DMP title

Project Name C3-project URGENT - DMP title

Project Identifier C3/21/056

Principal Investigator / Researcher Greet Langie

Project Data Contact Sofie Craps (sofie.craps@kuleuven.be) and Mieke Cannaerts (mieke.cannaerts@kuleuven.be)

Description In this interdisciplinary project we will examine how the existing PREFER professional roles tests and the knowledge of social psychology can be combined and contribute to the development of the future disciplinary self, and more specifically to the attraction and retention of female students and students with a migration background in engineering programmes. The research questions are: 1. Are the initial beliefs and predispositions on engineering the same for men/women and students with/without a migration background? 2. Does the answer to the previous question helps to understand why some groups experience a limited fit and lower feeling and belonging? 3. Are the items and cases in the PREFER 1.0-tests gender and culturally neutral and valid? 4. Are there differences in professional role preferences between students whether or not they belong to one of the two target groups? 5. Can we broaden the stereotypical initial beliefs by redesigning brochures and information days, in order to improve recruitment in general and more specifically the attractiveness for female students? 6. Can we address studentsâ€™ feeling of belonging by stimulating reflection with the help of the PREFER 1.1- tests, in order to improve retention of students in general and more specifically the retention of students with a migration background? These research objectives will result in two new projects: a Horizon Europe project focusing on an inclusive society and a collaboration with industry to implement the optimised PREFER-tests in HR-applications. A mixed methodology will be applied to collect the data: 1. Quantitative research based on an existing validated survey and the PREFER tests. 2. Qualitative research based on semi-structured interviews with students from both target groups and a narrative study in which final-year pupils with a migration background are followed from their choice to start an engineering program at KU Leuven through their first year in the engineering program. 3. Design-based research to optimize the tests and to develop and implement the interventions.

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Greet Langie

Internal Funds Project number & title

C3/21/056 - URGENT - UnderRepresented Groups of ENgineering sTudents: Role of professional awareness in inclusion

2. Data description

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

- Generate new data
- Reuse existing 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.

Part of our research will reuse data generated by the completed European project PREFER, brought together within the PhD-thesis of the postdoc who is part of this project team.

Additionally, we will collect different types of new data:

Type of data	Kind of data	Format	Volume	How created?
Qualitative data	multimedia and textual	Audio-recorded and transcribed into MS Word format and Nvivo.	?	interviews and focusgroup discussions
Qualitative data	textual	Word	?	narrative study
Quantitative data	numeric and textual	Excel, SPSS and R	?	validated survey
Quantitative data	numeric and textual	Excel, SPSS and R	?	PREFER tests
Quantitative data	numeric and textual	Excel, SPSS and R	?	background information of the students (academic achievement, secondary education information, socio-economic status, ..) collected through KU Leuven SAP

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.

Yes, we will use personal data.

Our study is inclusive, but we are specifically interested in how the values and beliefs of **female students** and **students with an imigration background** differ from the other first-year engineering students.

We will collect the following existing and new information: programme at university, study track during secondary education, academic achievement, GPA secondary education, LASSI-results, socio-economic status, etc.

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).

We will submit ethical approval to SMEC since our research is performed on human subjects, but is not related to health science practices or does not include medical or pharmacological procedures.

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 research project is a C3 project. By consequence valorisation is essential.

IP restrictions will be discussed with LRD.

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?

There are no existing 3rd party agreements yet.

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 provide documentation in the following way:

1. Qualitative data (interviews and focusgroup discussions): each original and transcribed audio-recorded file will be accompanied by a standardized overview sheet containing the following elements: date, setting, participants, informed consent, implemented guidelines, etc..

2. Qualitative data (narrative study): the Word document with the narratives of a student will be accompanied by a standardized overview sheet containing the following elements: name student, duration, setting, informed consent, guidelines, etc.

3. Quantitative data (validated survey): a codebook will accompany the collected survey data, containing the date the survey was organized, the population, the setting, variable information, etc.

4. Quantitative data (PREFER tests): a codebook will accompany the collected data, containing the date the test was organized, the population, the setting, variable information, etc.

This documentation will be kept in the same folder as the data. The name of the folder will contain the name of the method and the date.

We will anonimise our data where possible, pseudonomise where necessary for longitudinal follow up and work with the non-anonimised data if pseudonomised data is really insurmountable in the long run. The file with the participants ID's in the case of pseudonomised data will be in the appropriate folder.

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.

It's our goal to use a project-specific metadata standard to guarantee a feasible, efficient and effective DMP. This metadata standard will be based on DDI and it will be determined by mutual agreement within the project group before the first data are collected.

5. Data storage and backup during the project

5.1. Where will the data be stored?

Since we will be working with sensitive personal data that will be pseudonymized (or anonymized where possible), the data will be stored at the university's secure environment for private data: the j-disk of KU Leuven. A working copy for analyses will be stored at the KU Leuven Teams of the research group FIIW-ETHER that is only accessible for the researchers of the project. A private channel for URGENT, only accessible for the researchers of this project will store the datafiles.

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.

We do not expect problems with storage capacity.

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

We do not expect costs for data storage and backup during the project.

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

The j-disk on the university's central servers is a secure environment. Also 'Teams' is a KU Leuven environment with restricted accessibility.

6. Data preservation after the end of the 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 after the end of the project.

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

The data will be stored on the university's central server (j-disk) for at least 10 years, conform the KU Leuven RDM policy. Automatic back-up procedures are present.

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

There are no expected costs.

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)?

At this moment there are no agreements yet with 3rd parties in the context of the future IP. However this will be necessary in the final phase of the project once we will start collecting data outside KU Leuven. This will be an important point of focus during the valorization phase of the project.

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

All collected data will remain available for future reuse after the end of the project for the members of the research unit.

Future reuse by the broader scientific community is only possible after anonymisation of the data since we will be working with sensitive personal data.

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

- Upon request by mail

Data will be available on request after signing a data sharing agreement.

7.4. When will the data be made available?

- Upon publication of the research results

Since the analysis of the collected data will result in a PhD thesis, the data will only become available after the publication of the PhD thesis.

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

The anonymised data can be obtained upon request. This request should contain the planned reuse. At this moment, we prefer to limit this reuse to research purposes.

However, once the valorization of the PREFER tests is realized and the necessary IP is arranged in collaboraion with LRD, commercial reuse will be possible. At that moment, we have to reconsider this DMP.

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

There are no costs at this moment.

8. Responsibilities

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

The final responsible is the project promotor.

The PhD student and postdoc will continuously collect data, copy it on the Teams' research unit 'FIIW - ETHER' and add the necessary metadata based on the 'standardized overview sheet' and the 'codebook'.

The format of the standardized overview sheet and the codebook will be discussed and finalized in full agreement with all members of the project team.

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

The final responsible is the project promotor.

At the moment a publication of some research results is accepted, the PhD student and postdoc will copy the considered data for the long term on the j-disk. This secures a systematic follow-up

of the collected data, such that all data are securely stored at the end of the project.

8.3. Who will be responsible for ensuring data preservation and sharing?
The final responsible is the project promotor.

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).