Adaptive testing and learning environments accounting for domain-specific and domain-general skills to foster basic numerical skills

A Data Management Plan created using DMPonline.be

Creators: Flore Maricau, n.n. n.n.

Affiliation: KU Leuven (KUL)

Funder: Bijzonder Onderzoeksfonds

Template: KU Leuven BOF-IOF

Principal Investigator: Flore Maricau, n.n. n.n.

ID: 202213

Start date: 01-10-2023

End date: 01-10-2027

Project abstract:

For many school children, mathematics is a source of frustration because of the experienced difficulties with it. This led to increased research on the underlying cognitive processes of different mathematics skills and at the same time, research on personalized learning approaches to foster particular math skills. However, both research traditions are often too little geared to one another. In this project, we focus on numerosity and fraction comparison, both important precursors for math and affected by domain-specific and domain-general processes, that is, processes involved in many other – also non-mathematical – skills like inhibition. First, we will examine whether there are different profiles of children that excel or underperform on one or both (domain-specific and domain-general) dimensions. Second, we will construct adaptive diagnostical tests, that is a test in which the problems are adapted to the child in real time, taking into account both domain-specific and domain-general dimensions. Third, we will evaluate whether children learn better when they are presented with exercises matched to their current ability on both dimensions.

Last modified: 11-01-2024

Adaptive testing and learning environments accounting for domain-specific and domain-general skills to foster basic numerical skills

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.

Workpackage	Dataset name/ ID	Description	New or Reuse	Type of Data	File Format	Data Volume	How created
WP1 (investigating the role of inhibition)	Stimuli set: Numerosity and Fraction comparison	Stimuli set with all trials presented in the comparison task	New	Numerical	.csv	Estimated 50 kb	Mathlab and Python algorithm
	Questionnaire data	Collected via paper pencil task and digitalize later	New	Textual and Numerical	.txt .csv	Estimated 100 kb	General numerosity comparison task used in educational sentings
	Behavioral data: Logdata participants Numerosity and Fraction	Computer task programmed with Phyton/ PsychoPy	New	Logdata (numerical)	.csv	Estimated 400 kb	Logdata while solving computer tasks (PsychoPy)
	Behavioral data: performance dataset: Numerosity and fraction	All data combined of the computer tasks of all the participants	New	Numerical	.csv	Estimated 50 kb	Logdata while solving computer tasks (PsychoPy)
	Software script: numerosity and fraction task	PsychoPy script	New	Numerical	Python Source File	Estimated 50 kb	
	Statistical analyses: numerosity and fraction task	R-script and SPSS-script with analyses	New	Numerical	.R .spsv	Estimated 400 kb	
WP2 (development of an adaptive testing environment)	Stimuli set: Numerosity and Fraction comparison	Stimuli set with all trials presented in the comparison task	New	Numerical	.csv	Estimated 50 kb	Mathlab and Python algorithm
	Questionnaire data	Collected via (online) survey program	New	Textual and Numerical	.txt .csv	Estimated 100 kb	Self-reported data from likert-scale items (Math anxiety and motivation); validated questionnaires
	Behavioral data: Logdata participants Numerosity and Fraction	Computer task programmed with Phyton/ PsychoPy	New	Logdata Numerical	.csv	Estimated 400 kb	Logdata while solving computer tasks (PsychoPy)
	Behavioral data: performance dataset: Numerosity and fraction	All data combined of the computer tasks of all the participants	New	Numerical	.csv	Estimated 50 kb	Logdata while solving computer tasks (PsychoPy)
	Software script: numerosity and fraction task	PsychoPy script	New	Software	Python Source File	Estimated 50 kb	
	Statistical analyses: numerosity and fraction task	R-script and SPSS-script with analyses	New	Numerical	.R .spsv	Estimated 400 kb	

WP3 (development and evaluation of the adaptive learning environment)	Stimuli set: Numerosity and Fraction comparison	Stimuli set with all trials presented in the comparison task	New	Numerical	.csv	Estimated 50 kb	Mathlab and Python algorithm
	Questionnaire data	Collected via (online or paper pencil) survey program	New	Textual and Numerical	.txt .csv	Estimated 100 kb	Self-reported data from likert-scale items (Math anxiety and motivation); validated questionnaires
	Questionnaire data	Collected via (online or paper pencil) survey program	New	Textual and Numerical	.txt .csv	Estimated 100 kb	General preparatory math test, based on the curriculum tests for math
	Behavioral data: Logdata participants Numerosity and Fraction	Computer task programmed with Phyton/ PsychoPy	New	Logdata Numerical	.csv	Estimated 1 GB	Logdata while solving computer tasks (PsychoPy)
	Behavioral data: performance dataset: Numerosity and fraction	All data combined of the computer tasks of all the participants	New	Numerical	.csv	Estimated 400 kb	Logdata while solving computer tasks (PsychoPy)
	Software script: numerosity and fraction task	PsychoPy script	New	Software	Python Source File	Estimated 50 kb	
	Statistical analyses: numerosity and fraction task	R-script and SPSS-script with analyses	New	Numerical	.R .spsv	Estimated 400 kb	

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.

- Yes, human subject data (Provide SMEC or EC approval number below)

There are two main ethical applications regarding the two subjects of this research project:

- Numerosity comparison: G-2024-7545 (status: submitted)
- Fraction comparison: G-2023-7528 (status: submitted)

For each substudy within the project there will be an extra ethical approval code. We will add these when it will be available.
Study 1:

- Numerosity comparison: G-2023-7575 (status: submitted)
- Fraction comparison: G-2023-7528 (status: submitted)

Will you process personal data? 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)

As personal data, for each participant we will collect: name, sex, dominant hand, age (years and months), grade at school, and

name of the school.

In the pseudomized data, each participant and school will receive and 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 a informed consent form.

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 or 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

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keepdata 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).

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.

Information on research data-level:

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.

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.

- No

There is no metadata standard in our field.
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.
Controlled vocabulary (e.g., APA Thesaurus of Psychological Index Terms, MeSH-terms) will be used for the keywords.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Sharepoint online
- OneDrive (KU Leuven)

During the project al the digital data will be stored in and shared in a private team created in the Teams App. This will only be accesible for the principal investigators (Irene Oeo Morin and Flore Maricau), their supervisors (Prof. Dr. Bert Reynvoet and Prof. Dr. Fien Depaepe) and the team and research manager (Ine Windey and Frederik Cornillie).
All the physical data (the inform consent forms) with personal data of the participants will be stored in a locker in the principal investigators office. Only the two principal investigators have a key of this office.
The two principal investigators will use a shared One Drive folder to facilitate the collaboration.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

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

Only the principal investigators (Irene Oeo Morin and Flore Maricau) and their supervisors (Prof. Dr. Bert Reynvoet and Prof. Dr. Fien Depaepe) will have direct access to the collected data, since they will be stored in the folder within the Teams channel.

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

None.

Data Preservation after the end of the Research Project

Which data will be retained for 10 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...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

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

- Shared network drive (J-drive)

Non-digital data and informed consent forms will be seperately stored (during the research) and archived (after the research) in a locked room.
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.

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

There is no additional cost for data storage.

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.

- Yes, as open data

The datasets with pseudonomized data will be the open data.

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

The data is intended to be open access.

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.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- KU Leuven RDR (Research Data Repository)

Only the researchers working at the research centers to which the PI and collaborator belong can have access to the anonymized data upon publication of the project results and after the project.
Only the PI and collaborator have access to the pseudonymized data.
In case international researchers are willing to use the anonymized data we can provide access to this anonymized data set through Belnet file sender. The data will be saved in open formats (e.g., csv or txt) in view of fostering reuse on a longer term.

When will the data be made available?

- Upon publication of research results

The pseudonomized dataset used for particular publications will be made available to other researchers (cf. supra) upon publication of the research results. Furthermore, all pseudonomized datasets of the project will be made available through the protected university network system to other researchers immediately after the end of the project.

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

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

None

Responsibilities

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

The principal investigators (Irene Oeo Morin and Flore Maricau) are responsible for the data management during the time of their PhD research. The supervisors and IT staff will support the principal investigators in this task. After the research project ends,

preservation of data is the responsibility of the supervisors.

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

The principal investigators (Irene Oeo Morin and Flore Maricau) are responsible for the data management during the time of their PhD research. The supervisors and IT staff will support the principal investigators in this task. After the research project ends, preservation of data is the responsibility of the supervisors.

Who will manage data preservation and sharing?

The principal investigators (Irene Oeo Morin and Flore Maricau) are responsible for the data management during the time of their PhD research. The supervisors and IT staff will support the principal investigators in this task. After the research project ends, preservation of data is the responsibility of the supervisors.

Who will update and implement this DMP?

The principal investigators (Irene Oeo Morin and Flore Maricau) and the supervisors of the project (Fien Depaepe and Bert Reynvoet)