Plan Overview

A Data Management Plan created using DMPonline.be

Title: Spatial skills and mathematics, towards a causal understanding of mechanisms

Creator: Ilse Coolen

Affiliation: KU Leuven (KUL)

Funder: European Commission (Horizon)

Template: Horizon Europe DMP +

Project abstract:
Difficulties in learning maths usually begin at the start of education and are robust over time. The number of students not reaching basic levels of maths achievement remains a major societal challenge, at a time when societies are increasingly demanding maths skills due to scientific and industrial progress. Research examining the key cognitive foundations for successful maths learning has identified spatial skills as a promising potential contributor to maths development. Nevertheless, how and why SPatial skills contribute to mATHS learning remains elusive. SPATHS will therefore develop a new framework that aim to understand the causal mechanisms via which spatial skills impact maths learning. Various spatial skills and maths abilities will be tested and causal pathways between the different abilities will be estimated in children at the start of formal school (final year of preschool and first year of primary school). SPATHS will unravel long-awaited knowledge about the mechanism underlying the spatial-maths association. This is essential for future research into efficient interventions to improve maths and to make spatial training programs more effective.

ID: 211211

Start date: 01-10-2024

End date: 30-09-2026

Last modified: 13-03-2025

DPIA

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

* Not applicable

Created using DMPonline.be. Last modified 13 March 2025

2 of 12

Spatial skills and mathematics, towards a causal understanding of mechanisms
Full DMP

Version information

Action number

101150769

Action acronym

SPATHS

Action title

Spatial skills and mathematics, towards a causal understanding of mechanisms

DMP version number

v1.0.

Date

03.03.2025

1. Data summary

1.1 Will you re-use any existing data and what will you re-use it for?

NA

1.2 What types and formats of data and other research outputs will the project generate or re-use?

The SPATHS study will collect behavioural data from 360 5-7-year-old-children in Belgium (180) and England (180). We will collect data using a numerical and spatial task battery on paper or with physical objects (e.g., blocks). This battery will measure four early mathematical skills (i.e., calculations, proportional reasoning, ordering, place-value) and four spatial tasks (visuospatial memory, visuospatial attention, mental rotation, and, mental transformations). Each task will yield a mean performance (based on accuracy), sometimes accompanied by time scores.
Additional measures will include the responses from a short parent questionnaire, including sex, age, mother tongue, parent education level.
Data will be collected on paper and digitalised in csv-files. The csv-file will be for all participants separated per country, hence the total data size for both csv-files is unlikely to exceed 100 KB.

1.3 What is the purpose of the data generation or re-use and its relation to the objectives of the project?

The overarching goal is to develop and test a conceptual framework to understand the causal mechanisms that associate spatial and maths skills. The empirical evidence of this project will further refine and extend theories on the association between spatial skills and maths in mathematical cognition and it will enable the design and evaluation of effective interventions to improve maths at the start of education.
There are three actionable objectives:

Objective 1 – Mechanisms: I aim to understand the underlying mechanisms of the spatial-maths association.
Objective 2 – Causality: I aim to test, via observational data and using advanced statistical techniques, whether there is a causal association between different spatial and maths skills and whether or not causality runs in two directions.
Objective 3 – Potential moderators: I aim to understand how the association between spatial skills and maths is moderated by age, sex, differences in the educational environment, and socio-economic status (SES), to find potential foundations for inequalities in maths.

To address these objectives, we will measure various maths measures and spatial measures that have demonstrated or hypothesised to be relate to eachother (correlational in observational data or causational in training studies). The association of each unique spatial measure to each separate maths measure will be tested in causal modelling to provide a detailed framework of which spatial skills associate to which maths skills.
Data that has been identified in previous literature to be moderators between spatial skills and maths skills will be collected as well, including age, sex, mother tongue, year group in education, and SES.

1.4 What is the expected size of the data that you intend to generate or re-use?

Data will be collected on paper and digitalised in csv-files. The csv-file will be for all participants separated per country, hence the total data size for both csv-files is unlikely to exceed 100 KB.

1.5 What is the origin/provenance of the data, either generated or re-used?

We will collect data using a task battery on paper or with physical objects (e.g., blocks).
Additional measures will include the responses from a short parent questionnaire, including sex, age, mother tongue, parent education level.
Data will be digitalised in csv-files and analysed through R, generating an R script and an R-Markdown analyses output.

1.6 To whom might your data be useful ('data utility'), outside your project?

The data will be suitable for sharing with bona fide researchers and will be available after locking the analytical databases. The study will generate data useful for any bona fide researchers examining spatial and maths skills as well as researchers interested in designing more effective interventions or training studies to improve maths based on spatial skills at the start of education.

2.1 FAIR data: Making data findable, including provisions for metadata

2.1.1 Will data and other research outputs be identified by a persistent identifier?

- Yes: describe below

This project will generate two separate datasets: one collected in England and one in Belgium. Due to data security regulations and ethical considerations, these datasets will be stored and managed separately. Each dataset will be handled in accordance with the respective country's data protection policies (e.g., UK GDPR for England and EU GDPR for Belgium). Consequently, dataset sharing will also be conducted separately, ensuring compliance with local regulations and ethical requirements.
Data from Belgium
The data will be shared with bona fide researchers and will be available after locking the analytical database. The pseudonymised research data along with the codebook, will be deposited in the KU Leuven RDR (Research Data Repository). Each published datasets receives a DOI.
Data from England
The data will be shared with bona fide researchers and will be available after locking the analytical database, in collaboration with the repository UK Data Service, and the Open Science Framework. The dataset will have a DOI.
All output from the dataset will use the DOIs from the datasets, the authors' ORCID and the grant identification information will be provided.

2.1.2 Will rich metadata be provided to allow discovery?

What metadata will be created?
What disciplinary or general standards will be followed?
In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Data from Belgium

The pseudonymised research data along with the codebook, will be deposited in the KU Leuven RDR (Research Data Repository). Each published datasets will be accompanied by metadata (e.g., instruments of data collection, codebook, protocols) available via the User Interface. For a full overview of the potential metadata on the KU Leuven RDR, see the [full metadata model (chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.kuleuven.be/rdm/en/rdr/full-metadata-model](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.kuleuven.be/rdm/en/rdr/full-metadata-model)). Instructions on how to access study data, will be articulated on the project's website.

Data from England

We will make discovery metadata (describing the data) publicly available in collaboration with UK Data Service. The UK Data Service uses DDI to structure catalogue records, which records contain mandatory and optional metadata elements on the following:

- Study description – information about the context of the data collection, such as bibliographic citation of the study and data, the scope of the study (topics, geography, time), methodology of data collection, sampling and processing, data access information, and information on accompanying materials.

- Data file description – information on data format, file type, file structure, missing data, weighting variables and software.

- Variable descriptions.

A summary of the data use will be published on the project's website for transparency and to avoid duplication of access requests. Information about the data, including metadata will be available on the project website. Instructions on how to access study data along with the codebook, will be articulated on the project's website.

Datasets and accompanying metadata will be as identical as possible between the data collected in Belgium and England.

2.1.3 Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

- Yes: describe below

Data in Belgium

The KU Leuven RDR makes the addition of keywords manadatory as metadata. The interface can be used for building keyword indexes and for classification and retrieval purposes, and allows for the choice of *Controlled vocabulary names* for the specification of the keyword controlled vocabulary in use ([LCSH](#), [AAT](#), [MeSH](#),...).

Data in England

The UK data service uses the DDI as a rich and detailed metadata standard originally designed for describing social, behavioural and economic sciences data. It is used by most social science data archives in the world. The UK data service assigns keywords from their own HASSET thesaurus.

The use of keywords will be kept as constant as possible between the dataset from Belgium and England.

2.1.4 Will metadata be offered in such a way that it can be harvested and indexed?

- Yes: describe below

Data from Belgium

The KU Leuven RDR can be used for building keyword indexes and for classification and retrieval purposes. The RDR repository is a trusted data repository that ensure metadata fields and digital object identifiers (DOIs) to uploaded data. In such way, research data will be archived and indexed in a controlled manner, and thus, easy to discover.

Data from England

The UK Data Service uses DDI to structure catalogue records. The use of standardised records in eXtensible Mark-up Language (XML) brings key data documentation together into a single document, creating rich and structured content about the data. Metadata can be harvested for data sharing through the [Open Archives Initiative Protocol for Metadata Harvesting](#) (OAI-PMH).

2.2 FAIR data: Making data accessible

2.2.1 Will the data and other research outputs be deposited in a trusted repository?

- Yes: describe below

This project will generate two separate datasets: one collected in England and one in Belgium. Due to data security regulations and ethical considerations, these datasets will be stored and managed separately. Each dataset will be handled in accordance with the respective country's data protection policies (e.g., UK GDPR for England and EU GDPR for Belgium). Consequently, dataset sharing will also be conducted separately, ensuring compliance with local regulations and ethical requirements.
Data from Belgium

The data will be shared with bona fide researchers and will be available after locking the analytical database. The pseudonymised research data along with the codebook, will be deposited in the KU Leuven RDR (Research Data Repository). Instructions on how to access study data, will be articulated on the project's website.
Data from England
The data will be shared with bona fide researchers and will be available after locking the analytical database, in collaboration with the repository UK Data Service, and the Open Science Framework. A timestamped static database will be deposited with UK Data Service at the end of the grant period. Instructions on how to access study data along with the codebook, will be articulated on the project's website. The Investigator (Dr. Coolen) will be responsible for approving the sharing of research data to interested parties. Data requests will be submitted, via an application form (hosted on the website) and reviewed by the investigator.
A summary of the data use will be published on the project's website for transparency and to avoid duplication of access requests.

2.2.2 Have you explored appropriate arrangements with the identified repository where your data and other research outputs will be deposited?

- Yes

The selected repositories are appropriate and have the capacity to host all research data and outputs.

2.2.3 Does the repository ensure that the data and other research outputs are assigned an identifier? Will the repository resolve the identifier to a digital object?

Each published datasets receives a DOI.
All output from the dataset will use the DOIs from the datasets, the authors' ORCID and the grant identification information will be provided.

2.2.4 Will all data and other research outputs be made openly available?

- No, certain datasets cannot be shared openly for the following reasons:

Bona fide researchers will contact the study team to request access to the data. Instructions on how to access study data will be articulated on the project's website. The Investigator (Dr. Coolen) will be responsible for approving the sharing of research data to interested parties. Data requests will be submitted, via an application form (hosted on the website) and reviewed by the investigator.
A summary of the data use will be published on the project's website for transparency and to avoid duplication of access requests.
Restricted access is appropriate given that data contains certain personal information including age and gender for a volnurable population (children).

2.2.5 Is an embargo applied to give time to publish or seek protection of the intellectual property (e.g. patents)?

- No

2.2.6 If an embargo is applied (see question 2.2.5), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

NA

2.2.7 Will the data and other research outputs be accessible through a free and standardized access protocol?

Data in Belgium

Data accessibility is ensured by the data management policy of the RDR repository.
Data in England
Data accessbility is ensured by the UK data service.
Additionally, instructions on how to access study data will be articulated on the project's website.

2.2.8 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

Data will be shared on the KU Leuven RDR and the UK data service with restricted access. Bona fide researchers can request access through the repositories.
Access to data can also be requested by contacting the project data responsible (Dr Ilse Coolen) via the project's website.

2.2.9 How will the identity of the person accessing the data be ascertained?

Data in Belgium
The identification of the person accessing the data is managed directly by the KU Leuven RDR platform. If the access grant applicant doesn't have a KU Leuven login number, they have to log in with an ORCID ID. If the access grant applicant doesn't have an ORCID ID or a KU Leuven login number, they have to use the contact button to explain and send their request.

Data in England
The UK Data Service verifies user identity primarily through institutional login credentials from UK academic institutions. Users affiliated with UK universities and research institutions authenticate via their institutional accounts. Non-UK users or those without an institutional login must register for a UK Data Service account, providing an email address and verifying their identity through a registration process. Once registered, they log in using their UK Data Archive credentials.
An additional data access request will be put into place.

2.2.10 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

- No

The Investigator (Dr. Coolen) will be responsible for approving the sharing of research data to interested parties. Data requests will be submitted through the data repositories (KU Leuven RDR or UK data service), or via an application form (hosted on the project's website) and reviewed by the investigator.

2.2.11 Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why.

- Yes

2.2.12 Will metadata contain information to enable the user to access the data?

- Yes

2.2.13 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

In light of the KU Leuven research data management [policy](#) expecting that relevant research data generated are retained for a period of minimally 10 years after the end of the project, KU Leuven RDR ensures storage for at least 10 years.
The UK data service does not have a time limit and ensures long-term data accessibility.

2.2.14 Will documentation or reference about any software needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

Details about software will be documented in study preregistrations/protocols deposited on the OSF page, as well as in the Materials and Methods sections of open access preprints and publications. Analysis code will be provided in R scripts; readMe files will be provided if necessary to provide instruction on how to use the scripts in the appropriate software.

2.3 FAIR data: Making data interoperable

2.3.1
What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines?
Will you follow community-endorsed interoperability best practices? Which ones?

The UK Data Service employs the **Data Documentation Initiative (DDI)** standard, which is widely adopted in the social sciences for documenting and managing data. This standard enhances the richness and structure of metadata, facilitating effective data sharing and reuse.
The KU Leuven RDR allows for the use of the DDI standard, which will be employed to align with the UK data service.
The data will be in the same format for both the data in Belgium as England and will be in csv format and will be prepared and documenten using R Markdown, which allows for reproducible research by integrating code, analysis, and narrative text. This aligns with common practices in the social sciences to ensure compatibility and ease of use.

2.3.2 In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies:
Will you provide mappings to more commonly used ontologies?
Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

If and when specific vocabularies will be generated during the project, these will be openly available.

2.3.3 Will your data and other research outputs include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

- Yes

When applicable cross-references will be reported.

2.4 FAIR data: Increase data re-use

2.4.1 How will you provide documentation needed to validate data analysis and facilitate data re-use?

Details about data and data analysis will be documented in study preregistrations/protocols deposited on the OSF page, as well as in the Materials and Methods sections deposited on the OSF page and of open access preprints and publications. Analysis code will be provided in R Markdown scripts, which allows for reproducible research by integrating code, analysis, and narrative text. readMe files will be provided if necessary to provide further instruction on how to use the scripts in the appropriate software.

2.4.2
Will your data and other research outputs be made freely available in the public domain to permit the widest re-use possible?
Will your data and other research outputs be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

The metadata will be openly available under a CC0 (Public Domain Dedication) license, ensuring that dataset descriptions can be freely accessed, indexed, and reused.
For the datasets themselves:

- The data will be openly available whenever possible but subject to controlled access due to sensitive variables (e.g., gender, age).
- Access will be granted upon request, in line with ethical and legal requirements.
- To ensure responsible reuse, the data will be licensed under:
  - CC-BY (Attribution) on the KU Leuven RDR, allowing reuse with attribution.
  - The UK Data Service End User License (EUL), which requires applicants to agree to responsible data use.

Preprints will be deposited in PsyArXiv under a CC-BY license, ensuring early open access to research findings. Peer-reviewed publications will be made openly available in accordance with the funder's open-access policy and will be deposited in Lirias, KU Leuven's institutional repository, under a CC-BY license to ensure compliance with open-access requirements and long-term availability.

2.4.3 Will the data and other research output produced in the project be useable by third parties, in particular after the end of the project?

- Yes

Yes, data will be available for use by bona fide researchers. A summary of the data use (based on the data requests) will be published on the project's website for transparency and to avoid duplication of access requests.
Open-access tasks created and validated through this research will be made openly available with the required documentation for reuse on the OSF page and the project's website.

2.4.4 Will the provenance of the data and other research outputs be thoroughly documented using the appropriate standards?

- Yes

The appropriate standards will be followed.

2.4.5 Describe all relevant data quality assurance processes.

The quality of data will be ensured during raw data generation, data processing, and analysis. To do so, well-established data processing methods and protocols will be adopted which are in agreement with common requirements in our scientific discipline. These procedures will be well documented in the Methods and Methodologies section of each open-access publication.
Access to data will be free but modification will not be possible by others to ensure data quality and security when sharing data.

- **Standardised Data Collection** – Data will be collected following a predefined protocol to minimise errors and inconsistencies. Where applicable, validated instruments and surveys will be used.
- **Data Cleaning and Validation** – Raw data will be checked for inconsistencies, missing values, and outliers. Automated scripts in R Markdown will be used to document data cleaning steps, ensuring transparency and reproducibility.

- **Version Control** – Data files will be managed using version control systems in OneDrive with frequent back-ups to track changes and prevent data loss or unintended modifications.
- **Metadata and Documentation** – Comprehensive metadata following repository-specific standards (DDI metadata) will be created to facilitate accurate interpretation and reuse of the data.
- **Storage and Backup** – Data will be securely stored in KU Leuven RDR and the UK Data Service, ensuring long-term preservation.
- **Peer Review and Internal Checks** – Before data deposition, internal quality checks will be conducted, including peer verification of data integrity and reproducibility of analyses.

3. Other research outputs

3.1 Do you have any additional information, that was not addressed in the previous sections, which you wish to provide regarding other research outputs that are generated or re-used throughout the project?

Preprints will be deposited in PsyArXiv under a CC-BY license, ensuring early open access to research findings.

Peer-reviewed publications will be made openly available in accordance with the funder's open-access policy and will be deposited in Lirias, KU Leuven's institutional repository, under a CC-BY license to ensure compliance with open-access requirements and long-term availability.

All original material, including testing protocols, created for the purpose of this project will be shared and made accessable through the OSF page and the project's website, except where restricted by third-party licences, for example, standardised tests that require a paid licence will not be made publicly available.

Analysis code accompanying publications or data curation procedures will be provided in R Markdown scripts, which allows for reproducible research by integrating code, analysis, and narrative text and deposited on the OSF page.

4. Allocation of resources

4.1 What will the costs be for making data and other research outputs FAIR in your project?

There are no expected costs to make data and research output FAIR, as adopted repositories (KU Leuven RDR, UK data service, OSF, and Liras) and in-house storage servers are free of charge.

4.2 How will these be covered?

NA

4.3 Who will be responsible for data management in your project?

The research fellow, Dr. Ilse Coolen, is responsible for the data management during the duration of the project and for the data requests to access data through the KU Leuven RDR and the UK data service. In addition, the supervisors of the project, Prof. Bert De Smedt and Prof. Gaia Scerif will ensure the long-term storage and preservation of the data and outputs of this project in their in-house computer clusters.

4.4 How will long term preservation be ensured?

Data in Belgium

Data generated throughout the research project will be retained in the KU Leuven RDR repository for at least 10 years in

accordance with the FAIR principle and KU Leuven policies.
Data in England
Data generated will be deposited in the UK data service indefinitely.

5. Data security

5.1 What provisions are or will be in place for data security?

Data in Belgium
Digital data will only be stored on the network drives of the KU Leuven. Offline data will be stored (in a closed drawer) in the KU Leuven Office of Prof. Bert De Smedt. Data shared open-access will be pseudonymous and non-identifiable. Restricted access will be given upon request to the data that contains certain personal information including age and gender as it concerns a volnurable population (children).
KU Leuven has IT specifications for data storage and management, including standard backups in secure locations. The IT department provides tailored solutions to ensure that data is securely stored, and cannot be altered by an unauthorised entity.
Read, write, and execute access privileges to data are protected by password and antivirus software managed directly by the KU Leuven's central Active Directory Domain called LUNA-SET (https://set.kuleuven.be/set-it/docs/luna-set), the central authentication and authorization system releasing rights and security settings in the KU Leuven network.
Data in England
Digital data will only be stored safely in the Oxford Nexus 364 OneDrive for Business project folder. Offline data will be stored in a locked filing cabinet in the Department of Experimental Psychology at the Uniersity of Oxford. Data shared open-access will be pseudonymous and non-identifiable. Restricted access will be given upon request to the data that contains certain personal information including age and gender as it concerns a volnurable population (children).
The management and support of the users and data on the Oxford Nexus 364 OneDrive for Business folders is done by the IT Services' Nexus Team, and by default, access to all personal OneDrive for Business documents is restricted to the investigators on the project.
OneDrive for Business is certified against the internationally recognised ISO/IEC 27001:2013 standard for managing information security and approved by the University for all data.

5.2 Will the data be safely stored in trusted repositories for long term preservation and curation?

- Yes

Data in Belgium

Data will be stored on the KU Leuven RDR repository for the minimal duration of 10 years.
Data in England
Data will be stored on the UK data service indefinitely.

6. Ethics

6.1 Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?

- Yes

Personal data of a volnurable population (children) will be collected as part of the data collection. We will pseudonymise the data by replacing the name of participants with unique codes and replace date of borth by the individual's age in months. The file linking the code and personal identifiers will only be accessible to the investigators on the project and will be password protected and stored in a password protected folder only accesssible to the team members. This personal information will never be shared, instead the pseudonymous data that will not include variables that risk participant identification (location, name, date of birth) will be deposited on the repositories (KU Leuven RDR or UK Data Service). Nevertheless, data containing certain sensitive data such as age and gender, will be restricted and only shared upon request. Statements regarding the prevention of re-identification of

participants will be included in the data access policy, and data will be shared when recipients state that they will not make any attempt to identify any research participants.

6.2 Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

- Yes

With parental consent, the data will be suitable for sharing with bona fide researchers and will be available after locking and pseudonymisation. No identifiable personal data will be shared.

7. Other issues

7.1 Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

- No

Data management will be in accordance with this Data Management Plan, and the data management from the KU Leuven RDR, the UK Data Service, and the OSF data management as described in this DMP.