DMP title

Project Name My plan (C1-C2-IDN DMP) - DMP title

Project Identifier 3H210379

Grant Title C3/21/008

Principal Investigator / Researcher Joke Torbeyns

Description Preschoolersâ€™ mathematical competencies are important predictors for academic achievement in primary and secondary education and socio-economic status in adulthood. Cumulative evidence indicates that preschool teachers do not systematically stimulate preschoolersâ€™ acquisition of early number and patterning competencies, due to limitations both in the availability of effective research-based materials and in teachersâ€™ pedagogical content knowledge (PCK) in these domains. Following a design-based research methodology, this C3 project will address these limitations by designing and implementing (a) effective research-based instructional materials to enhance preschoolersâ€™ early number and patterning competencies, and (b) a research-based PCK guide for (future) preschool teachers focusing on the developmental trajectories and effective instructional strategies in these domains, complemented with a continuous professional development program. The effectiveness of the research-based instructional materials to enhance preschoolersâ€™ early number and patterning competencies will be evaluated via two intervention studies in Flemish classrooms with 4-5-year-old children.

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Joke Torbeyns

C1-C2 Project number & title

C3 3H210379

Towards powerful early math instruction: Designing materials to improve (future) preschool teachers’ and preschoolers’ number and patterning competencies.

2. Data description

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

The two intervention studies with preschoolers involve the same WP’s and same type of data to be collected. Only the intervention materials offered to the preschoolers will differ (with materials offered in Study 1 optimized on the basis of the findings of Study 1 and offered in optimized version in Study 2).

WP1: Recruitment. This WP involves demographic data and background characteristics of the participants (preschoolers) collected via a questionnaire. These will be scanned and stored as .pdf files and further inputted in a spreadsheet that is saved as .xls file. Data formats will be textual and numerical. Estimated data volume: 200MB per study.

WP2: Child competencies. This WP involves data about the participating children’s early mathematical and language competencies collected via paper-and-pencil and digital tasks. These will be inputted in a spreadsheet that is saved as .xls file. Data formats will be textual and numerical. Estimated data volume: 200MB per study.

WP3: Interaction during the intervention. This WP involves data about the teacher-child interactions during the intervention collected via audio-registered observations. These will be stored as mp4 files and literally transcribed as structured text that is stored as .docx file. Data formats will be audio and textual. Estimated data volume: 200GB per study.

3. Ethical and legal issues

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

Personal data will be used.

Short description of the kind of personal data that will be used: Personal data of the participants are name, sex, date-of-birth, contact information and signed informed consents. As we will observe and audio-register teacher-child interactions during the intervention, this also involves audio data. All personal information is only available to researchers involved in recruitment, data collection and data analysis. The file linking the code and personal identifiers is only accessible to these researchers. It is stored on the J:\ share of the project on KU Leuven network drives, according to the ICTS storage guidelines. For the remainder of each study, all derivative data will be coded, and thus pseudonymised. An ethics application for SMEC was already submitted and approved for the project as a whole (*G-2021-3882-R2(MAR)*). For each of the individual studies, a new ethics application for SMEC will be submitted and registered via the PRET tool.

3.2. 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, add the reference to the formal approval by the relevant ethical review committee(s).

An ethics application for SMEC was already submitted and approved for the project as a whole (*G-2021-3882-R2(MAR)*). For each of the individual studies, a new ethics application for SMEC will be submitted and registered via the PRET tool.

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

Our studies will result in effective instructional materials for preschoolers in the domain of early mathematics, to be offered to and published by an educational publisher. We are in contact with our colleagues with expertise related to IP at KU Leuven and will discuss all IP issues with them upon end of the project (before starting to contact educational publishers).

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

No

4. Documentation and metadata

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

We will make a codebook documenting the study design, sampling, measures, variables and analysis schemes that allows a secondary data analyst to use the data accurately and effectively. Metadata about the analytic techniques applied that is available in the statistical analysis software will be included as well. For the observation data, the codebook will be complemented with the data listing template available on https://www.ukdataservice.ac.uk/manage-data/document/data-level/qualitative.aspx.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

We will use the metadata standards from the Data Documentation Initiative for describing behavioral data and data collected via observation methods, see https://ddialliance.org/about-the-alliance.

5. Data storage and backup during the C1-C2 project

5.1. Where will the data be stored?

The data will be stored on the university's central servers with automatic daily back-up procedures.

5.2. How will the data be backed up?

The data will be stored on the university's central servers with automatic daily back-up procedures.

5.3. 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.

The J:\ and K:\ drives have sufficient storage for this project. The total estimated amount of data

is 480 GB.

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

The estimated cost for back-up at J:\ for 480 GB of data is approximately 385 euro. The estimated cost for storage at K:\ during the project is approximately 50 euro. This will be paid by the budget of the C3-project.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data will be stored at the university's secure environment for private data, J:\ and K:\ drives, which will only be accessible by the involved researchers.

6. Data preservation after the end of the C1-C2 project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained for the expected 10 year period.

6.2. Where will these data be archived (= stored for the long term)?

Offline copies (questionnaires, paper-and-pencil data) and informed consents will be separately archived in a locked room for 10 years after the end of the project, as required by the KU Leuven RDM policy. They will be destroyed after the 10 year period. Digital data will be stored on the university's central servers (with automatic back-up procedures) for a period of 10 years, conform the KU Leuven RDM policy. This will be done on the K:\ drive, bound to the KU Leuven ICT code of conduct.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The estimated cost for archiving on K:\ of about 480 GB is approximately 960 euro. This will covered by the FWO-project.

7. Data sharing and re-use

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

All intervention sessions will be audio-registered, resulting in personal data that cannot be shared with researchers that are not directly involved in the project. We will not share these audio-registered observation data. Only pseudonymized data will be available for sharing.

7.2. Which data will be made available after the end of the project?

The pseudonimized dataset of each study will be uploaded on the OSF account of the PI in a csv format (https://osf.io/cmvdh/).

7.3. Where/how will the data be made available for reuse?

- In a restricted access repository

The pseudonimized dataset will be uploaded in a csv format at the OSF on the account of the PI upon publication of a study.

7.4. When will the data be made available?

- Upon publication of the research results

The pseudonimized dataset will be uploaded in a csv format at the OSF on the account of the PI upon publication of a study.

7.5. Who will be able to access the data and under what conditions?

The pseudonimized dataset of a study will be shared in a csv format on the OSF platform. It will be available to anyone provided that they give appropriate credit.

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

There are no costs expected for data sharing.

8. Responsibilities

8.1. Who will be responsible for the data documentation & metadata?

The PI, Joke Torbeyns

8.2. Who will be responsible for data storage & back up during the project?

The PI, Joke Torbeyns

8.3. Who will be responsible for ensuring data preservation and sharing?

The PI, Joke Torbeyns

8.4. Who bears the end responsibility for updating & implementing this DMP?

The end responsibility for updating and implementing the DMP is with the supervisor (promotor).