

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	Bert Reynvoet (0000-0002-4898-2475)
Contributor name(s) (+ ORCID) & roles	Arnaud Szmalec (UC Louvain) (0000-0003-3903-3953)
Project number ¹ & title	THE ROLE OF STATISTICAL LEARNING IN THE DEVELOPMENT OF MATHEMATICS
Funder(s) GrantID ²	FWO G007224N
Affiliation(s)	<input checked="" type="checkbox"/> KU Leuven <input type="checkbox"/> Universiteit Antwerpen <input type="checkbox"/> Universiteit Gent <input type="checkbox"/> Universiteit Hasselt <input type="checkbox"/> Vrije Universiteit Brussel <input type="checkbox"/> 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	<p>Mathematics is a fundamental tool in life, however there is a lack of a clear scientific understanding about how math learning is precisely achieved. The central premise underlying this project is that mathematics is fundamentally a language that uses meaningful symbols which must be organized according to a finite set of rules in order to exchange ideas, concepts and theories among people. One of the most impactful insights that emerged from the psychology of language over the last decade, is that language acquisition seem to be largely acquired through a simple, robust and rapid learning mechanism, which is called statistical learning. Statistical learning (SL) refers to the human ability to extract regularities in how stimuli co-occur in the environment over space and time. The goal of this project is to make a case for the integration of SL and mathematics and to make a first, large-scale empirical attempt at investigating how sensitivity to distributional properties in the environmental input (i.e., SL) can subserve the development of mathematical skills. The present project presents a systematic inquiry into the importance of SL for mathematics learning, in particular mental arithmetic and geometry. We will set up developmental studies, compare typical and different atypical populations and use experimental training paradigms as well as neuroscientific techniques in children and adults to investigate the SL mechanisms underlying math skills.</p>
--	--

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.

The project consist of 3 WPs and 5 studies (WP1, WP2a, WP2b, WP3a, WP3b). For each study, the following materials and data will be collected and stored in a folder corresponding to the study.

Dataset Name	Description	New or Reused	Digital or Physical	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA
				Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
Personnel data	paper	New	digital	Spreadsheet (csv)	.csv	Estimated 100kb	
Math tests	(Standardized) math test via paper/pencil task	New	physical				2 tests per person (each test will consist of a couple of pages with different math exercises)
Software script (computer task)	Computer experiment with psychopy	New	Digital	Program code	Python code (psychopy)	Estimated 100kb	
Behavioral data: performance data	Data gathered by computer tasks	New	Digital	Numerical	.Csv	Estimated 400kb	
Analysis script	R script with analysis	New	Digital	Program code	.R	Estimated 400kb	
linked datafile	File including all pseudonymized info of the participant (math test + behavioral	new	digital	Spreadsheet (csv)	.csv	Estimated 100kb	

³ Add rows for each dataset you want to describe.

	performance on computer tests						
				ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA

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 be 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.	We do not use existing data
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.	<input checked="" type="checkbox"/> Yes, human subject data; provide SMEC or EC approval number: the personnel hired on the project will start in September/October 2024. SMEC application will be submitted once the study designs are finalized. For WP3b (TMS study) a separate application will be submitted <input type="checkbox"/> Yes, animal data; provide ECD reference number: <input type="checkbox"/> Yes, dual use; provide approval number: <input type="checkbox"/> No Additional information:

<p>Will you process personal data⁴? 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).</p>	<p><input checked="" type="checkbox"/> Yes (provide PRET G-number or EC S-number below): SMEC application will be submitted in fall 2024. The personal data will consist of name, dominant hand, age (years and months), name of school. In the pseudomized data, each participant and school will receive an ID number and we will remove their names. Because we collect data from minors, we will ask the parents/ guardians approval to participate throughout an active informed consent form. The child will give this document signed to the researcher before the experiment start. The child will also have to give approval to participate throughout an informed consent form.</p> <p><input type="checkbox"/> No</p> <p>Additional information:</p>
<p>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.</p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please comment:</p>
<p>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.</p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>
<p>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.</p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>

⁴ See Glossary Flemish Standard Data Management Plan

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.](#)

Data files will be accompanied by a .csv or .txt file describing all the information a third part requires in order to be able to find and use the data (i.e., variable list, measurement units, conditions under which the data are collected, scale information, etc.).

The readme file created per data file will include:

General information on project-level:

1. Title of the dataset
 2. Contact information of the PI
 3. Date of data collection
 4. Geographic location of data collection
 5. Keywords used to describe the data topic
 6. Language information
 7. Information about funding sources that supported the collection of the data
 8. The detailed experimental protocol will be described in the scientific papers (or preprints) generated based on this project.
 9. The ethical application will be saved as a PDF document and the approval code will be added to the project documentation.
- An empty informed consent form will be provided as a word file.

Project materials:

1. Details about the questionnaires
2. Details about the experiment program
3. Details about the stimulus set
4. A codebook will be created in R. In addition, information on the variables will be also provided in SPSS codebook or Readme file.

	<p>Information on research data-level:</p> <ol style="list-style-type: none"> 1. The data management plan will be provided as a pdf file. 2. The data preparation and statistical analyses will be documented in an annotated analysis code file (SPSS syntax and R code file). The version of the used software will be documented in in the scientific paper.
<p>Will a metadata standard be used to make it easier to find and reuse the data?</p> <p>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.</p> <p><i>REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.</i></p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:</p> <p>If no, please specify (where appropriate per dataset or data type) which metadata will be created:</p> <p>There is no metadata standard in our field.</p> <p>Project metadata of the numerical dataset in .csv or .txt will be created manually. For the three experiments, the metadata file will contain all information on the actual data to enable discoverability and reuse.</p> <p>Controlled vocabulary (e.g., APA Thesaurus of Psychological Index Terms, MeSH-terms) will be used for the keywords.</p>

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

<p>Where will the data be stored?</p> <p><i>Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.</i></p>	<p> <input type="checkbox"/> Shared network drive (J-drive) <input type="checkbox"/> Personal network drive (I-drive) <input checked="" type="checkbox"/> OneDrive (KU Leuven) <input type="checkbox"/> Sharepoint online <input type="checkbox"/> Sharepoint on-premis <input type="checkbox"/> Large Volume Storage <input type="checkbox"/> Digital Vault <input type="checkbox"/> Other: </p>
<p>How will the data be backed up?</p> <p><i>WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?</i></p>	<p> <input checked="" type="checkbox"/> Standard back-up provided by KU Leuven ICTS for my storage solution <input type="checkbox"/> Personal back-ups I make (specify) <input type="checkbox"/> Other (specify) </p>
<p>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.</p>	<p> <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No If no, please specify: </p>
<p>How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?</p> <p><i>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.</i></p> <p>Guidance on security for research data</p>	<p>Only the PhD student and their supervisor (Prof. Dr. Bert Reynvoet) will have direct access to the collected data.</p>

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	none
---	------

5. Data Preservation after the end of the Research Project	
<p>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...).</p> <p>Guidance on data preservation</p>	<p> <input checked="" type="checkbox"/> All data will be preserved for 10 years according to KU Leuven RDM policy <input type="checkbox"/> 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 <input type="checkbox"/> Certain data cannot be kept for 10 years (explain) </p>
<p>Where will these data be archived (stored and curated for the long-term)?</p> <p>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.</p>	<p> <input type="checkbox"/> KU Leuven RDR <input type="checkbox"/> Large Volume Storage (longterm for large volumes) <input checked="" type="checkbox"/> Shared network drive (J-drive) <input type="checkbox"/> Other (specify): </p> <p>Non-digital data and informed consent forms will be separately stored (during the research) and archived (after the research) in a locked room.</p> <p>All the digital data will be stored in a folder in the shared J-drive. Before the projects ends, the supervisor and principal investigator will have a meeting, where they discuss how the data is structured within these folder.</p>

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	No additional cost expected
--	-----------------------------

6. Data Sharing and Reuse	
<p>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.</p> <p><i>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/#INFOEU-REPO-ACCESSRIGHTS</i></p>	<p><input checked="" type="checkbox"/> Yes, as open data</p> <p><input type="checkbox"/> Yes, as embargoed data (temporary restriction)</p> <p><input type="checkbox"/> Yes, as restricted data (upon approval, or institutional access only)</p> <p><input type="checkbox"/> No (closed access)</p> <p><input type="checkbox"/> Other, please specify:</p>
<p>If access is restricted, please specify who will be able to access the data and under what conditions.</p>	

<p>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.</p>	<p> <input type="checkbox"/> Yes, privacy aspects <input type="checkbox"/> Yes, intellectual property rights <input type="checkbox"/> Yes, ethical aspects <input type="checkbox"/> Yes, aspects of dual use <input type="checkbox"/> Yes, other <input checked="" type="checkbox"/> No </p> <p>If yes, please specify:</p>
<p>Where will the data be made available? If already known, please provide a repository per dataset or data type.</p>	<p> <input checked="" type="checkbox"/> KU Leuven RDR <input type="checkbox"/> Other data repository (specify) <input type="checkbox"/> Other (specify) </p>
<p>When will the data be made available?</p>	<p> <input checked="" type="checkbox"/> Upon publication of research results <input type="checkbox"/> Specific date (specify) <input type="checkbox"/> Other (specify) </p>
<p>Which data usage licenses are you going to provide? If none, please explain why.</p> <p><i>A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENSE 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.</i></p> <p>Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.</p>	<p> <input checked="" type="checkbox"/> CC-BY 4.0 (data) <input type="checkbox"/> Data Transfer Agreement (restricted data) <input type="checkbox"/> MIT licence (code) <input type="checkbox"/> GNU GPL-3.0 (code) <input type="checkbox"/> Other (specify) </p>

<p>Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here.</p> <p><i>INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.</i></p>	<p><input checked="" type="checkbox"/> Yes, a PID will be added upon deposit in a data repository</p> <p><input type="checkbox"/> My dataset already has a PID</p> <p><input type="checkbox"/> No</p>
<p>What are the expected costs for data sharing? How will these costs be covered?</p>	<p>none</p>

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	PhD students are responsible for the data documentation during the period of the PhD. When there are questions, they will consult the PI. After the project, preservation of data is the responsibility of the supervisor
Who will manage data storage and backup during the research project?	PhD students are responsible for the data storage during the period of the PhD. When there are questions, they will consult the PI. After the project, preservation of data is the responsibility of the supervisor
Who will manage data preservation and sharing?	PhD students are responsible for the data preservation during the period of the PhD. When there are questions, they will consult the PI. After the project, preservation of data is the responsibility of the supervisor
Who will update and implement this DMP?	PhD student and supervisor

