

DMP title

Project Name REPRESENTATIONAL CHANGES UNDERLYING HUMAN LEARNING: A COMPUTATIONAL COGNITIVE NEUROSCIENCE PERSPECTIVE - DMP title

Project Identifier C14/21/047NAFWO

Grant Title C14/21/047NAFWO

Principal Investigator / Researcher Hans Op de Beeck

Project Data Contact hans.opdebeeck@kuleuven.be

Description The human mind has been trained intensively in a wide variety of complex information processing tasks. In addition, many individuals develop expertise in specific domains such as chess or microscopy. Previous studies have supported the hypothesis that expert learners do not develop super-powers, such as a larger working-memory capacity, but instead develop a more efficient information processing that is domain specific (see e.g. Bilalic, 2017). To understand the underlying representational changes at the neural level, we need to characterize neural representations in detail, which for a long time has not been possible due to a lack of suitable experimental and computational methods. The goal of this project is to investigate the neural basis of human learning through a state-of-the-art computational cognitive neuroscience approach. We will compare the unfolding of representational structure across brain space (functional magnetic resonance imaging combined with multi-voxel pattern analyses) and across time (representational dynamics as studied with multivariate electroencephalography) between human experts and non-experts, and between human and artificial expert systems (deep neural networks). Through this multidisciplinary approach we will obtain an unprecedented interdisciplinary understanding of how universal learning and domain-specific expertise alter the human mind and brain.

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Hans Op de Beeck

C1-C2 Project number & title

C14/21/047NAFWO

REPRESENTATIONAL CHANGES UNDERLYING HUMAN LEARNING: A COMPUTATIONAL COGNITIVE NEUROSCIENCE PERSPECTIVE

2. Data description

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

Type 1:

Type: Behavioural responses of participants to visual tasks on computer or using online platform.

Format: Mostly .mat (Matlab files) or .csv

Size: 1-10 GB

How created: Output of experimental scripts written in e.g. Matlab (PsychToolbox) or Python (PsychoPy)

Type 2:

Type: Magnetic Resonance Images (MRI) of the brain, Structural & Functional

Format: Mostly NIfTI or DICOM

Size: 300 GB - 1 TB (the latter including intermediate processing steps)

How created: MRI research scanner

Type 3:

Type: Electroencephalography (EEG)

Format: Depending on software, often .py and .mat

Size: 40 GB - 100 GB (the latter including intermediate processing steps)

How created: EEG equipment (electrodes attached to the skin)

Type 4:

Type: Computer simulation data

Format: Depending on software, often .py and .mat

Size: 500 GB

How created: By implementing artificial neural networks & training them to classify visual images (e.g. using Python & Tensorflow/PyTorch)

3. Ethical and legal issues

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

Yes, we use personal data. This is registered through our ethical approvals with EC and SMEC.

For all studies, we obtain the names, email address, personal address, and bank account information. This information is needed to pay participants.

This identifiable information is kept separate from the actual research data (see types 1-4 under Section 2.2). The research data are coded.

3.2. 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, add the reference to the formal approval by the relevant ethical review committee(s).

Yes, there are ethical considerations, and they are covered by several ethical approvals.

SMEC: G-2020-2910; G-2020-2379

EC: S62131

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

No

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

NA

4. Documentation and metadata

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

The raw data files for all four types are automatically stored with relevant meta-data.

For each experiment a detailed Methods section is written that allows to replicate the experiment, and re-analyse the obtained data.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

We use data acquisition and analysis software that is internationally used (e.g., fMRIPREP, SPM, cosmoMVPA toolbox), and where relevant the applicable standard data formats such as BIDS (which also standardizes directory structure & experimental information).

5. Data storage and backup during the C1-C2 project

5.1. Where will the data be stored?

The coded research data are stored on the professional KU Leuven Onedrive for Enterprises serves, using the drive of the main experimenter per experiment (up to 3 experimenters can be involved in this project). Copies can be made and kept on personal professional devices that fall under the university's secure environment. All people with access to these data use multi-factor authentication.

5.2. How will the data be backed up?

5.2. How will the data be backed up?

The Onedrive assures a storage using online cloud services. In addition, the coded research data might be backed-up on local external hard drives that are encrypted and password-protected.

5.3. Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

Yes, the expected size of the research data is smaller than the 2 TB per person provided through Onedrive.

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

No substantial costs expected, except the purchase of a few external hard drives.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The research data are always on password-protected devices. These research data are coded. The identifiable information (name, email, address) is not saved with the research data. The link between the identifiable information and the code of the research data is stored together with the informed consent form on paper (= not electronically). This paperwork is stored in KU Leuven offices within a locked closet.

6. Data preservation after the end of the C1-C2 project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

The coded research data will be preserved for 10 years.

The informed consents on paper will also be preserved for 10 years.

6.2. Where will these data be archived (= stored for the long term)?

Currently we archive data on a central storage RAID system of our research group when a lab member's contract ends, with additional backup on encrypted & password-protected external hard drives stored in a different building.

Towards the end of this project, we hope to move towards using the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

With current policies, we expect a total cost of around 5 000 euro, which can be covered from the C1 budget.

7. Data sharing and re-use

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

The research data can be shared (anonymously) with other researchers, also on online databases. This is explicitly mentioned in the informed consent forms signed by the participants.

7.2. Which data will be made available after the end of the project?

We share the final analyses files and further experimental material (stimuli etc.) using the OSF platform.

7.3. Where/how will the data be made available for reuse?

- In an Open Access repository

We use the Open Science Framework.

Partially also by mail, because not all the raw data files might be on OSF.

7.4. When will the data be made available?

- Upon publication of the research results

7.5. Who will be able to access the data and under what conditions?

Summary data & analysis files and experimental material will be fully open access for all that are registered on OSF.

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

OSF has no costs (at the moment).

8. Responsibilities

8.1. Who will be responsible for the data documentation & metadata?

The researchers hired on the project and the supervisor/promotor (the latter is the first contact point).

8.2. Who will be responsible for data storage & back up during the project?

The researchers hired on the project and the supervisor/promotor (the latter is the first contact point).

8.3. Who will be responsible for ensuring data preservation and sharing?

The supervisor/promotor.

8.4. Who bears the end responsibility for updating & implementing this DMP?

The end responsibility for updating and implementing the DMP is with the supervisor (promotor).