DMP title

Project Name Youth mental health and relationships with teachers and peers: Unravelling their interplay - DMP title

Grant Title FWO G0D6222N

Principal Investigator / Researcher Hilde Colpin

Description Youth mental health problems are on the rise worldwide and threaten the future of our children and society at large. Positive social relationships can play a key role in preventing mental problems and promoting well-being. This project focuses on social relationships at school. More specifically, it digs into the understudied interplay among different dimensions of teacher-student and peer relationships in early adolescence, and their links with studentsâ€™ mental (ill-)health, by addressing two main research objectives: (1) to investigate the unique and joint effects of different dimensions of teacher-student relationships and peer relationships on different student mental health outcomes, (2) to unravel the interplay among teacher-student relationships and peer relationships. The project comprises two longitudinal studies with existing data sets in grade 4-6 of elementary school (Study 1: N=1051 students, 62 teachers; Study 2: N=1067 students, 63 teachers) and an experimental study in pre-service teacher training (Study 3: N to be recruited =320). Instruments consist of validated student and teacher reports, peer nominations, and observations. State-of-the-art statistical techniques for analyzing longitudinal and hierarchical data and for experimental data are used.

Institution KU Leuven

1. General Information

Name applicant

Supervisor: Hilde Colpin

Co-supervisors: Karine Verschueren, Jantine Spilt

FWO Project Number & Title

G0D6222N - Youth mental health and relationships with teachers and peers: Unravelling their interplay.

Affiliation

- KU Leuven

2. Data description

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

- Generate new data
- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

The project comprises three studies based, each based on a different data set. In this DMP, as in the project proposal, the three studies (and hence the corresponding data sets) are referred to as STUDY 1, STUDY 2, and STUDY 3. STUDY 1 and 2 make use of existing data sets; for STUDY 3, a new data set will be collected. For STUDY 2, a DMP is already in place (FWO G057220N, approved by KU Leuven National Funds, e-mail June 3th, 2020). We refer to that DMP for the information about STUDY 2, but have repeated the main information in this new DMP.

For STUDY 1, data of the longitudinal Teachers4Victims project are used (FWO G071317N). Students in grade 4-6 of elementary school and their teachers and parents completed questionnaires at three occasions (wave 1, 2, 3) in school year 2018-2019. For students and parents, paper and pencil questionnaires were used. Teacher questionnaires were administered in Qualtrics (an online survey platform); after each data collection session, digital teacher data were downloaded fromQualtrics and stored as a .sav file (SPSS) and .xlsx file. (Excel). In a few cases of technical problems, a paper version of the teacher questionnaire was administered. In addition, classroom interactions were observed on wave 1. Observations were coded on paper by the researchers in the classrooms; no videorecordings were made. All data on paper were digitalized and stored as a .sav file (SPSS) and .xlsx file (Excel). Later, the different files were

merged into one final data set (sav. file and xlsx file). Data are analyzed by statistical software, such as SPSS, R, MPlus. The estimated data volume does not exceed 1 GB.

For STUDY 2, data from FWO G057220N, a randomized controlled trial with a teacher training in grade 4-6 of elementary school, are used. In line with the approved DMP for that project (cf. supra), all teachers and students completed questionnaires at the pre-test (October/November 2021) and the post-test assessment (March/April 2022). Teachers in the training condition were additionally be asked to fill out questionnaires about their evaluation of the training session (after each session) and about their implementation of the training (before the second and third session). The questionnaires were administered in Qualtrics (an online survey platform). After each data collection session, data were downloaded from Qualtrics and stored as a .sav file (SPSS) and xlsx file (Excel). In a few cases of technical problems, a paper version of the questionnaire was used and responses digitalized afterwards *(*typed in an Excel file (.xlsx). After the data collection, the dataset containing the digitalized paper-and-pencil data were merged with the dataset containing the Qualtrics-data. Data will be analyzed with statistical software packages such as SPSS, R and MPlus. The estimated data volume does not exceed 1 GB.

For **study 3**, data will be collected in a sample of pre-service teachers. Participants will be asked to fill out questionnaires at the pre-test and the post-test assessment of an experimental study. The questionnaires will be administered in Qualtrics (an online survey platform). After each data collection session, data will be downloaded from Qualtrics and stored as a .sav file (SPSS) and/or .xlsx file (Excel). In case of technical problems, a paper version of the questionnaire will be used and responses will be digitalized afterwards (typed in an Excel file (.xlsx) and/or in a SPSS file (.sav)). After the data collection, the dataset containing the (eventual) digitalized paper-and-pencil data will be merged with the dataset containing the Qualtrics-data. Data will be processed and analyzed in Excel and SPSS and in other statistical packages such as R and MPlus. The estimated data volume is not expected to exceed 1 GB.

ORIGIN	HOW CREATED	TYPE	FORMAT	ESTIMATED VOLUME
STUDY 1 secondary	questionnaires filled out by teachers in Qualtrics or on paper questionnaires filled out by students and parents on paper class observations on paper paper questionnaires and observations later digitalized N teachers=62 students=1051	observational data (survey data)	textual/numerical · .sav (SPSS) · .xlsx (Excel) · .csv for archiving purposes	Max. 1 GB
STUDY 2 secondary	questionnaires filled out by teachers and students in Qualtrics or on paper paper questionnaires later digitalized N teachers=63 students=1067	observational data (survey data)	textual/numerical · .sav (SPSS) · .xlsx (Excel) · .csv for archiving purposes	Max. 1 GB
STUDY 3 primary	questionnaires filled out by pre-service teachers in Qualtrics or on paper N teachers = +/- 320	observational data (survey data)	textual/numerical · .sav (SPSS) · .xlsx (Excel) · .csv for archiving purposes	Max. 1 GB

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- Yes

STUDY 1

Privacy Registry Reference: G-2022-4832
Short description of the kind of personal data that will be used:

- We collected ordinary personal data using questionnaires (namely participants' name and birthdate or age and data on a variety of constructs relevant for our research objectives) and classroom observations
- For students and parents, the data were administered on paper and later digitalized.
- Both on paper and in the digital data set, the data were pseudonymized. Prior to data collection, all participants were assigned a unique ID-number, which was noted on their questionnaires. In the digital data set, participants' responses on the questionnaires are not saved with their names but with the unique ID-numbers. Moreover, in the final dataset information on participants' birthdate was deleted and replaced by their age.
- Paper data (with ID-codes, not participants' names) and parental informed consents are saved separately in a cabinet (within a building of the KU Leuven) that is closed with a lock.

STUDY 2
Privacy Registry Reference: G-2020-1612-R5(AMD)

Short description of the kind of personal data that will be used:
In line with the DMP that is in place for FWO-project G057220N:

- We collected ordinary personal data using questionnaires (namely participants' name and birthdate/age and data on a variety of constructs relevant for our research objectives)
- The data were pseudonymized. Participants' responses on the questionnaires were not saved with their names but with unique ID-numbers which we assigned to all participants prior to the start of the baseline assessment. Moreover, in the final dataset information on participants' birthdate will be deleted and replaced by their age (data cleaning ongoing).
- The questionnaires were administered on laptops or tablets of the university; thus no identifying IP address was recorded electronically.
- Paper questionnaire data and the paper informed consent forms are saved separately in a cabinet (within a building of the KU Leuven) that is closed with a lock.

STUDY 3
In this study, anonymous data will be collected. We will collect ordinary personal data using questionnaires (e.g., participants' age, gender and data on a variety of constructs relevant for our research objectives), but we will not ask for participants' names.

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)

- Yes

STUDY 1: G-2016 12 700
STUDY 2: G-2020-1612-R5(AMD)
STUDY 3: to be obtained before the start of this study (2023)

Does your work 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?

- No

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 are in place?

- Yes

The studies' participants are involved as a third party. Agreements as such are part of the Informed Consent form, mentioning the publication of results in scientific communications.

STUDY 2 is based on a project in collaboration with Prof. dr. Heidi Vandebosch (co-PI; University of Antwerp); dissemination or exploitation of the data will be discussed with her and will take place in mutual agreement, as outlined in the DMP in place for FWO-project G057220N.

4. Documentation and metadata

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

For STUDY 1, we created created data documentation at different levels:

1. Project level

- Project description and aims: short description of the main objectives of the project and research methodology together with grant number (.docx, .pdf, or .txt)
- Recruitment: documentation on the methodology used to recruit schools (and within schools: teachers and students) (.docx, .pdf, .xlx); and recruitment materials (recruitment letter for schools, information leaflet for teachers and for parents) (.pdf)
- Data collection: documentation on the methodology used to collect the data and all relevant research documents (data collection protocol, blank copy of the informed consent forms, blank copy of the questionnaires) (.docx, .pdf); documentation on the data collection (e.g., relevant notes made during data collection) (.docx, .pdf, or .txt)
- Project management: reports of the meetings of the research team (.docx); the data management plan itself (.docx, .pdf)
-
- 2.File or database level

- Dataset: documentation on the creation of the (final) dataset(s) (e.g., data log with documentation on data cleaning and on the different versions of data files; syntax used for deriving variables from original variables) (.docx, .pdf, .sps)
- Data-analysis: for each manuscript documentation on the (final) analyses (e.g., analysis logbook) (depending on statistical software used: .sps)
-
- 3. Item level
- Codebook dataset: explanatory information (i.e., annotation metadata) about the variables and how they were coded (e.g., variable labels, value labels, missing values) (in the dataset (.sav), and a variable legend in a separate text file: docx, .pdf, or .txt)

For STUDY 2, a DMP is in place (FWO-project G057220N) (cf. supra). Data collection for this project was finished in May 2022 and data cleaning is ongoing. In line with this DMP we either did or will create -as specified in the DMP- data documentation at various levels:

1. Project level

- Project description and aims: short description of the main objectives of the project and research methodology together with grant number (.docx, .pdf, or .txt)
- Training: documentation on the development of the training and information session about bullying, and a master file containing the final training and information session materials (.pdf)
- Recruitment: documentation on the methodology used to recruit schools (and within schools: teachers and students) (.docx, .pdf, .xlx); and recruitment materials (recruitment letter for schools, information leaflet for teachers and for parents) (.pdf)
- Data collection: documentation on the methodology used to collect the data and all relevant research documents (data collection protocol, blank copy of the informed consent forms, blank copy of the questionnaires) (.docx, .pdf); documentation on the data collection (e.g., relevant notes made during data collection) (.docx, .pdf, or .txt)
- Project management: reports of the meetings of the research team (.docx); documentation on the folder structure, used file naming conventions and used abbreviations (.docx, .pdf, or .txt); the data management plan itself (.docx, .pdf), and documentation on the project progress (time line with main dates, e.g. of ethical application, ethical approval, start of data collection) (.docx, .pdf, or .txt)
-
- 2. File or database level

- Dataset: documentation on the creation of the (final) dataset(s) (e.g., data log with documentation on data cleaning and on the different versions of data files; syntax used for

deriving variables from original variables) (.docx, .pdf, .sps); and administrative metadata about the data namely a ReadMe text file with basic information about the dataset (including the location(s) of data storage) (.txt)
- Data-analysis: for each manuscript documentation on the (final) analyses (e.g., analysis logbook) (depending on statistical software used: .sps)

3. Variable or item level

- Codebook dataset: explanatory information (i.e., annotation metadata) about the variables and how they were coded (e.g., variable labels, value labels, missing values) (in the dataset (.sav), and a variable legend in a separate text file: docx, .pdf, or .txt)

For STUDY 3, we will follow the same procedure as for STUDY 2.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

There is no metadata standard available in our discipline.
For STUDY 1 and STUDY 2, data-input was conducted mainly in SPSS (or Excel). We used the metadata characteristic for these programs. This ascertains that data are understandable for all collaborators in the project and for researchers interested in conducting secondary analyses using these data. We will use the same procedure for STUDY 3.

5. Data storage and backup during the FWO project

Where will the data be stored?

Like for STUDY 2, for which a DMP is in place (FWO-project G057220N) for STUDY 1 and 3, in line with the guidelines of our faculty and with the GDPR, the digital data are stored in a secured, limited access network drive operated by the faculty's ICT service. Data are pseudonymised (STUDY 1&2) or anonymous (STUDY 3). For STUDY 1&2, the files containing the participants' names and ID-codes are stored separately from the files that contain their data. The files with the names can only be accessed by the data managers of the research group (K. Luyckx, J. Spilt). All computers working with the data have encrypted hard drives (Bitlocker). The cost for this secure data storage (at least 5 years after the end of the research) is budgeted.
Paper data and the paper informed consent forms are stored separately at the KU Leuven in a cabinet that is closed with a lock.

How is backup of the data provided?

- The shared and non-shared network drives (managed by the faculty’s ICT personnel (PPW DICT)) where we store the data are automatically backed-up on a frequent basis.

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

We do not need much storage volume to store the data because the studies only collect questionnaire data and (for study 1) coded data about classroom observations. The estimated volume per study does not exceed 1 GB. There is sufficient storage and backup capacity available on the non-shared and shared network drives (I:, J: drive, K: drive, KU Leuven One drive) to store the data.

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

The costs for storage and backup of the data of the current project (estimated volume: max 3 GB; estimated costs max. 500 euro) will be covered by the FWO-project.

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

As for STUDY 2, for which a DMP is in place (FWO G057220N), for STUDY 1 and 3:

- To prevent unauthorized persons to have access to our data, data are not stored locally on the laptops of the involved researchers, but are stored on the secure network drives which are managed by the faculty's IT staff (cf. supra, question "where will the data be stored"). The use of these network drives allows for secured storage, management and sharing of files, and avoids loss of data and data breaches.
- (Part of the) final pseudonymized (Study 1 and 2)/anonymous (Study 3) dataset are/will be shared with KU Leuven students who need the data for their master thesis. They are requested to encrypt the hard drives of their laptops before receiving the data and to delete the data after having defended their master thesis.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained.

Where will the data be archived (= stored for the longer term)?

For STUDY 2, the DMP in place for FWO-project G057220N specifies that all data (paper and digital) will be stored for five years after publication of manuscripts based on these data and that, after this period:

- The eventual paper data and paper informed consent forms will be destroyed
- The key will be deleted
- The pseudonymized dataset will be archived with limited access on a restricted area of the K: archive drive of the PI (for archiving purposes and for allowing secondary data analyses). Meanwhile, the PI's archive moved to KU Leuven One drive and data will be stored there on the longer term.

For STUDY 1 and 3, the data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the current KU Leuven RDM policy. After this period, the same procedure as for STUDY 2 will be followed.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The costs for storage and backup of the data of the current project (max 3 GB) are estimated costs at max. 500 euro and will be covered by the FWO-project.

7. Data sharing and reuse

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)?

- No

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

Like specified in the DMP in place for STUDY 2 (FWO G057220N), for STUDY 1 and 3, studies will be registered on OSF and interested researchers can send a motivated request to the PI (H. Colpin) if they would like obtain access to the final, pseudonymized (STUDY 1) or anonymous (STUDY 3) dataset. Hence there will be restricted or controlled access to the data.

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

- In a restricted access repository

Studies will be registered on OSF and interested researchers can send a motivated request to the PI (H. Colpin) if they would like obtain access to the final dataset (pseudonymized for Study 1 and 2, anonymous for Study 3). Hence there will be restricted or controlled access to the data.

When will the data be made available?

- Upon publication of the research results

Like for STUDY 2 (see DMP for FWO-project G057220N), once all results have been published on certain data, these data can be obtained from the PI (H. Colpin) in its pseudonymized form (STUDY 1) or its anonymous form (STUDY 3).

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

For STUDY 1, access is available for the PI Hilde Colpin and co-PI's Karine Verschueren and Jantine Spilt, post-docs Karlien Demol and Chloë Finet, PhD-student Fleur Van Gils, master student Anouck Lubon; in the future: thesis students, future PhD students and/or postdocs of the PI and co-PI's research group can send the PI a motivated request to obtain access to the pseudonymized data. They will only have access to the pseudonymized data, and only if they agree to treat the data confidentially.

For STUDY 2, as specified in the DMP in place, access is available for the PI Hilde Colpin, co-PI Heidi Vandebosch, and post-doc C. Finet. Meanwhile, master student Anouck Lubon collaborated to the data collection and will have access to the data as well. In the future, master thesis students, future PhD students and/or postdocs affiliated to the research group of the (co-)PI) can send the PI and the co-PI a motivated request to obtain access to the pseudonymized data. They will only have access to the pseudonymized data, and only if they agree to treat the data confidentially.

For STUDY 3, access is available for the PI H. Colpin, co-PI's K. Verschueren and Jantine Spilt, and post-doc K. Demol; in the future: thesis students, future PhD students and/or postdocs of the PI and co-PI's research group can send the PI a motivated request to obtain access to the anonymous data. They will only have access if they agree to treat the data confidentially.

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

No costs are expected.

8. Responsibilities

Who will be responsible for data documentation & metadata?

The main responsibility for data documentation and metadata lays with the PI (H. Colpin).

Others involved in this process:

-For STUDY 1: co-PI Karine Verschueren, post-doc Karlien Demol

-For STUDY 2, co-PI Heidi Vandebosch, post-doc Chloë Finet

-For STUDY 3: co-PI's Karine Verschueren and Jantine Spilt, post-doc Karlien Demol

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

The main responsibility for data storage and back up during the project lays with the PI (H. Colpin). The co-PI's (STUDY 1: K. Verschueren, STUDY 2: H. Vandebosch, STUDY 3: J. Spilt, K. Verschueren) and post-docs K. Demol and C. Finet will also be involved in this process. The PPW Dict Personnel are responsible for the automatic backup system of the network drives.

Who will be responsible for ensuring data preservation and reuse ?

The main responsibility for ensuring data preservation and use lays with the PI (H. Colpin). The co-PI's (STUDY 1: K. Verschueren, STUDY 2: H. Vandebosch, STUDY 3: J. Spilt, K. Verschueren) will also be involved in this process.

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

The PI bears the end responsibility of updating & implementing this DMP.