FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	Lev Kiar Avberšek & https://orcid.org/0000-0002-3086-5166
Contributor name(s) (+ ORCID) & roles	Agnes Moors & main supervisor
Project number [1] & title	11PCY24N & Testing goal-directed explanations of perseverative and overly exploratory behavior in obsessive-compulsive disorder with behavioral and neuroscientific methods
Funder(s) GrantID [2]	
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	Obsessive-compulsive disorder (OCD) is a severe psychiatric disorder characterized by perseverative symptoms – obsessions and compulsions. Accumulating evidence suggests that behavior in OCD can also go into the opposite extreme and become overly exploratory. Explanations of behaviors in OCD are usually based on dual-process models, which differentiate between goal-directed and habitual behavior. Some authors suggest that perseverative behavior is habitual, but these explanations suffer from empirical and epistemic limitations. Other authors explain perseverative and overly exploratory behavior in terms of aberrant recruitment of goal-directed processes. These explanations lack direct empirical evidence and an integrative theoretical framework. To fill these gaps, we propose a goal-directed model that accounts for both types of behaviors in OCD and we test its assumptions. Specifically, we assume that two types of perseverative behavior (rigid and repetitive) and overly exploratory behavior result from the interaction between three deficiencies in a goal-directed cycle (rigidity in the action repertoire, aversion to uncertainty in response-outcome contingencies and to uncertainty in stimulus-goal discrepancies) and two environmental factors (deterministic vs. probabilistic task environment and un/certainty of outcome feedback). To examine our hypotheses, we use advanced behavioral paradigms and combine them with neuroimaging via electroencephalography (EEG) neuromodulation via transcranial magnetic stimulation (TMS). The data collected in this project will be used to answer the research questions of my FWO PhD project. It will consist of experimental behavioural data (participants make decisions based on observed data), self-report questionnaires about a variety of psychopathological symptoms, medical data (diagnosis, treatment, disorder onset), brain imaging data (EEG recordings) and neural intervention data (TMS).

2. 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 [3].

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*	*ONLY FOR PHYSICAL DATA*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
Predictive inference experiment	Computer experiment, in which participants have to make decisions to gain rewards in a volatile (probabilistic) environment. These data will be collected online using the Prolific platform.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☒ Model ☐ Software ☐ Other:	.csv	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Clinical questionnaires	Self-report questionnaires that assess symptoms of different mental	Generate new data	Digital	Numerical	.csv	< 1GB	

[3] Add rows for each dataset you want to describe.

	conditions, including anxiety, depression, and OCD. These data will be collected for the online experiments as well as for the lab experiments with clinical participants.						
Clinical data	Psychatric diagnosis, severity, onset, treatment type. These data will be used for the lab experiments with clinical participants.	Reuse existing data	Digital	Numerical	.csv	< 1GB	
EEG data	Brain activity recording during lab experiments with clinical participants.	Generate new data	Digital	Numerical, images	.edf, .png	< 100GB	
TMS data	Repetitive Transcranical Magnetic	Generate new data	Digital	Numerical, textual	.csv	< 1GB	

	Stimulation features (TMS device, coil type, stimulation parameters, stimulation protocol)						

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

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.	The clinical data will come from patients with OCD, treated at the University Hospital (UZ Leuven), handled by my co-supervisor Prof. Chris Bervoets.
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: Ethical approval will be applied for shortly and the approval number will be added. ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: We will apply for EC approval of TMS intervention in the future.

Will you process personal data[4]? 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) ☐ No Additional information: Clinical data contains information about participants' psychiatric condition, including diagnosis, disorder onset, and treatment. Clinical questionnaires collect information about participants' symptoms of different mental disorders. Privacy register number will be obtained in the future.
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.	☐ Yes ☒ No If yes, please comment:
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. Documentation and Metadata

[4] See Glossary Flemish Standard Data Management Plan

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). *RDM guidance on documentation and metadata*.	Each study will be dedicated a separate folder containing the following elements: 1. Readme.txt file: study overview, purpose and objectives, methodology, data description, data processing and cleaning. 2. Codebook.xlsx file: anonymized information about participants, variable list and data types, variable descriptions and units, summary statistics. 3. Analyses.R/.py/.m file: thoroughly commented statistical analysis code. 4. Experiment.py file: thoroughly commented experiment code. 5. All the study documents: Ethical application and approval, informed consent example, the instructions we gave participants. The PDF of all questionnaires (both baseline and ESM) will be included 6. Raw data In addition, all the experiments will be preregistered at Open Science Framework (OSF).
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. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: EEG data – European Data Format (.edf) allows us to store metadata. If no, please specify (where appropriate per dataset or data type) which metadata will be created:

4. Data Storage & Back-up during the Research Project

Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☒ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☒ OneDrive (KU Leuven) ☒ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☐ Other:
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)
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 ☐ NoThe size of data does not exceed KU Leuven and/or personal storage capabilities (<10GB)
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	All data will be pseudonymized with unique identifiers. The latter will be stored separately in encrypted files, which can be accessed only by the main researcher. The data will be continuously and automatically backed-up. The data will be shared with the involved researchers, who will need to provide identification to access them. The (internal and external) hard drives of the researchers' laptops will be encrypted using specialized software. Identifiers will be removed after 5 years.

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

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five 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...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☒ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specifiy):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	No expected costs.

6. 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. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☒ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify:
If access is restricted, please specify who will be able to access the data and under what conditions.	
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.	☒ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No If yes, please specify: Openly accessible data will be anonymous.
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☒ KU Leuven RDR ☒ Other data repository (specify): Open Science Framework ☐ Other (specify)

When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☒ GNU GPL-3.0 (code) ☐ Other (specify)
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	No expected costs.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Lev Kiar Avberšek

Who will manage data storage and backup during the research project?	Lev Kiar Avberšek
Who will manage data preservation and sharing?	Lev Kiar Avberšek
Who will update and implement this DMP?	Lev Kiar Avberšek