Flemish Standard Data Management Plan

1. General Project Information	
Name Grant Holder & ORCID	Ellen Van Houtven | PhD candidate | http://orcid.org/0000-0003-3988-247X
Contributor name(s) (+ ORCID) & roles	Kathleen Beullens | Supervisor | http://orcid.org/0000-0002-0530-7947 Desiree Schmuck | Co-Supervisor | http://orcid.org/0000-0002-9492-6052
Project number & title	3H220735 | The Role of Digital 'Super Peers' in Gender and Sexual Identity Achievement Among Adolescents
Funder(s) GrantID	11Q2B24N
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☒ Other: University of Vienna ROR identifier KU Leuven: 05f950310 ROR identifier University of Vienna: 03prydq77

Please provide a short project description	Adolescence is a crucial phase for the exploration and achievement of gender and sexual identity. During this period, parents, peers, but also the media can act as important socialization agents who shape norms and values. In particular, digital environments offer unique opportunities to turn to media personae, so-called 'super peers', who can serve as role models and information sources regarding gender and sexual identity in a way that is often not available in adolescents' own social environment. Against this background, this PhD project makes three important contributions: First, drawing from a multidisciplinary theoretical framework, it introduces the novel concept of digital super peers (DSPs), and investigates the role of fictional (i.e., TV show characters) and non-fictional (i.e., social media influencers) DSPs in the exploration and achievement of adolescents' gender and sexual identity. Second, it explores how the socializing impact of DSPs ties in with adolescents' intersecting identities, developmental phase, and social context. Third, it studies how gender non-conforming and non-heteronormative identities are portrayed within adolescents' digital media repertoires and how they affect perceptions of their own and others' gender and sexual identity. These contributions will be realized by using an innovative mixed-method approach combining interlinked longitudinal surveys and qualitative content analyses, in-depth interviews, and a multiple-exposure experiment.

2. Research Data Summary

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*	*ONLY FOR PHYSICAL DATA*
Dataset Name	**Description**	**New or Reused**	**Digital or Physical**	**Digital Data Type**	**Digital Data Format**	**Digital Data Volume (MB, GB, TB)**	**Physical Volume**
Interview_Guide	Interview guide containing topics and questions for the semi-structured interviews (WP1	New	Digital	Textual	.docx	< 1 GB	/
Interview_Recordings	Audio recordings of ±20 qualitative interviews (WP1)	New	Digital	Sound	.mp3	< 1 GB	/
Interview_Transcripts	Ad verbatim transcripts of ±20 qualitative interviews (WP1)	New	Digital	Textual	.docx	< 1 GB	/
Interview_Analysis	Qualitative data analysis by coding the interview transcripts in Nvivo (WP1)	New	Digital	Software	.nvp	< 1 GB	/
Survey_Data_W1	Longitudinal survey data wave 1 (WP1)	New	Digital	Software	.sav	< 1 GB	/

Survey_Data_W2	Longitudinal survey data wave 2 (WP1)	New	Digital	Software	.sav	< 1 GB	/
Survey_Data_W3	Longitudinal survey data wave 3 (WP1)	New	Digital	Software	.sav	< 1 GB	/
Survey_Script_CS	R markdown file containing the R script for the quantitative data analysis of the longitudinal survey | Cross-sectional after wave 1 (WP1)	New	Digital	Software	.rmd	< 1 GB	/
Survey_Script_LG	R markdown file containing the R script for the quantitative data analysis of the longitudinal survey | Longitudinal after waves 1-3 (WP1)	New	Digital	Software	.rmd	< 1 GB	/
CA_SM_Content_Recordings	Screen recordings of social media content included in the	New	Digital	Audiovisual	.mp4	< 100 GB	/

	qualitative content analysis (WP2)						
CA_SM_Content_Pictures	Screenshots of social media content included in the qualitative content analysis (WP2)	New	Digital	Images	.png	< 100 GB	/
CA_SM_Transcripts	Transcripts of the social media content that will be qualitatively analyzed (WP2)	New	Digital	Textual	.docx	< 1 GB	/
CA_TV_Descriptions	Descriptions of the digital TV show episodes that will be qualitatively analyzed (WP2)	New	Digital	Textual	.docx	< 1 GB	/
CA_Codebook	Codebook containing all themes and codes for the qualitative content analysis (WP 2)	New	Digital	Textual	.docx	< 1 GB	/
CA_Analysis_W1	Qualitative data analysis of the first coding	New	Digital	Software: Data analysis in Nvivo	.nvp	< 1 GB	/

	round by coding the media content in Nvivo (WP2)						
CA_Analysis_W2	Qualitative data analysis of the second coding round by coding the media content in Nvivo (WP2)	New	Digital	Software: Data analysis in Nvivo	.nvp	< 1 GB	/
Experiment_Data_T1	Multiple-exposure experiment data wave 1 (WP3)	New	Digital	Software	.sav	< 1 GB	/
Experiment_Data_T2	Multiple-exposure experiment data wave 2 (WP3)	New	Digital	Software	.sav	< 1 GB	/
Experiment_Data_T3	Multiple-exposure experiment data wave 3 (WP3)	New	Digital	Software	.sav	< 1 GB	/
Experiment_Data_T4	Multiple-exposure experiment data wave 4 (WP3)	New	Digital	Software	.sav	< 1 GB	/
Experiment_Data_T5	Multiple-exposure experiment data wave 5 (WP3)	New	Digital	Software	.sav	< 1 GB	/

Experiment_Script	R markdown file containing the R script for the quantitative data analysis of the multiple-exposure experiment (WP 3)	New	Digital	Software:	.rmd	< 1 GB	/

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.	/
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: G-2023-7181 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: (Special category) personal data will be collected and processed for WP1 and WP3 of this research project. For example: age, gender, sexual orientation, ethnicity and religion will be asked for. Moreover, research participants include minors. Ethical approval from SMEC will be asked for the entire project (G-2023-7181).
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) ☐ No Additional information: WP1 and WP3 involve the processing of (special category) personal data, including gender, sexual orientation, ethnicity, religious beliefs, etc. We will also ask for participants' email address in the longitudinal survey and multiple-exposure experiment, so that we can recontact them for the following steps. Additionally, respondents' in WP1 and WP3 will be asked to indicate their email if they want to receive the aggregated study results. In any case, email addresses will always be stored separately from the survey responses. SMEC number: G-2023-7181

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.	☐ Yes ☒ No If yes, please comment: /
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain: /
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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain: /

3. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep **data 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).	On the one hand, I will provide project-level documentation via a README.txt file. On the other hand, I will provide file-level documentation through additional README.txt files for each study. Furthermore, comments will be added within the R syntax for the data analysis of WP1 and WP3. The PhD candidate will store the collected data securely on the KU Leuven drives based on the different work packages within the project. WP1 exists of 9 datasets with the names 'Interview_Guide' (with as data the interview guide based on topics and questions to discuss during the semi-structured interviews), 'Interview_Recordings' (with as data the audio recordings of ±20 qualitative interviews), 'Interview_Transcripts' (with as data the ad verbatim transcripts of the ±20 qualitative interviews), 'Interview_Analysis' (with as data the coding of the interview transcripts in Nvivo), 'Survey_Data_W1' (with as data the survey data of wave 1), 'Survey_Data_W2' (with as data the survey data of wave 2), 'Survey_Data_W3' (with as data the survey data of wave 3), 'Survey_Script_CS' (with as data the R script for the quantitative data analysis of the cross-sectional first survey wave), and 'Survey_Script_LG' (with as data the R script for the quantitative data analysis of the longitudinal survey waves). WP2 exists of 7 datasets with the names 'CA_SM_Content_Recordings' (with as data the screen recordings of the social media content), 'CA_SM_Content_Pictures' (with as data the screenshots of the social media content), 'CA_SM_Transcripts' (with as data the transcripts of the social media content), 'CA_TV_Descriptions' (with as data the descriptions of the digital TV show episodes), 'CA_Codebook' (with as data the themes and codes for the qualitative content analysis), 'CA_Analysis_W1' (with as data the first set/wave of coded media content), and 'CA_Analysis_W2' (with as data the second set/wave of coded media content). WP3 exists of 6 datasets with the names 'Experiment_Data_T1' (with as data the survey data of T1), 'Experiment_Data_T2' (with as data the survey data of T2), 'Experiment_Data_T3' (with as data the survey data of T3), 'Experiment_Data_T4' (with as data the survey data of T4), 'Experiment_Data_T5' (with as data the survey data of T5), and 'Experiment_Script' (with as data the R script for the quantitative data analysis of the multiple-exposure experiment).

	A README.txt file will be provided on OSF, where anonymized datasets will be provided for reuse. Additionally, comments on data transformations and analyses will be further noted on within the .rmd files.
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.	☐ Yes ☒ No If no, please specify (where appropriate per dataset or data type) which metadata will be created: / Data will be uploaded in the data repository OSF, where additional metadata will be provided. For WP1, we will upload a file containing more detailed information on the set-up of the qualitative interviews. Moreover, for the longitudinal survey, we will provide a pseudonymized metadata file, in which, e.g., the codes linking participants across surveys have been deleted. For WP2, the codebook will be made available and will contain descriptions of all themes. That way, it can be reused in future studies. For WP3, additional info will be provided on the concrete procedure that was used. Additionally, it will describe the stimulus materials and selection thereof.

4. Data Storage & Back-up during the Research Project

Where will the data be stored?	☒ Shared network drive (J-drive) ☒ Personal network drive (I-drive) ☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☐ Other: All collected data within the scope of this project will be stored on the secure KU Leuven drives (I- drive and J-drive). So, access to this data is personal and password-protected/accessible only through VPN. Furthermore, the identifying data will be stored in separate datasets than pseudonymized data so that it cannot be linked.
How will the data be backed up?	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)
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 ☐ No If no, please specify: /
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?	Once processed, all data will be stored on KU Leuven's shared network drive (i.e., J-drive). This drive can only be accessed through the password-protected intranet or VPN, and is only accessible to KU Leuven employees who have been granted access. Additionally, data will be upload to the drive with read-only access, so that unauthorized individuals may not modify the data. Also, the identifying data will be stored in separate datasets than pseudonymized data so that it cannot be linked.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	There are no additional costs expected. The cost to access the I- and J-drives have already been covered by the research group.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five 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 ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☒ Certain data cannot be kept for 10 years (explain) All collected research data will be stored for 10 years on secure drives, complying to KU Leuven's RDM policy, before being permanently deleted. However, email addresses will not be kept for this 10-year period. Rather, these will be deleted as soon as they are no longer needed (a) to recontact the research participants for follow-up surveys within the study (i.e., longitudinal survey + multiple-exposure experiment), and/or (b) to send respondents the anonymized and aggregated research results, in case they requested to receive these.
Where will these data be archived (stored and curated for the long-term)?	☐ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☒ Shared network drive (J-drive) ☒ Other (specify): Apart from the shared KU Leuven drive (J-drive), data will also be put on the OSF repository.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	There are no additional costs expected. The cost to access the I- and J-drives have already been covered by the research group.

6. 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 ☒ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☒ Other, please specify: The interview guide and codebooks (WP1 & WP2) will be made available for reuse. Identifying data will be deleted before datasets are shared for reuse. Identifying information will at no point be accessible to people other than the PhD candidate and supervisors.
If access is restricted, please specify who will be able to access the data and under what conditions.	/
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.	☒ Yes, privacy aspects ☐ Yes, intellectual property rights ☒ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No If yes, please specify: Identifying information of the research participants will never be shared.
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ KU Leuven RDR ☒ Other data repository (specify): OSF ☐ Other (specify) All data that will be made available, will be uploaded on the Open Science Framework (OSF).
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)

Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify) This will allow others to reuse the data, but only when giving credit.
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here.	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	There are no additional costs expected.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Ellen Van Houtven
Who will manage data storage and backup during the research project?	Ellen Van Houtven
Who will manage data preservation and sharing?	Kathleen Beullens Desiree Schmuck
Who will update and implement this DMP?	Ellen Van Houtven