FWO SENIOR RESEARCH PROJECT FUNDAMENTAL RESEARCH - G047023N- LEMMA - THE ROLE OF LANGUAGE IN EARLY MATHEMATICAL ACHIEVEMENT

A Data Management Plan created using DMPonline.be

Creators: Wim Van Dooren, Karen De Keersmaeker

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Karen De Keersmaeker

Data Manager: Karen De Keersmaeker

Grant number / URL: G047023N

ID: 198608

Start date: 01-01-2023

End date: 31-12-2026

Project abstract:
Mathematics has always been a central curricular domain in elementary and secondary education worldwide (De Corte, Greer, & Verschaffel, 1996). In the preschool years, the focus is mostly on very basic numerical abilities, such as reciting the counting words or identifying or comparing the numerosities of objects, and in the first years of primary school this is further systematized in the form of simple arithmetic. In recent years, however, mathematics education researchers have started to plea for also giving attention to more complex kinds of mathematics at the preschool level and in the first years of primary school (e.g., Bailey, Geary, & Siegler, 2014; Dede, 2010; English & Mulligan, 2013). They argue that even in the beginning years, mathematics education can and should already represent a broader coverage of the discipline, including early reasoning about mathematical relations, shapes, and patterns and structures (Clements & Sarama, 2013; Mulligan & Mitchelmore, 2009). Recent studies have also shown that children are already able to reason about more complex mathematical ideas at a younger age than traditionally assumed (Nunes, Bryant, Barros, & Sylva, 2012). By now, it is well known that language plays an important role in the learning of various school subjects, including mathematics (e.g., De Smedt & Boets, 2010; Lin, Peng, & Zeng, 2021). Mathematics as a school discipline uses a specific kind of vocabulary that children need to acquire. If more complex mathematics is addressed already in the early years, the role of such specific mathematical vocabulary may become even more important. In the current research project, we therefore want to investigate the role of vocabulary knowledge in such more advanced mathematical domains (and in addition to more foundational mathematical vocabulary) in young children's development of advanced mathematical abilities. We will focus on the abilities of mathematical patterning, proportional reasoning, and probabilistic reasoning. We will also investigate the mediating role of children's home environment (more specifically SES) in this relation. Finally, we will not only study the longitudinal effect of the mathematical vocabulary knowledge on children's later performance, but also vice versa the effect of performance on the later development of vocabulary. Before describing the research objectives and methodology of the project in detail, we provide a state of the art regarding the role of language (and mathematical vocabulary in particular) in mathematics learning, insights in the development of more advanced mathematical ideas in young learners, and the mediating role of SES in linking mathematical vocabulary to mathematical performance.

Last modified: 13-04-2023

FWO SENIOR RESEARCH PROJECT FUNDAMENTAL RESEARCH - G047023N- LEMMA - THE ROLE OF LANGUAGE IN EARLY MATHEMATICAL ACHIEVEMENT

Application DMP

QUESTIONNAIRE

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

We will be generating new data (see table below). Some of this data will be personal data (i.e. date of birth), which will be stored separately from the research data. Permission for the collection and analyzation of this personal data is provided by the Social and Societal Ethics Committee KU Leuven (G-2022-5497-R3)
Because we will collect different types of data on pupils at different times and we want to relate these data to each other, it will not be possible to work with completely anonymous data. However, the names of the participants (children) will not be used while processing the data and/or further steps in the research process. The principal investigator will pseudonymise the collected data; she will create a file in which the children's names will be linked to a unique identification number, consisting of the child's school (e.g. S1) and an individual number (e.g. P24). Only this unique identification number (and not the names of the participants) will be used during further processing and analysis of the data and all further steps in the research process. Moreover, both files, (1) file with names of participants linked to the unique identification number and (2) file with identification number linked to the data, will be kept in two different locations. After pseudonymisation, the blinding key (key to convert codes back into names) will be delivered to the data manager (i.e. Marc Vlecken) who will manage it and delete it after completion of the study. The analyses will be done using the pseudonymised file. Thus, from the moment the data are digitised, the researchers involved will only work on data files without names. In this way, all data collected will be kept confidential during each stage of the research process. The names and e-mail addresses of the principals and teachers will only be collected in function of communication about the research, i.e. making practical arrangements and communicating the results of the research. These names will therefore not be used/processed in the study itself.

Data	Study	Additional information
Grade	1, 2, and 3	
Date of birth	1, 2, and 3	This will only be used to determine the children's age. This data will thus not be reported in the research findings, therefore making it impossible to identify the participating children.
Gender	2 and 3	
Educational level mother	2 and 3	As a measurement of SES
Mathematical language abilities	1, 2, and 3	Interview study in study 1 Paper and pencil task in study 1, 2, and 3
General language abilities	2 and 3	Paper and pencil task
General mathematical abilities	2 and 3	Paper and pencil task

We will use the two data types: interviews and paper and pencil tasks. In a piloting phase, we will conduct and record interviews in which children have to explain their thinking, this to assure the validity of our test instrument. The recordings will be deleted as soon as possible (thus after deciding if we want to change certain items). All other data will be collected in either paper-and-pencil tasks or by asking the parents together with the informed consents. This data will be gathered in CSV files.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

We will collect physical data which will be transformed to digital data in order to analyse the data. All data will be stored on the storage facilities of the research unit. This data can only be accessed by the researchers of this project. The Data will be stored for 10 years in line with the KU Leuven RDM policies. During the research.
During this research project Karen De Keersmaeker will be responsible for the data preservation. During this term, the physical data will be stored in a locked cupboard, the digital data will be stored on a shared folder on OneDrive. This digital data will automatically back-upped since the data is saved on the shared drives of the research unit. Data will never be stored on a personal device.
After the research, Wim Van Dooren will be responsible for the data preservation. The physical data will be persevered in the archive of the research unit, whereas the digital data will be stored on OneDrive.
The costs for data storage during and for ten years after the project are covered by the project/by the faculty.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

We do not wish to deviate from the principle of preservation of data nor the minimum preservation term of 5 years.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Ethical issues have been considered. As stated above, we do use personal data, i.e. date of birth. We need this information to calculate the age of the participating children. This date of birth will be stored separately from the research data. For the data-analysis we will only work with the age of the children (in months), and will thus not report on this date of birth.
It is not possible to work with completely anonymous data because we will collect different types of data on pupils at different times and we want to relate these data to each other. Therefore, all data will be pseudonymised (as explained above).

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

FWO SENIOR RESEARCH PROJECT FUNDAMENTAL RESEARCH - G047023N- LEMMA - THE ROLE OF LANGUAGE IN EARLY MATHEMATICAL ACHIEVEMENT

DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

FWO SENIOR RESEARCH PROJECT FUNDAMENTAL RESEARCH - G047023N- LEMMA - THE ROLE OF LANGUAGE IN EARLY MATHEMATICAL ACHIEVEMENT

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

FWO SENIOR RESEARCH PROJECT FUNDAMENTAL RESEARCH - G047023N- LEMMA - THE ROLE OF LANGUAGE IN EARLY MATHEMATICAL ACHIEVEMENT

FWO DMP (Flemish Standard DMP)

1. 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.

All data will be physical, but transferred into a digital format.

Data	New or reused data	Data type	Data format	Digital data volume
Grade	New	Survey	.cvs	<100MB
Date of birth	New	Survey	.cvs	<100MB
Gender	New	Survey	.cvs	<100MB
Educational level mother	New	Survey	.cvs	<100MB
Mathematical language abilities	New	Experimental	.cvs audio data	<100MB
General language abilities	New	Experimental	.cvs	<100MB
General mathematical abilities	New	Experimental	.cvs	<100MB

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)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data

Ethical issues have been considered. This project has been approved by the Social and Societal Ethics Committee KU Leuven (G-2022-5497-R3).

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

We will collect information on the date of birth in order to calculate the age of the participating children. Information on the date of birth will be stored separately from the research data.

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.

- No

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 in the comment section to what data they relate and what restrictions are in place.

- No

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

2. 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).

Karen De Keersmaeker will provide the necessary documentation for each dataset, which gives insights in the used methodology, the analytical and procedural information, definitions and abbreviations of variables, etc. This will allow all researchers involved in this project to understand the various datasets.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

3. DATA STORAGE & BACK-UP DURING THE RESEARCH PROJECT

Where will the data be stored?

We will collect physical data which will be transformed to digital data in order to analyse the data. All data will be stored on the storage facilities of the research unit. This data can only be accessed by the researchers of this project. The Data will be stored for 10 years in line with the KU Leuven RDM policies. During the research.
During this research project Karen De Keersmaeker will be responsible for the data preservation. During this term, the physical data will be stored in a locked cupboard, the digital data will be stored on a shared folder on OneDrive. This digital data will automatically back-upped since the data is saved on the shared drives of the research unit. Data will never be stored on a personal device.
After the research, Wim Van Dooren will be responsible for the data preservation. The physical data will be persevered in the archive of the research unit, whereas the digital data will be stored on OneDrive.

How will the data be backed up?

The digital data will automatically be back-upped.

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

For the digital data, there is still 2039.55 GB of storage left. The physical data will be stored in a cupboard which also has sufficient storage.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Physical data will be stored in a locked cupboard. Digital data will be stored on a shared folder on OneDrive, which can only be assessed by the researchers involved in this project. This is a shielded drive, preventing the access or modification by unauthorized persons.

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

The costs for data storage during and for ten years after the project are covered by the project/by the faculty.

4. 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...).

Yes, all data will be retained for 10 years, following the KU Leuven RDM policy.

Where will these data be archived (stored and curated for the long-term)?

After the research, Wim Van Dooren will be responsible for the data preservation. The physical data will be persevered in the archive of the research unit, whereas the digital data will be stored on OneDrive.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The costs for data storage during and for ten years after the project are covered by the project/by the faculty.

5. DATA SHARING AND REUSE

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- No (closed access)

It is not allowed to reuse the collected data.

If access is restricted, please specify who will be able to access the data and under what conditions.

NA

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 in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

It is possible that the data will be made available (e.g. Open Science Framework).

When will the data be made available?

The pseudonymised data will be made available for reviewers of manuscripts during the revision phase, and to other researchers upon publication of the research results.

Which data usage licenses are you going to provide? If none, please explain why.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

We intend to add a persistent and unique identifier in order to identify and retrieve the data.

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

As the datasets are quite small, this is usually covered by the repository.

6. RESPONSIBILITIES

Who will manage data documentation and metadata during the research project?

Karen De Keersmaeker

Who will manage data storage and backup during the research project?

Karen De Keersmaeker

Who will manage data preservation and sharing?

Karen De Keersmaeker (during research), Wim Van Dooren (after research)

Who will update and implement this DMP?

Karen De Keersmaeker