



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

ABSTRACTION: UNLOCKING MEANING FROM EXPERIENCE THROUGH LANGUAGE

(GRANT AGREEMENT: ERC-2021-STG-101039777)

UNIVERSITY OF BOLOGNA, ITALY.



European Research Council
Executive Agency

Established by the European Commission

Horizon Europe

D1.1 - Data Management Plan

Project Acronym**Project Number**

ABSTRACTION	101039777
-------------	-----------

Project start date: 1 June 2022

Project duration: 60 months

PI: Marianna Bolognesi (m.bolognesi@unibo.it)

Project websites: <https://www.abstractionproject.eu/>

<https://site.unibo.it/abstraction/en>

Document History

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
1	30/10/2022	Marianna Bolognesi	-

Data Management Plan (DMP) Updates

This document will be updated whenever important changes in the data or the data management policy occur.



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Content

The Data Management Plan (DMP)	1
1. Data Summary	1
2. FAIR Data.....	5
2.1 Making data findable	5
2.2 Making data accessible	6
2.3 Making data interoperable	7
2.4 Increasing data re-use	7
3. Other research outputs.....	8
4. Allocation of resources	8
5. Data security	9
6. Ethical and legal aspects.....	10
7. Other issues.....	11
Annex I: "README" file template	12
Annex II: Informed Consents for participation to research and Data treatment information	13

The Data Management Plan (DMP)

This DMP is a document that provides details on all the research data collected and generated within the ABSTRACTION project. It explains the way research data are handled, organized, licensed and made available to the public, and how they will be preserved after the project is completed. The DMP also provides motivations when versions or parts of the project research data cannot be openly shared on account of third-party copyright issues, confidentiality or personal data protection requirements or when open dissemination could jeopardize the project achievements.

This DMP reflects the current state of the art of the ABSTRACTION project. However, the details and the final number of data sets may vary during the research project. The variations will be recorded in updated versions of this DMP.

DMPs are mandatory for Horizon Europe's projects.

1. Data Summary

The aim of the project ABSTRACTION is to study the cognitive representation/processing and communicative effects of words that vary in concreteness and specificity. To achieve this goal, a series of tasks for data collection are envisioned, and a series of existing lexical and textual resources will be used.

The envisioned tasks that involve human participants are hereby summarized (and we refer to these tasks and the related datasets in Table 1).

TASK 1 - Word Ladders

- Type: Language production task (generating word associations through the Word Ladders app)
- Participants: adults and children (age 10-14), English and Italian native speakers

TASK 2 - Judging texts

- Type: Reading and judging quality of short texts (on 5-point scales)
- Participants: adults and children (age 10-14), English and Italian native speakers

TASK 3a - Creating metaphors

- Type: Language production task (generating creative metaphors)
- Participants: adults English and Italian native speakers

TASK 3b - Judging metaphors

- Type: Reading and judging quality of creative metaphors from task above (on 5-point scales)
- Participants: adults English and Italian native speakers



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

We therefore envision the management of three types of data:

1. Human judgments collected through crowdsourcing tasks and through the Word Ladders app (see tasks description above, and description of the app in *section 3*)
2. Existing databases of human judgments and lexical variables (e.g., word frequency, word length, etc)
3. Existing texts, which will be used as input to collect human judgments in reading tasks, and as input to assess automatically their readability through algorithms.

About the **first type of data**, the procedures involving collection of human judgements have been previously discussed and agreed upon during the Grant Agreement Preparation phase with the dedicated Ethics Officer. In addition, the ABSTRACTION team hired an external and independent ethics' consultant, Dr Pinelopi Trillou, from [Trilateral Research](#), who will assist the team in the data collection, especially for those tasks that involve also minors (Tasks 1 and 2, see below).

About the **second type of data**, the project will use existing resources of human-generated judgments of word concreteness, word average age of acquisition, and other variables that are commonly used when investigating linguistic and cognitive phenomena. These datasets will be used to run statistical analyses (correlations with the variable of interest, namely word specificity, and used as control variables in linear models). Existing resources on word properties such as these will be retrieved from the [Open Science Framework](#) platform (OSF), where these resources are commonly shared by the authors, in compliance with GDPR and FAIR principles.

About the **third type of data**, the project ABSTRACTION will use existing short texts (educational material, scientific articles, narrative, etc), to investigate text readability. These texts will be collected in compliance with copyright laws (see *section 6* for more details). The selection of the texts will be performed by PostDoc2, throughout year 1. The collection of texts, however, will not be publicly released. We will release the link to the original text, for replication purposes but we will not reproduce copyrighted texts online. These texts (or parts of them) will be used as inputs for the application of formulas of text readability. Therefore, they will be automatically scored in terms of their readability, using formulas based on the resources described above (datasets of word concreteness, etc). Moreover, the texts will also serve as input for surveys with human participants (Task 2), who will be asked to read, judge and give a score to texts in terms of their clarity and informativeness.

Table 1 summarizes and describes all datasets, the timeframe in which the dataset will be collected, and when it will be released, if applicable. Each dataset will be released on Open Science Framework, and linked to the publication in which the dataset is analyzed and discussed. Please note that in the "Type" column we refer to the three types of data listed at p. 1.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Table 1. Overview of datasets involved in the implementation of the project ABSTRACTION

#	Dataset title	Brief description	Type	Period of collection	Envisioned release
1	<i>ABSTRACTION. Task1 Specificity Ratings ITA collected through World Ladders</i>	Specificity scores associated to 30K Italian words, collected through an app (WordLadders game) from participants. The dataset will consist of a spreadsheet (csv) containing average specificity scores on a Likert scale for each of the 30K words.	1	Months 6-18	Month 27
2	<i>ABSTRACTION. Task1 Specificity Ratings ENG collected through World Ladders</i>	Specificity scores associated to 30K English words, collected through an app (WordLadders game) from participants. The dataset will consist of a spreadsheet (csv) containing average specificity scores on a Likert scale for each of the 30K words.	1	Months 6-18	Month 27
3	The adaptation of the Affective Norms for English Words (ANEW) for Italian	Human-generated ratings of word meaning on various dimensions (e.g., concreteness, familiarity, etc.)	2	Existing dataset on OSF: https://osf.io/eu7a3/	na
4	Italian age of acquisition norms for a large set of words (ItAoA)	Human-generated ratings of word meaning on various dimensions (e.g., age of acquisition, etc.)	2	Existing dataset on OSF: https://doi.org/10.17605/OSF.IO/3TRG2	na
5	Perceptual strength norms in Italian	Dataset containing behavioral data (reaction times collected during lexical decision task) for a large sample of Italian words	2	Existing dataset on OSF https://osf.io/zdg59/	na
6	The Calgary semantic decision project: concrete/abstract decision data for 10,000 English words	Dataset containing behavioral data (reaction times collected during lexical decision task) for a large sample of English words	2	Existing dataset on journal's platform https://doi.org/10.3758/s13428-016-0720-6	na
7	<i>ABSTRACTION. Task2 Text Readability Ratings ITA</i>	Dataset containing human judgments expressed on Likert scales on text readability for a sample of texts used as stimuli (in Italian)	1	Months 21-24	Month 30

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

8	<i>ABSTRACTION. Task2 Text Readability Ratings ENG</i>	Dataset containing human judgments expressed on Likert scales on text readability for a sample of texts used as stimuli (in English)	1	Months 21-24	Month 30
9	<i>ABSTRACTION. Task2 Corpus ITA</i>	A collection of Italian texts annotated with metadata (genre, target audience etc) to be used as stimuli for the previous task	3	Months 6-20	Month 30
10	<i>ABSTRACTION. Task2 Corpus ENG</i>	A collection of English texts annotated with metadata (genre, target audience etc) to be used as stimuli for the previous task	3	Months 6-20	Month 30
11	<i>ABSTRACTION. Task3a Creative Metaphors ITA</i>	A collection of creative metaphors generated by participants, in response to a specific prompt and a specific communicative goal, in Italian	1	Months 40-44	Month 52
12	<i>ABSTRACTION. Task3a Creative Metaphors ENG</i>	A collection of creative metaphors generated by participants, in response to a specific prompt and a specific communicative goal, in English	1	Months 40-44	Month 52
13	<i>ABSTRACTION. Task3b Creative Metaphors Evaluation ITA</i>	A dataset of crowdsourced judgments expressed on Likert scales about the quality and effectiveness of the metaphors collected in the previous task, in Italian	1	Months 44-48	Month 52
14	<i>ABSTRACTION. Task3b Creative Metaphors Evaluation ENG</i>	A dataset of crowdsourced judgments expressed on Likert scales about the quality and effectiveness of the metaphors collected in the previous task, in English	1	Months 44-48	Month 52

The research team has agreed to convert research data from proprietary formats to well-known and documented open formats in order to facilitate accessibility and reusability. All datasets will therefore be processed in xml, csv, or txt formats and released in these same formats, on the dedicated OSF repository. Data will be published in aggregated form only. Because the datasets will be released together with the dedicated publications (scientific articles) the envisioned dates of release may change, depending on the publication process.

2. FAIR Data

This DMP follows EU guidelines¹ and describes data management procedures according to FAIR principles². The acronym FAIR stands for “findable, accessible, interoperable and re-useable”, and describe data management practices that allow maximum knowledge circulation and return on investment.

2.1 Making data findable

Whenever project results are published the team members deposit and describe the relative underlying data sets in trusted³ data repositories; we plan to use the OSF platform, which is a OpenAIRE compliant⁴ repository and attributes persistent unique identifiers (DOI) to the deposited items. DOIs are automatically generated by the OSF platform, where we plan to deposit all the datasets produced by ABSTRACTION. The unique identifiers are then used to cite the data sets within all research publications.

The data sets generated by the project and stored on OSF are accompanied by standard descriptive metadata (brief description, date of deposit, contributing authors, funding grant project, licensing terms, persistent identifiers for the dataset) to ensure data sets indexing and discoverability. Specific keywords will be associated to each data set to enhance semantic discoverability.

Research data are organized in data sets, which are named collections of data units with the same focus and scope. In this DMP we set out common rules for data set naming, in order to improve data visibility, discoverability, citation and permanent online tracking.

The recommended data set title structure consists of:

*PROJECT ACRONYM. Task title or description. Additional information (if necessary).
Version number*

Example:

ABSTRACTION. Specificity Ratings ITA collected through Word Ladders.

The version number of the data set will be added at the end of the title in case of data revisions to help identifying the data set updates especially in repositories that do not track versioning automatically.

This DMP also recommends the following rules for file naming:

- for data set file(s)

*PROJECT ACRONYM_TaskNumber_Coverage or other content
specifications_Date(YYYYMMDD)_VersionNumber.fileExtention*

1 Horizon Europe Programme Guide: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf

2 The FAIR data principles (Force11 discussion forum), <https://www.force11.org/group/fairgroup/fairprinciples>

3 <https://doi.org/10.25504/FAIRsharing.g4z879>

4 <https://explore.openaire.eu/search/dataprovider?pid=10.25504%2FFAIRsharing.g4z879>



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Example:

ABSTRACTION_Task1_SpecRatingsITA_20220930.txt

- for readme file(s)⁵

*PROJECT ACRONYM_TaskNumber_Coverage or other content
specifications_Date(YYYYMMDD)_VersionNumber_README.fileExtention*

Example:

ABSTRACTION_Task1_SpecRatingsITA_20220930_README.txt

2.2 Making data accessible

At the time of publication of results, researchers deposit the project data that can be shared in a data repository in order to guarantee their discoverability, access and preservation beyond the project end. This will be done on the Open Science Framework platform: <https://osf.io/>

As a guiding principle, ABSTRACTION seeks to make all research data openly available as soon as possible and ensure open access in order to allow dissemination, validation and re-use of research results. To this purpose, all the files are converted to standard and well-documented open formats and the data sets are deposited on OSF together with all relevant documentation and explanation.

In case of copyright on raw data derived, collected or elaborated from pre-existing databases or from other original sources (i.e. papers, journal articles, book chapters, reports, video and audio sources), collected data will be made available if the reproduction and sharing are allowed by expressed permission of the right holders or by applicable copyright exceptions and exemptions. Otherwise, only aggregate data resulting from the analysis will be openly published. When the sources are freely available on-line in their original repositories, but direct reproduction is not allowed, a detailed account on how the data set was created from the original data will be provided, together with the specification of open repositories from where the original data sets are available.

Whenever possible, aggregated data will be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement. The raw data (e.g., the survey outputs of each individual participant) can be made available upon request directed to the PI of ABSTRACTION. Raw data will be shared in such a way that the participants' ID are replaced with codes indicating their native language and a number, e.g., participant ID: "ENG001"). Names/codes list used for pseudonymization are handled by the PI (password-protected and stored in a dedicated offline device) and never shared as to avoid re-identification (see *section 5*). OSF guarantees long term preservation and automatically attributes persistent

⁵ A "README" file is a document containing relevant information about data set authorship, terms of reuse and responsibilities, explaining data set content and structure, collection procedures and analysis (such as file specifics, methodologies, codebooks of variables, data sources, and further necessary notes). (See Annex I to visualize the suggested README file template).

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

unique identifiers to the archived data sets (DOI). It supports a variety of open licenses and uses a custom metadata schema, closely matching Dublin Core with a few additional elements from DataCite⁶. Metadata are released publicly; they are also sent to SHARE⁷, a OSF-run external service with its own metadata schema⁸.

2.3 Making data interoperable

All data sets will be described using standard descriptive metadata, in order to ensure metadata interoperability for indexing and discoverability. For each deposited data set, all relevant documentation explaining data collection procedures and analysis (such as methodologies, etc.) is made available along with the data, in order to guarantee intelligibility, reproducibility and the validation of the project findings. Moreover, the deposited documentation specifies the tools and software recommended to reproduce and reuse the data, where necessary (see also Tab. 2 for a list of tools and software enabling reuse of the dataset).

As mentioned, the team will convert all shareable data from proprietary formats to well-known and documented open formats (xml, csv or txt). This allows data exchange and re-use between different researchers, institutions, organisations and countries.

In case specific software is used during data processing, full explanation and instructions will be included in the deposited documentation. See Table 2 below for a summary of the tools and software necessary to reuse our data.

Table 2 – Summary of tools and software for enabling re-use of the data sets

Tools/software
R (open) - scripts to re-run our analyses will be shared in the OSF repository, for replicability
Python (open) - scripts to re-run our analyses will be shared in the OSF repository, for replicability

2.4 Increasing data re-use

ABSTRACTION distributes the shareable data by adopting licenses that allow re-use of the data and of the data sets in their entirety by other scholars and stakeholders. The data sets are made available, unless otherwise stated, under the latest available version of the Creative Commons License (to be decided whether CC 0 or CC BY) following the principle ‘as open as possible as closed as necessary’.

⁶ They are currently working on a project to better document their metadata and link them more closely to standard schemas.

⁷ <https://www.share-research.org/>

⁸ <https://github.com/CenterForOpenScience/SHARE/blob/HEAD/share/schema/schema-spec.yaml>



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

In general, data are made openly available to validate the research results immediately at the time of publication of public reports and scientific papers. Data are given full citation from official project publications and website, and they are made available through the OSF platform, which is compliant with OpenAIRE requirements⁹.

3. Other research outputs

The implementation of the project ABSTRACTION involves the development of a mobile application, through which linguistic data will be collected (datasets 1 and 2 in Table 1).

This research product is developed in its prototype form by the ABSTRACTION team and then an external gaming studio will improve it between February and May 2023, and will make it ready and available for download in EU-territory on Android and IOS stores. The application, which will be free of charge and without advertisements, is currently under development and the independent ethic consultant is working with the team for its implementation. The gaming studio that will help the team to make the app ready for Android and IOS stores and disseminate it on EU-territory will document all the measures taken to ensure that the data is collected, processed and stored safely and in compliance with EU regulations.

The application will be used to collect data and will be kept active during the whole duration of the project, even though (as described in Table 1) on month 27 we envision the release of the aggregated data collected up to that moment through the game. The data collected after month 27 will need to be evaluated. If the volume of the data will be as large as the primary dataset (released on month 27) we will consider releasing also this second dataset for replication studies. At the end of the project (month 60) the data collection through the app will be stopped. The application, wherever possible, will be managed in line with FAIR principles. More details will be added as the app development progresses.

4. Allocation of resources

Making data FAIR requires an investment of money and researchers' time. ABSTRACTION is not an exception.

The cloud storage solution adopted to share data among team members during the project implementation is OneDrive, provided by the hosting institution. The costs are covered by the hosting institution. The service for hosting the database containing the data coming from the app (cloud server) is currently ongoing (26th October 2022). The team is currently comparing service providers for hosting this database, making sure that the provider is GDPR compliant, and has a physical server within the EU territory and server

⁹ <https://explore.openaire.eu/search/dataprovider?pid=10.25504%2FFAIRsharing.g4z879>



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

maintenance is performed within the EU-territory. Updates will follow, once the service provider has been identified.

The hosting institution covers also the costs related to the [project website on the University portal](#)¹⁰. This website is addressed to the academic community. ABSTRACTION has also an [external website](#), linked to the university website, in which the objectives and tasks involved in the project are explained in a more accessible language (both, English and Italian), to a public of non-experts¹¹. The ABSTRACTION project dedicated a sum of 14500 euros for the development, hosting and maintenance of the website throughout the duration of the project, the development of the visual system and brand identity of the project (logo, visual assets, templates for communicating results through various channels, etc), cookies and data policies related to the website and support for communication strategies to be adopted on social media for sharing the project results through the 5 years of duration.

The costs related to data management and documentation, conversion of proprietary data files into open formats, and deposit procedures are estimated about 3-5% of the amount of Person-Months assigned to research activities.

Moreover, the activities related to the DMP are estimated to cost about 2 Person-Month (PI time), over the 5 years.

The costs of data preservation after the project ends are null because the chosen repositories do not apply fees for archiving and data curation.

Responsible for data management are data sets creators and in particular the PI.

5. Data security

Storage, retention and destruction of data during active data management: in compliance with EU and national legislation, research data are stored in computers, laptops, hard-drives, cloud storage systems (Microsoft OneDrive) of the research institution accessible through institutional password modified periodically (every 3 months in case of storage of personal data), and protected by regularly updated antiviruses.

None of the project data will be left inadvertently available by being left on desks or in unlocked rooms. All the research materials stored in computers are subjected to back up regularly in order to safeguard them from accidental losses.

As a general principle, all materials that could lead to an identification of the person (e.g., informed consent, names/codes list used for pseudonymization) are stored separately from research data that are

¹⁰ <https://site.unibo.it/abstraction/en>

¹¹ <https://www.abstractionproject.eu/>



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

handled by the PI. The matrix file containing the associations between participants' names and participants' codes will be accessible only to the PI, protected by password and stored offline in a dedicated hard disk, with a backup copy on a separate, offline device.

In particular the data are password protected (see above), accessible only to authorized the team members (authorized and controlled by the team leader) when they are no more necessary for the research, they are destroyed in according with art. 5 par. 1 letter "e" of GDPR (storage elimination).

The research data are contained on a separate file and do not contain any personal data.

All these resources are managed in compliance with University security policies, regularly subjected to backup procedures, and are accessible only to authorized members of the research teams, protected with University authentication credentials (see above).

A detailed description of the anonymization/pseudonymization techniques used to ensure personal data protection is described in Section 6.

6. Ethical and legal aspects

As previously described, the ABSTRACTION project will benefit from the support of an external and independent ethic advisor, as indicated in the Ethics Screening Report that preceded the Grant Agreement Signature. The appointed consultant, Dr Pinelopi Trillou from Trilateral Research, has relevant expertise in GDPR involving children as research participants and she will monitor the ethical concerns related to this project. Dr Trillou will draft the ethics reports in which she will discuss how ethical issues were handled by the research team during the reporting period, if any new issues emerged, and how they can be addressed (if applicable).

Informed consent forms for participation in research and for sharing data will be included in each data collection task dealing with personal data. In Annex II are reported the informed consent forms prepared for English speaking participants. These have been previously approved by the hosting institution Ethics committee, and by the Ethics Officer and reviewers who performed the ethics screening during the grant agreement preparation. Equivalent forms have been drafted in Italian, for Italian participants. The information included in the forms needs to be accepted by all participants before the data collection may start. Acceptance will be pursued by means of signature of participants or legal guardians, for those tasks that involve data collection in person, which is envisioned for tasks involving minors (these may be tasks described on rows 7 and 8 in Table 1), or by means of clicking on the relevant boxes for accepting terms and conditions of use (in the case of the data collected through the app). For those tasks that involve minors as participants, in collaboration with the independent ethics advisor we are working on the development of the best communicative (visual and verbal) strategies for informing minors about the scope of the research and the data treatment policy to be used in the app.

Data collected by the app will be associated to a user's ID, which is a made up code, randomly generated by the app. The user's identity will remain anonymous, because we will not collect the user's name, email nor IP of the device.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Participants' profiling is not an issue. Profiling is defined as “any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements”. Our analyses imply simple classification of users based on characteristics that they shared (age, sex, etc). These classifications will be used to acquire an aggregated overview of the users, without making any predictions or drawing any conclusion about an individual. In this case, the purpose is not assessing individual characteristics and is therefore not profiling, as indicated in the Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (retrieved here: https://ec.europa.eu/info/law/law-topic/data-protection_en).

For the data collection tasks involving adults, participants will be recruited through various channels and in particular through Prolific (<https://www.prolific.co/>), a GDPR compliant platform for the recruitment of research participants.

All data will be released in anonymized and aggregated forms.

Finally, for the construction of the corpus of texts (datasets 9 and 10 in Table 1), as described in the Resources section in the GA, a legal consultant may be appointed to discuss potential issues that may arise during the construction of the corpora. An effort will be made to use texts that are released online under CC-BY license, but some of the texts may be protected by copyright. Therefore, a strategy needs to be developed to ensure that the texts can be included in the corpora (e.g., payment of a fee, or collection of specific consent written by the copyright owner etc). These issues will be discussed with the legal expert of copyright law. We reserved a budget of 10K euros for this issue.

In any case, we do not plan to release these texts publicly. We will release the link to the original text, for replication purposes but we will not reproduce copyrighted texts online (see also *Section 1*).

7. Other issues

So far, there are no other data management issues to report.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)**Annex I: “README” file template**

A “README” file is a document that will be deposited with each dataset, containing relevant information about data set authorship, terms of reuse and responsibilities, explaining data set content and structure, collection procedures and analysis (such as file specifics, methodologies, codebooks of variables, data sources, and further necessary notes).

ReadMe file template

Data Set Title: “[insert title as defined in the DMP]”

Data Set Author/s: Name Surname (Affiliation), ORCID (if available);

Data Set Contributor/s: Name Surname (Affiliation), ORCID (if available);

Data Set Contact Person/s: Name Surname (Affiliation), ORCID (if available), email;

Data Set License: this data set is distributed under a [insert LICENSE]

Publication Year: [insert YEAR]

Project Info: [insert PROJECT ACRONYM] ([project full title], funded by European Union, Horizon 2020 Programme. Grant Agreement num. [insert grant agreement number]; [insert project website url])

Data set Contents

The data set consists of:

- 1 tabular quantitative file saved in .csv format: “title”
- 1 README file: “title”

Data set Documentation

Abstract:

[Insert data set abstract]

Content of the files:

- file [Insert filename] contains ...
- file [Insert filename] contains ...
- ...

Notes

[Related to the whole dataset]

Data sources

[Optional]

Methodologies

[If necessary to understand how to reuse data]

Codebook of variables

[If necessary to understand the meaning of the variables]

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)**Annex II: Informed Consents for participation to research and Data treatment information**

All the forms hereby reported have been approved by the hosting institution Ethics' Committee, as well as by the Ethics Officer of ABSTRACTION during the GAP phase.

All the forms hereby reported have been also approved in their Italian format, which will be used with Italian participants.

TASK 1:**ABSTRACTION. Unlocking meaning from experience through language.****1. Introduction**

Dear Madam/Sir,

You are invited to take part in a research conducted within the framework of the project "ABSTRACTION". Before deciding whether to participate, it is important that you have all the information necessary to adhere in an informed and responsible manner. We ask you to read this document and to ask the person who proposed this study all the questions that it considers appropriate.

2. Brief description and Objectives

The research project "ABSTRACTION" aims to explain how we extract meaning from experience through language. To achieve this, the researchers involved in the project aim to collect and analyze judgments of word specificity and word concreteness provided by English native speakers.

3. What does participation in the study involve?

Participation in this task of the research project involves the generation of "word ladders", namely, word associations produced in relation to an initial word. The words that will be associated, shall be linked to the initial words by relations of category inclusion.

4. Benefits, inconveniences and/or potential risks of participation

Participation in the research is voluntary and free. For the participants the collaboration does not involve any type of risk or discomfort. It could however be the case that some of the words suggested as prompts in the task may trigger in the participants negative emotions associated to past experiences.

5. Withdrawal from research

Participants have the right to withdraw at any time your consent to participation in this research, without notice or specific reason.

6. Information on research results and incidental findings

Participants (and their legal guardians, in case of minors) have the right to request information on the results of the research and to know any unexpected incidental finding emerged on their account.

7. Measures planned to protect anonymity

The processing of the collected data will be carried out in such a way as to eliminate any reference that may allow the reconnection of individual word associations to a specific person. The results of the research will be published in summary form and in no case any brief quotations will be ascribable to single persons.

8. Contacts

For any information and clarification on this research or for any need, please contact the PI of the researcher, Marianna Bolognesi (m.bolognesi@unibo.it), who is at your disposal for further information or clarification.

Informed consent to participate in the research

The undersigned _____, born on _____ in _____,

DECLARES

- to have read the aforementioned information sheet received; to have obtained satisfactory answers to potential questions from the staff in charge and to be ready to receive information on any unexpected incidental finding;
- to have understood that participation in the study is completely voluntary, that one can withdraw from the study at any time, without having to give explanations and without this involving any disadvantage or prejudice;
- to have understood the nature and activities that participation in the study entails and the related risks;
- to have understood that participation in this study will not entail the recognition of any direct or indirect economic advantage.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Consequently, the undersigned

[] CONSENT [] NOT CONSENT

to participate *In the case of minors, replace with "to the participation of my own daughter/son"* in the research, in the knowledge that such consent is freely expressed and can be revoked at any time without this leading to any disadvantage or prejudice.

(place and date)

(signature of the participant *In case of minors, replace with "legal representative of the participant"*)

(signature of the person collecting the consent)

TASK 2:**ABSTRACTION. Unlocking meaning from experience through language.****1. Introduction**

Dear Madam/Sir,

You are invited to take part in a research conducted within the framework of the project "ABSTRACTION". Before deciding whether to participate, it is important that you have all the information necessary to adhere in an informed and responsible manner. We ask you to read this document and to ask the person who proposed this study all the questions that it considers appropriate.

2. Brief description and Objectives

The research project "ABSTRACTION" aims to explain how we extract meaning from experience through language. To achieve this, the researchers involved in the project aim to understand how words that vary in concreteness and specificity affect the perceived clarity and informativeness of English written texts.

3. What does participation in the study involve?

Participation in this task of the research project involves reading short texts and judging their degree of clarity and informativeness (on a scale from 1 to 5).

4. Benefits, inconveniences and/or potential risks of participation

Participation in the research is voluntary and free. For the participants the collaboration does not involve any type of risk or discomfort. It could however be the case that some of the texts provided in the task may trigger in the participants negative emotions associated to past experiences.

5. Withdrawal from research

Participants have the right to withdraw at any time your consent to participation in this research, without notice or specific reason.

6. Information on research results and incidental findings

Participants (*and their legal guardians, in the case of minors*) have the right to request information on the results of the research and to know any unexpected incidental finding emerged on their account.

7. Measures planned to protect anonymity

The processing of the collected data will be carried out in such a way as to eliminate any reference that may allow the reconnection of individual word associations to a specific person. The results of the research will be published in summary form and in no case any brief quotations will be ascribable to single persons.

8. Contacts

For any information and clarification on this research or for any need, please contact the PI of the researcher, Marianna Bolognesi (m.bolognesi@unibo.it), who is at your disposal for further information or clarification.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Informed consent to participate in the research

The undersigned _____, born on _____ in _____,

DECLARES

- to have read the aforementioned information sheet received; to have obtained satisfactory answers to potential questions from the staff in charge and to be ready to receive information on any unexpected incidental finding;
- to have understood that participation in the study is completely voluntary, that one can withdraw from the study at any time, without having to give explanations and without this involving any disadvantage or prejudice;
- to have understood the nature and activities that participation in the study entails and the related risks;
- to have understood that participation in this study will not entail the recognition of any direct or indirect economic advantage.

Consequently, the undersigned

CONSENT NOT CONSENT

to participate *In the case of minors, replace with "to the participation of my own daughter/son"* in the research, in the knowledge that such consent is freely expressed and can be revoked at any time without this leading to any disadvantage or prejudice.

_____,
(place and date)

(signature of the participant *In case of minors, replace with "legal representative of the participant".*)

(signature of the person collecting the consent)

TASK 3a:

ABSTRACTION. Unlocking meaning from experience through language.

1. Introduction

Dear Madam/Sir,

You are invited to take part in a research conducted within the framework of the project "ABSTRACTION". Before deciding whether to participate, it is important that you have all the information necessary to adhere in an informed and responsible manner. We ask you to read this document and to ask the person who proposed this study all the questions that it considers appropriate.

2. Brief description and Objectives

The research project "ABSTRACTION" aims to explain how we extract meaning from experience through language. To achieve this, the researchers involved in the project aim to collect creative metaphors, freely generated by English native speakers on the basis of initial prompts.

3. What does participation in the study involve?

Participation in this task of the research project involves the generation of creative metaphors, starting from an initial prompt provided by the researchers (e.g., "SCIENCE IS LIKE...") and a communicative goal to which the metaphor may adhere (e.g., "please create a metaphor that has the goal to amuse your audience").

4. Benefits, inconveniences and/or potential risks of participation

Participation in the research is voluntary and free. For the participants the collaboration does not involve any type of risk. It could however be the case that some of the words suggested as prompts in the task may trigger in the participants negative emotions associated to past experiences.

5. Withdrawal from research

You have the right to withdraw at any time your consent to participation in this research, without notice or specific reason.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

6. Information on research results and incidental findings

You have the right to request information on the results of the research and to know any unexpected incidental finding emerged on your account.

7. Measures planned to protect anonymity

The processing of the collected data will be carried out in such a way as to eliminate any reference that may allow the reconnection of individual word associations to a specific person. The results of the research will be published in summary form and in no case any brief quotations will be ascribable to single persons.

8. Contacts

For any information and clarification on this research or for any need, please contact the PI of the researcher, Marianna Bolognesi (m.bolognesi@unibo.it [mailto:](mailto:m.bolognesi@unibo.it)), who is at your disposal for further information or clarification.

Informed consent to participate in the research

The undersigned _____, born on _____ in _____,

DECLARES

- to have read the aforementioned information sheet received; to have obtained satisfactory answers to potential questions from the staff in charge and to be ready to receive information on any unexpected incidental finding;
- to have understood that participation in the study is completely voluntary, that one can withdraw from the study at any time, without having to give explanations and without this involving any disadvantage or prejudice;
- to have understood the nature and activities that participation in the study entails and the related risks;
- to have understood that participation in this study will not entail the recognition of any direct or indirect economic advantage.

Consequently, the undersigned

CONSENT NOT CONSENT

to participate in the research, in the knowledge that such consent is freely expressed and can be revoked at any time without this leading to any disadvantage or prejudice.

(place and date)

(signature of the participant)

(signature of the person collecting the consent)

TASK 3b:

ABSTRACTION. Unlocking meaning from experience through language.

1. Introduction

Dear Madam/Sir,

You are invited to take part in a research conducted within the framework of the project "ABSTRACTION". Before deciding whether to participate, it is important that you have all the information necessary to adhere in an informed and responsible manner. We ask you to read this document and to ask the person who proposed this study all the questions that it considers appropriate.

2. Brief description and Objectives

The research project "ABSTRACTION" aims to explain how we extract meaning from experience through language. To achieve this, the researchers involved in the project aim to understand how people judge the quality and effectiveness of creative metaphors that have been generated in a previous study, by English native speakers.

3. What does participation in the study involve?

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Participation in this task of the research project involves reading and judging the quality, effectiveness, comprehensibility and aesthetic appeal of creative metaphors provided in a list. Judgments will be expressed on a scale from 1 to 5.

4. Benefits, inconveniences and/or potential risks of participation

Participation in the research is voluntary and free. For the participants the collaboration does not involve any type of risk. It could however be the case that some of the metaphors provided may trigger in the participants negative emotions associated to past experiences.

5. Withdrawal from research

You have the right to withdraw at any time your consent to participation in this research, without notice or specific reason.

6. Information on research results and incidental findings

You have the right to request information on the results of the research and to know any unexpected incidental finding emerged on your account.

7. Measures planned to protect anonymity

The processing of the collected data will be carried out in such a way as to eliminate any reference that may allow the reconnection of individual word associations to a specific person. The results of the research will be published in summary form and in no case any brief quotations will be ascribable to single persons.

8. Contacts

For any information and clarification on this research or for any need, please contact the PI of the researcher, Marianna Bolognesi (m.bolognesi@unibo.it), who is at your disposal for further information or clarification.

Informed consent to participate in the research

The undersigned _____, born on _____ in _____,

DECLARES

- to have read the aforementioned information sheet received; to have obtained satisfactory answers to potential questions from the staff in charge and to be ready to receive information on any unexpected incidental finding;
- to have understood that participation in the study is completely voluntary, that one can withdraw from the study at any time, without having to give explanations and without this involving any disadvantage or prejudice;
- to have understood the nature and activities that participation in the study entails and the related risks;
- to have understood that participation in this study will not entail the recognition of any direct or indirect economic advantage.

Consequently, the undersigned

CONSENT NOT CONSENT

to participate in the research, in the knowledge that such consent is freely expressed and can be revoked at any time without this leading to any disadvantage or prejudice.

_____, _____
(place and date)

(signature of the participant)

(signature of the person collecting the consent)



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Assent form for minors (task 2)

With this form, I give my assent to participate to the research project ABSTRACTION.

I, _____, born on _____ in _____,

DECLARE

- to have understood what is this research project about, as explained to me by the research team, and to have obtained answers to my questions from the researchers;
- to have understood that participation in the study is completely voluntary, that I can withdraw from the study at any time, without having to give explanations and without this involving any disadvantage or prejudice;
- to have understood the nature and activities that participation in the study entails and the related risks.

Thus, I

CONSENT

NOT CONSENT

to participate in the research, and I know that I can take back my consent at any time.

_____, _____
(place and date)

(signature of the participant)

(signature of the person collecting the consent)

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

DATA PROCESSING INFORMATION SHEET (for all tasks)

Information on the processing of personal data pursuant to art. 13 of Regulation (EU) 2016/679

You are invited to take part in a research conducted within the framework of the project "ABSTRACTION".

Pursuant to art. 13 of Regulation (EU) 2016/679 (General Data Protection Regulation), please note that Alma Mater Studiorum – University of Bologna is the Data Controller and will process your personal data in compliance with the requirements of Regulation (EU) 2016/679 (General Data Protection Regulation) and Decree 196 dated 30 June 2003 as amended (Data Protection Code).

PARTIES INVOLVED IN PROCESSING

- **Data Controller**
Alma Mater Studiorum – University of Bologna (registered office: via Zamboni 33, 40126 - Bologna, Italy; e-mail: privacy@unibo.it; PEC: scriviunibo@pec.unibo.it).
- **Data Protection Officer of Alma Mater Studiorum – University of Bologna**
Registered office: via Zamboni 33, 40126 - Bologna, Italy; e-mail: dpo@unibo.it

PURPOSES AND METHODS OF PROCESSING

Your personal data and, in particular, the data relating to *[task-specific aim to be added as in related Informed consent]*, will be processed by specifically authorised persons, both with and without the use of automated equipment, for the following purposes: understanding how different groups of English native speakers (children and adults, avid readers and non-avid readers, experts and non-experts) make sense of our daily experiences to construct meaning, and encode such meaning into words.

(A) PARTICIPATION IN THE RESEARCH AND RELATED OPERATIONS AND ACTIVITIES

If you give your consent to participate in the study in question, some of your personal data will be collected through the *[depending on the task, this can be the Word Ladders game or a survey]*.

The information that you will provide about yourself will be processed by the PI of the project, Dr Marianna Bolognesi (m.bolognesi@unibo.it) who will replace any identification data with a code.

The data will be collected without identifying data (name, surname); however, the registration of the IP address through the software used for the data collection could lead to the possibility of indirectly bring back to you the answers provided. It should be noted, however, that for no reason the data will be processed by authorized personnel with the aim of proceeding to the mere identification of the participant.

The data collected and stored for the realization of the study in question, will be kept for the entire duration of the research project (until 2028). After this date, the file in which participants' names and codes are stored, will be destroyed.

(A1) INFORMATION CONCERNING THE DISSEMINATION

The data will be disseminated only in strictly anonymous and aggregated form, for example through scientific publications, statistics and scientific conferences.

(B) CONSERVATION FOR ADMINISTRATIVE PURPOSES

Your data will also be processed for administrative accounting purposes and kept for the time strictly necessary for the pursuit of these purposes, except for the ten-year period to ensure the fiscal, accounting and administrative requirements required by law and, possibly, longer terms, not determinable a priori, as a consequence of different

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

conditions of lawfulness of the treatment (for example, legal actions that make the treatment necessary for over ten years).

LEGAL BASIS AND NATURE OF THE PROVISION OF DATA

The legal basis of the treatments referred to in point (A) described above lies in the consent, pursuant to art. 6, first paragraph, lett. a) of the Regulation (EU) 2016/679 (General Regulation on Data Protection) and, in the case of special categories of personal data, of art. 9, second paragraph, lett. a) of Regulation (EU) 2016/679. The consent to the processing for the purposes described above is optional.

However, it is specified that the provision of data for the purposes referred to in points (A) and (A1) is not mandatory but is essential for the achievement of the purposes described. The refusal to grant them will not allow you to participate in the study in question.

The provision of data for the purposes referred to in point (B) is necessary to comply with legal obligations regarding the conservation of administrative and accounting documents. In this case the legal basis can be found in the art. 6, paragraph 1, lett. c) and e), of Regulation (EU) 2016/679.

RIGHTS OF THE DATA SUBJECT

Without prejudice to the limitations placed on exercise of the rights of data subjects by arts. 2-undecies and 2-duodecies of the Data Protection Code (Decree 196/03), you, as a participant in this study and therefore a data subject, may exercise the rights granted pursuant and consequent to arts. 15-21 of the Regulation, including the right to request access to your personal data and its rectification or erasure, as well as to restrict the processing of your data, object to its processing and request its portability.

Please note that any consent given by you is given freely and may be revoked at any time, without any penalties or adverse effects, and without prejudicing the lawfulness of processing based on the consent given prior to revocation.

Requests to exercise the above rights may be presented to Alma Mater Studiorum – University of Bologna by contacting Dr Marianna Bolognesi, m.bolognesi@unibo.it.

Lastly, should you believe that the processing of your personal data is in infringement of the provisions of Regulation (EU) 2016/679 or Decree 196/03 and subsequent amendments and additions, you have the right to lodge a complaint with the Italian data protection authority (known as the "Garante per la protezione dei dati personali"), pursuant to Article 77 of the Regulation (UE) 2016/679, or to refer to the appropriate courts (art. 79 of the Regulation).

Consent for the processing of personal data

I, the undersigned _____, born on _____ in _____

pursuant to the provisions of Regulation (EU) 2016/679 and Decree 196/2003 and subsequent amendments and additions and having read the above "Information on the processing of personal data".

give consent

deny consent

for the processing - NECESSARY for the purpose of participating in the study in question – of my personal data for scientific research and statistical purposes in the manner and for the reasons described in the section entitled "Purposes and methods of processing" (point A).

give consent

deny consent



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

to the processing and the publishing - NECESSARY for the purposes of participating in the study in question – of interviews without my identification data, with the methods and for the purposes described in point (A1).

give consent

deny consent

to the storage and further use - NOT NECESSARY for the purposes of participating in the study in question - of my personal data for the purposes and in the manner set forth in point (B).

Date

Signature
