Unraveling Response Processes in Daily Life

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

In the original project proposal, we have given a concise overview of the data aimed to obtain. For convenience, this is repeated here.

Workpackage	Objective	Data
WP 1 Cognitive interview afterESM	Probe adherence to timeframe in momentary and short retrospective ESM affect items	Interview data Audio Files + Transcriptions
WP 2 ESM + Cognitive Probes	Use ESM items and ask follow up questions on response process during ESM study	Audio Files + Transcriptions Affect measures across time
WP 3 Emotion Vignettes	Ask intensity and rating of emotion scenarios. Translate scenarios into numbers order them by intensity	Ratings of Scenarios Interview data of response process
WP 4 ESM study - Between person design	Test different rating scales on two different timeframes	Response scale answers Information on demographic variables and mental health of participants

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: ***N****(ew data)* *or* ***E****(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
Cognitive Interviews (recordings)	Audio files of the interviews	N	D	S	wav	< 100GB	
Cognitive Interviews (transcripts)	Transcripts of the audio of interviews	N	D	T	nvp	< 5GB	
Experience Sampling Data Current and Retrospective Timeframe	participants report their current affect and retrospective affect via a smartphone application m-Path	N	D	N	csv	< 1GB	
Experience Sampling Data Affect	participants report their current emotional/ cognitive state, behaviour, and context throughout their daily life via a smartphone application m-Path 9 Days	N	D	N	csv	< 1GB	
Experience Sampling Audio Recordings	participants provide audio descriptions of how they filled out the questionnaires on the smartphone for 7 days	N	D	S	wav	< 100GB	
Experience Sampling Audio Transcripts	Transcripts of the audio description	N	D	T	nvp	< 5GB	
Affective scenario assessment ratings	Ratings of hypothetical scenarios, how they or others would feel	N	D	N	csv	< 1GB	
Affective scenario assessment labels	Participants provide labels for the affect intensity of scenarios and of endpoints of the scale	N	D	T	csv	< 1GB	
Cognitive Probes (recording)	Reasoning process, examples of higher/lower intensity scenarios	N	D	S	wav	< 100GB	
Cognitive Probes (transcript)	Transcript of cognitive probes interview	N	D	T	nvp	< 5GB	
Experience Sampling Data Rating Scales	Experiment between persons comparing different rating scales, numeric rating of affect on two response formats, collected via m-path	N	D	T	csv	< 1GB	

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:

N/A

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 below)

Ethical issues have been carefully considered as human participants will be directly interacting with the researchers in cognitive interviews. Information about thoughts, emotions and behaviors in daily life will be assessed quantitatively and qualitatively and quantitative evaluations. Participants will be informed and consent will be obtained. Participant's identity will be protected with in the app with an anonymous code. Pseudonymization measures will be applied to questionnaire and interview files, to protect participants' identities and personal data. As audio recordings may contain identifiable information, they will be transcribed and coded by the researcher and will not be shared with other

researchers. Ethical approval for the above mentioned datasets is still to be obtained. SMEC number will be provided after seeking ethical approval.

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)

Yes, Personal data like emails, and phone numbers will be included in the data. SMEC number will be provided after seeking ethical approval.

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.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

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).

- **File Naming and Organization**: All datasets will be named clearly and consistently, including metadata such as the study name, (for interviews) participant ID, and data type (e.g., ESM_Retrospective_vs_Momentary_Study.csv or Interview_ParticipantID_Date.wav).
- **Documentation**:
  - A README.txt file will accompany the dataset, providing a detailed description of the file structure, variable definitions, and any preprocessing steps.
  - A codebook (in .tsv format) will document all variables, their labels, and any transformations applied during preprocessing.
- **Reproducibility**: All preprocessing scripts for the ESM data (e.g., cleaning, scoring scales, and preparing analysis-ready datasets) will be version-controlled and made available on OSF or GitHub
- **Personal data** required for financial payments and informed consent will be stored in a secure, encrypted folder accessible only to authorized personnel. Identifiers linking personal data to anonymized datasets will be stored in a separate, secure file.
- **Backup and Archiving**: All data will be backed up regularly on institutional servers and external storage solutions, ensuring long-term preservation.

Interview Data

- Audio recordings will be anonymized by replacing participant identifiers with unique codes. Files will be named systematically by date and ID
- Transcripts will be stored in .nvp NVivo-format and linked to corresponding audio files with a codebook
- A README.txt file will explain the structure and content of the raw and processed data.

- Details about transcription preprocessing (e.g., tools used like NVivo) and conventions (e.g., handling of pauses, non-verbal sounds) will be included.
- A codebook will outline the coding scheme, including definitions of codes, examples, and interrater reliability metrics.

ESM Audio Data

- Audio files will follow the same anonymization and systematic naming conventions as the interview data.
- Transcriptions and participant-provided audio descriptions will be stored with clear metadata linking them to the relevant ESM entries.

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.

- No

There is no formally acknowledged metadata standard specific to our discipline, so we will rely on the DDI standard (Data Documentation Initiative) for the description on the project level. All information will be provided in txt, csv, or word-file. Some pseudonomyzed data sets (e.g., the data from the lab experiments) will be shared through the Open Science Framework (OSF) where they will be stored on servers within the EU.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- OneDrive (KU Leuven)

The data from this project will be stored on OneDrive for convenient access and on Large Volume Storage for long-term preservation. We will process the data exclusively on computers with encrypted hard drives. Data associated with publications may also be published on open data repositories such as OSF.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

To secure our data, we store it within KU Leuven's Large Volume Storage and OneDrive, both compliant with high and medium confidentiality standards, respectively. KU Leuven's ICT solutions adhere to stringent university-wide information security protocols. Access permissions for the raw network storage managed by the faculty's ICT service are strictly regulated, delegated, and audited by designated data managers who are trained for this role, regardless of their IT expertise. For the code files, which include pseudonyms, access is tightly controlled by the Principal Investigator. Researchers involved in the project are granted access exclusively to pseudonymized data, ensuring that personal identifiers are not disclosed.

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

The anticipated costs for data storage and backup for the duration of the research project are estimated at €2378.5, based on a block 5TB storage for 5 years. The costs for data storage are covered by the research group of quantitative psychology, with whom this project is associated with.

Data Preservation after the end of the Research Project

Which data will be retained for 10 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
- Certain data cannot be kept for 10 years (explain below)

Most data will be retained for 10 years to support reproducibility and potential follow-up research. However, personal data that is not essential for long-term analysis will be deleted when it is no longer necessary: IP addresses collected during online surveys, used to ensure each participant is unique, will be deleted immediately after verification. Contact details for participants will be deleted once they are no longer needed for ongoing communication or study invitations.

Where will these data be archived (stored and curated for the long-term)?

- Large Volume Storage (longterm for large volumes)

A data archive will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform KU Leuven policy. Some data will in addition also be available on the OSF and EMOTE servers (although we cannot guarantee a 10 years period). In the longterm the data will be

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

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 restricted data (upon approval, or institutional access only)
- Yes, as embargoed data (temporary restriction)

For the quantitative data, pseudonymized versions will be made open subject to an embargo period to secure the researcher's time to fully analyze the data and publish the findings. This measure is to prevent premature dissemination of the data which could lead to potential intellectual property issues or misappropriation of the research concepts by third parties.
Due to the sensitive nature of the interview data and difficulty of handling the information for non-familiar third parties, this data will not be made publicly available but can be available upon request, ensuring the protection of participant privacy and data confidentiality.

If access is restricted, please specify who will be able to access the data and under what conditions.

Participants' personal information (e.g., contact information, names, etc.) will never be shared.

Audio files will never be shared.

Embargoed data (ESM measures) will be available after publication of all planned research projects.

Embargoed interview data will be available after personal contact. This will require a personal data sharing agreement.

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

Qualitative data requires highly sensitive coding and effective pseudinymization.

Audio recordings will not be shared to preserve privacy of participants

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other data repository (specify below)

ESM and background data will be made available after a period of embargo. The data will be shared using various ways (e.g., OSF). Analysis syntax and code will be uploaded to the OSF and/or Gitlab.

When will the data be made available?

- Upon publication of research results
- Other (specify below)

The data will be made available upon request. Because the data contains a large amount of information from participants, the data will only be shared with individual researchers. We want to ensure the data is not used for non-scientific purposes (e.g. training large language models).

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)

Codebooks, syntax, and preprints will be made publicly available on the Open Science Framework.

When doing so, a CC-BY 4.0 license is provided.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

For the newly collected data, a permanent identifier is added to data upon deposit in a repository.

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

There are no costs

Responsibilities

Who will manage data documentation and metadata during the research project?

The principal investigator (LC)

Who will manage data storage and backup during the research project?

The principal investigator (LC)

Who will manage data preservation and sharing?

The principal investigator (LC)

Who will update and implement this DMP?

The principal investigator (LC)

Created using DMPonline.be. Last modified 25 February 2025

7 of 7