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	Laura Tibermont: 0000-0002-2609-7133
Contributor name(s) (+ ORCID) & roles	/
Project number [1] & title	KU Leuven Research Portal: 3M220478 Are you in tune with me? Understanding social functioning in autism via biobehavioral synchrony measurements
Funder(s) GrantID [2]	Flanders Research Foundation (FWO), Belgium: 11Q0R24N
Affiliation(s)	☒ KU Leuven - ROR identifier KU Leuven: 05f950310 ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other:
Please provide a short project description	During social interactions, behaviors and neurophysiological processes are bidirectionally synchronized between human interacting partners. In recent neuro-affective research, this phenomenon is referred to as biobehavioral synchrony and plays a crucial role by facilitating successful social exchanges. However, for individuals with autism spectrum disorder, who may experience challenges in social interaction and communication, this synchrony may present atypically. For my PhD project, I will investigate how behavioral patterns and neurophysiological processes of autistic versus neurotypical individuals synchronizes to the patterns and processes of their NT interaction partner, and vice versa. Moreover, I am interested to investigate these dynamics around the onset of real-life eye contract between the interacting partners. To measure this, I will apply dual multimodal biobehavioral measurements throughout a series of innovative real-life dyadic social interaction paradigms. This project constitutes a crucial step towards understanding the underlying mechanisms of social attunement and development, thereby providing new tools for the development of novel therapeutic strategies to help autistic individuals in social contexts.

[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.

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
Screening_log	During the first lab visit, participant's eligibility will be assessed based on inclusion and exclusion criteria. Eligibility will be recorded on the screening_log.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	1 large folder
Subject_log	Upon inclusion in the study, participant names will be Pseudonymize using codes. These will be recorded in the subject_log.	☒ Generate new data	☒ Physical	NA	NA	NA	1 large folder

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

Informed Consent	During the first lab visit, participant and their parents will receive a thorough explanation about the study. Following this discussion, the informed consent form will be discussed and signed.	☒ Generate new data	☒ Physical	NA	NA	NA	3 large folders
Demographics	Following inclusion, data regarding age, gender, IQ, concomitant medication or therapy, etc. will be collected.	☒ Generate new data	☒ Digital ☒ Physical	☒ Numerical ☒ Textual	☒ Electronic Case Report Form (eCRF) ☒ .csv ☒ .xlsx ☒ .pdf	☒ < 1 GB	NA
Questionnaires	During the first visit, I will apply the following questionnaires: - Social Responsiveness Scale 2 - Screen for Child Anxiety Related Disorders - Attachment Style Questionnaire - 3x3 Attachment questions	☒ Generate new data	☒ Digital ☒ Physical	☒ Numerical ☒ Textual	☒ eCRF ☒ .csv ☒ .xlsx ☒ .pdf	☒ < 1 GB	I will gather questionnaire data solely on paper if there would be an issue with the eCRF. 2 folders
ET_data	During the second lab visit, I will record eye tracking (ET) data from the two interacting partners simultaneously.	☒ Generate new data	☒ Digital	☒ Numerical ☒ Visual	☒ .csv ☒ .xlsx ☒ .pdf ☒ .txt ☒ .avi ☒ .png	☒ < 5 TB	NA
EEG_data	During the second lab visit, I will record encephalography (EEG) signals from the two interacting partners simultaneously.	☒ Generate new data	☒ Digital	☒ Numerical	☒ .bdf	☒ < 5 TB	NA

ECG_data	During the second lab visit, I will record electrocardiography (ECG) signals from the two interacting partners simultaneously.	☒ Generate new data	☒ Digital	☒ Numerical	☒ .bdf	☒ < 5 TB	NA
GSR_data	During the second lab visit, I will record Galvanic Skin Responses (GSR) from the two interacting partners simultaneously.	☒ Generate new data	☒ Digital	☒ Numerical	☒ .bdf	☒ < 5 TB	NA
Video_data	During the second lab visit (and as part of the ET_data), the interacting partners will be video recorded.	☒ Generate new data	☒ Digital	☒ Visual ☒ Audiovisual	☒ .avi ☒ .mp4	☒ > 5 TB	NA
Audio_data	During the second lab visit, the audio of the conversations between the interacting partners will be recorded.	☒ Generate new data	☒ Digital	☒ Sound	☒ .WAV	☒ < 5 TB	NA
OT_data	During the second lab visit, I will collect 3 saliva samples from the participant, for assessing endogenous oxytocin levels.	☒ Generate new data	☒ Physical ☒ Digital	☒ Numerical	☒ .csv ☒ .xlsx ☒ .pdf	☒ < 1 GB	2 freezer compartments NA
Worksheets	During the second lab visit, annotations regarding data collection will be made on a designated worksheet (on paper).	☒ Generate new data	☒ Physical	NA	NA	NA	2 folders
Visit_overview	Overview of completed sessions and collected data per participant.	☒ Generate new data	☒ Digital	☒ Textual	☒ .csv ☒ .xlsx ☒ .pdf	☒ < 1 GB	NA

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.	NA
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: s67699 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: Only information relevant for the project's research questions will be collected, as defined in section 2 "Research Data Summary". These (personal) data will be de-identified and pseudonymized, and will be collected, stored and processed in coded form.

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: The project involves the collection of a broad range of multimodal data, including participant characteristics (phenotypical and demographical information, questionnaires, clinical ratings, …), dual (neuro)physiology recordings (EEG, ECG, skin conductance), dual eye tracking data, video and audio recordings of child and experimenter during interactions, saliva samples, etc. Collected (personal) data will be de-identified and pseudonymized, and will be collected, stored and processed in coded form. Participant-identification information is only accessible by the researchers who need to know this information for the purpose of conducting the study. Personal data used for organizing the research (i.e. name, phone number, e-mail address) will not be included in the analysis and will be stored separately from the research data. Datasets: Screening_log, Subject_log, Demographics, Informed Consent G number: G-2023-6502 S number: S67699
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:

[4] See Glossary Flemish Standard Data Management Plan

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	
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*.	1. Study manual and Standard Operating Procedures (SOPs): All scientific operations that will be performed for collecting and processing data are documented in a study manual and/or study specific SOPs. 2. Readme files in each individual dataset folder: Every dataset is stored in a separate folder, which will be supplemented with a readme file. The first section of this file will contain the table of content. The following sections will include more information regarding that dataset in terms of how data was acquired, who was involved in the collection of the data (e.g. internship/master students), a short overview of the respective study experiment, etc. 3. Metadata labels in the KU Leuven ManGO platform: Raw (and analyzed) data will be stored on the KU Leuven ManGO platform (section 4 Data Storage & Back-up during the Research Project), which provides Metadata schema(s) for adding descriptive information about data objects and collections in the form of metadata. 5. (Pre-)processing and analysis documents: Manuals with the instructions to clean, (pre-)process and analyze the data will be written for every analyzed dataset. These documents will never contain sensitive (identifiable) participants information such as names, contact details, etc. 6. Markdown and comments in analysis scripts: Many scripts, such as RStudio and Python scripts, allow to add comments within the analysis script, in between the lines of code, allowing for clarifying/explaining in a step-by-step manner how the script operates.
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: If no, please specify (where appropriate per dataset or data type) which metadata will be created: I will not use any metadata standard, as there is no a priori intention to reuse the data that I will be collecting. Only in exceptional cases, such as severe illness for instance, the data could be further processed/reused by other members within the same research group as myself. In that case, that person(s) could rely on the types of documentation described above (1-6).

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: KU Leuven ManGO platform ☒ Other: Paper data will be stored in a locked room, next to the lab, in a storage cupboard or drawer, and can only be accessed by the researcher. Data collected on paper will be transferred to the eCRF (REDCap, a secure web-based application, running on the protected servers of the University). Entered data can be exported in .xlsx, .csv or .pdf format, in order to be stored on OneDrive (KU Leuven) or Large Volume Storage.
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): Digital data will be backed up on BIOMED L-Drive for large volume storage (secure KU Leuven facility for archiving) and/or the KU Leuven ManGO platform. ☐ 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 All data, except for the video (and audio) recordings, will stored on the KU Leuven One Drive and L-drive. The video data will be stored on the secured KU Leuven ManGO platform. In case storage space should not be sufficient, additional space will be purchased. KU Leuven One Drive offers 2 TB of storage space. The L-drive offers unlimited data storage in blocks of 5 TB, or 10 TB once storage has exceeded 100 TB. The KU Leuven ManGO platform offers 1 TB of storage for free with the option to expand storage against additional charges. At the moment we have 10 TB available. ☐ No If no, please specify:
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*	The personal nature of OneDrive ensures that files that are not explicitly shared, are not accessible to anyone else. As such, a separate folder will be created for each dataset separately. Only registered collaborators (i.e. colleagues withing the same research group, KU Leuven students, etc.) can be granted (reader only) access to this folder. The KU Leuven drives, including the L-drive, are incorporated within secured KU Leuven environments. Therefore, they are password-protected (including smartphone-based multi-factor identification) and are only accessible by registered KU Leuven researchers and/or students. Only the PI can request access to the network drive for study personnel.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The Department of Neurosciences provides our research group (Center for Developmental Psychiatry) with an L-drive. As such, costs will be covered by the department (€ 104,42/TB/year). The OneDrive storage comes free of charge for students and personnel of KU Leuven. At the moment, we have 10 TB of storage available on the KU Leuven ManGO platform (€35/TB/year). Any additional costs regarding storage and backup will be covered by research group-related funding.

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): Saliva samples will be used to analyze endogenous levels of oxytocin. Once analyzed, residual saliva will be destructed.
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: KU Leuven ManGo platform ☒ 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?	Large volume storage can be purchased in blocks of 5 TB. From 100 TB on, large volume storage can only be purchased in blocks of 10 TB. The cost per TB, per year is € 95.14. The KU Leuven ManGO platform offers 1 TB of storage for free with the option to expand storage against additional charges of €35 per TB per year. Given the maximum estimated used storages spaces will be as follows: - OneDrive: 2 TB: no cost - L-drive: - Mango: 8 TB: The expected cost will be € Any additional costs regarding storage and backup will be covered by research group-related funding.

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: Only published data (and associated scripts) can be made available in the form of publications or other dissemination of scientific work. In such case, all data will be pseudonymised when disseminated. More data can be made available or shared after permission of the responsible person (prof. Bart Boets). Non-published data will remain confidential until a final decision on publication of the data has been taken.
If access is restricted, please specify who will be able to access the data and under what conditions.	Data can potentially be (re)used by direct colleagues, after consultation and approval of the head of the research group (prof. dr. Bart Boets). External researchers will have to motivate why they want access to the data. When this data it is being used by other researchers, they are required to give credit to the original data creators.
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: I work with confidential data (e.g., name, sex, age, physiological data, etc.).

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): A specific repository will be chosen based on the publication strategy, as some journals request specific repositories. ☐ 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/DOI/accession number 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?	Given that most data repositories are free of charge, no costs are expected for data sharing.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PhD researcher (Laura Tibermont) will manage data documentation and metadata, under supervision of the head of the research group (Prof. dr. Bart Boets).
Who will manage data storage and backup during the research project?	Data storage and back up will be managed by the PhD researcher (Laura Tibermont), under supervision of the head of the research group (Prof. dr. Bart Boets) and with delegation to the dedicated data manager of the Psychiatry Research Group (dr. Martien Wampers).
Who will manage data preservation and sharing?	The head of the research group (Prof. dr. Bart Boets) will be responsible for ensuring data preservation and sharing, with delegation to the dedicated data manager of the Psychiatry Research Group (dr. Martien Wampers).
Who will update and implement this DMP?	The PhD researcher (Laura Tibermont) will update and implement this DMP, under supervision of the PI (Prof. Dr. Bart Boets).