Data Management Plan

1. General Information	
Name applicant	**Hans Op de Beeck**
Project Number & Title	**METH/24/003** **THE DOMAIN-SPECIFIC AND DOMAIN-GENERAL COGNITIVE AND NEURAL PROCESSES UNDERLYING HUMAN SUCCESS AND FAILURE**
Affiliation	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other:
2. Data description	
Will you generate/collect new data and/or make use of existing data?	☒ Generate new data ☐ Reuse existing data

Describe the origin, type and format of the data (per dataset) and its (estimated) volume *If you **reuse** existing data, specify the **source** of these data.* *Distinguish data **types** (the kind of content) from data **formats** (the technical format).*	**Type 1:** Type: Behavioural responses of participants to visual or cognitive tasks on university computers / tablets (on or off campus), or using online platforms (e.g., Pavlovia, Prolific). Format: Mostly .mat (Matlab files) or .csv Size: 1-10 GB How created: Output of experimental scripts written in e.g. Matlab (PsychToolbox), Python (PsychoPy), Open Sesame, and javascript **Type 2:** Questionnaires to collect demographic data Size: max 1 GB **Type 3:** Type: Magnetic Resonance Images (MRI) of the brain, Structural & Functional Format: Mostly NIfTI or DICOM Size: 10 TB (the latter including intermediate processing steps) How created: MRI research scanner at UZ Leuven **Type 4:** Type: Electroencephalography (EEG) recordings Format: depending on EEG machine & analysis software Size: 1 TB How created: EEG recording devices (e.g. BioSemi) **Type 5:** Type: eye-tracking data Format: depending on tracking device & analysis software Size: 100 MB How created: eye-tracking devices (e.g. Tobii; EyeLink)

	Type 6: Type: Computer simulation data Format: Depending on software, often .py and .mat Size: 1 TB How created: By implementing artificial neural networks & training them to classify visual images (e.g. using Python & Tensorflow/PyTorch).

3. Ethical and legal issues	
Will you use personal data? If so, shortly describe the kind of personal data you will use AND add the reference to your file in your host institution's privacy register. *In case your host institution does not (yet) have a privacy register, a reference is not yet required of course; please add the reference once the privacy register is in place in your host institution.*	☒ Yes ☐ No Yes, we use personal data. This is registered through our ethical approvals with the EC Research UZ / KU Leuven and the SMEC, see our registration in CMT. For studies with paid participants, we obtain the names, email address, personal address, and bank account information (of their legal guardians, where relevant). This information is needed to pay participants. For studies in children, we will collect date-of-birth to calculate their age. This identifiable information is kept separate from the actual research data (see types 1-6 under Section 2.2). The research data are coded.
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 ☐ No Yes, there are ethical considerations, and they will be covered by ethical approval. We have existing ethical approvals for all data types mentioned above, in healthy participants and in clinical populations, and will require further approval when necessary.

Does your work 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?	☐ Yes ☒ No If yes, please comment:
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 are in place?	☐ Yes ☒ No If yes, please comment:

4. Documentation and metadata	
What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?	The raw data files for all six 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. It is impossible to detail these methods before the start of the project, given that many design and implementation choices will be made together with the to-be-hired junior or senior researchers.
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.	☐ Yes ☒ No Where applicable we use data acquisition and analysis software that is internationally used (e.g., for data type 3: fMRIPREP, SPM, cosmoMVPA toolbox), and the relevant standard data formats such as BIDS for data types 3 and 4 (which also standardizes directory structure & experimental information).

5. Data storage & backup during the Methusalem project

Where will the data be stored?	The coded research data are stored using the professional KU Leuven Onedrive for Enterprises, using the drive of the main experimenter per experiment as well as for the PI (Op de Beeck, Desender, Gillebert, De Smedt) that is supervising the experiment. Copies can be made and kept on personal professional devices that fall under the university' secure environment. All people with access to these data use multi-factor authentication.
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. These hard drives do not contain personal data that are easily identifiable (e.g. no participant names etc.). ***Note:*** *This is the current situation in the main applicant's team. The project will include support staff that will have data management as part of the job description, but this dedicated staff has not started yet. We expect major changes and improvements during the project. This note applies to all information given in this initial plan*.
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 ☐ No Yes, the expected size of the research data per individual researcher on the project is smaller than the 2 TB per person provided through Onedrive.
What are the expected costs for data storage and backup during the project? How will these costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of* ***the allocated project budget*** *to be used to cover the cost incurred.*	No substantial costs expected, except the purchase of a few external hard drives.
Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?	Password protection and multi-factor autenthication.

6. Data preservation after the end of the Methusalem project	
Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (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.
Where will these data be archived (= stored for the long term)?	Currently we archive coded research data in full on encrypted & password-protected external hard drives stored in two different rooms, with partial archiving in openly accessible platforms. We expect this to change in the first part of the project duration, once a more specific plan can be developed with the dedicated support staff.
What are the expected costs for data preservation during these 5 years? How will the costs be covered?	With current policies, we expect a total cost of around 10 000 euro, and probably a higher cost with future policies. This can be covered from the project budget.

7. Data sharing and reuse	
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)?	☒ Yes ☐ No 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.

Which data will be made available after the end of the project?	We typically share the final analyses files and further experimental material (stimuli etc.) using the OSF platform, which is an international standard frequently used in the domain of psychology (e.g., https://osf.io/qabrn/). Other options will be explored. For two recent datasets which we co-authored we used https://doi.gin.g-node.org/, which is very suited for imaging data (type 2).
Where/how will the data be made available for reuse?	☒ In an Open Access repository ☐ In a restricted access repository ☒ Upon request by mail ☐ Other (specify): See previous field.
When will the data be made available?	Upon publication of the research results, potentially already in the public preprint phase.
Who will be able to access the data and under what conditions?	Fully open access as much as possible (if allowed by ethical committee)
What are the expected costs for data sharing? How will these costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of* ***the allocated project budget*** *to be used to cover the cost incurred.*	The solutions that we currently use (OSF, g-node) have no costs (at the moment).

8. Responsibilities	
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).
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).
Who will be responsible for ensuring data preservation and sharing?	The supervisor/promotor.

Who bears the end responsibility for updating & implementing this DMP? *Default response: The PI bears the overall responsibility for updating & implementing this DMP*	The end responsibility for updating and implementing the DMP is with the supervisor (promotor).