by formatting in BIDS data standard and enriched with BIDS compliant meta-data.
- README files and other documentation will accompany all major folders (experiment, analysis scripts, data)
- Experiment and analysis scripts will be thoroughly commented and reviewed by at least another lab member.

Sharing and Reuse

What data will you be sharing and in what form? (e.g. raw, processed, analyzed, final).

- Raw data will be shared openly as well as analysis scripts necessary to reproduced the results and figures from the raw data.

Have you considered what type of end-user license to include with your data?

Raw data will be shared under a CC-0 license.

What steps will be taken to help the research community know that your data exists?

- Data will be made publicly available on a national repository.
- Data will be associated with a persistent identifier and the dataset will be cited in each publication derived from it.
- We will advertise this freely available dataset and the associated publications on social media and on the laboratory website.

Responsibilities and Resources

Identify who will be responsible for managing this project's data during and after the project and the major data management tasks for which they will be responsible.

This is a relatively small project involving one graduate student as lead and a Research Associate as support staff.

- During the data collection and analysis, the graduate student will be responsible for the data management. The Research staff will be supporting and training the graduate student in how to achieve all the steps of this data management plan.
- After the project is over the graduate student will be responsible to submit the data to FRDR for data sharing.
- After the graduate student leaves the lab, the research staff will be responsible for any maintenance related to data preservation on the lab's infrastructure.

How will responsibilities for managing data activities be handled if substantive changes happen in the personnel overseeing the project's data, including a change of Principal Investigator?

- When the graduate student leaves, the research staff will be responsible for data management duties.
- If the research staff leaves, their replacement will be responsible for data management duties.
- If the PI leaves, the Unit will be responsible for data management duties.

What resources will you require to implement your data management plan? What do you estimate the overall cost for data management to be?

The lab with incur little to no extra costs for data management for this project

- Data deposition on FRDR is free for researchers at Canadian institutions
- Sharepoint is provided freely to the McGill community
- Dataset is small and only takes a small portion of our lab's NAS (estimated about 5$ based on the cost of 1Tb hard drive)
- No additional infrastructure or staff needed to format and analyze the data.

Ethics and Legal Compliance

If your research project includes sensitive data, how will you ensure that it is securely managed and accessible only to approved members of the project?

Sensitive data will is minimal and not critical to the dataset. They will be kept in paper files in a locked cabinet and only accessible to the research student and staff directly involved in the project for 7 years.

If applicable, what strategies will you undertake to address secondary uses of sensitive data?

- Participant will provide informed consent to share their data publicly without restriction (CC-0). No sensitive data will be shared.

How will you manage legal, ethical, and intellectual property issues?

- The research protocol and consent form were approved by the MUHC Neuro Ethics board and should thus be compliant with local legislation.

Created using DMP Assistant. Last modified September 09, 2022

4 of 4