

Data Documentation

Data documentation for the raw data.

The raw data are stored in Raw_Data/Data_MTurk_AN.csv (responses retrieved from MTurk), Raw_Data/Data_Social_Media_AN.csv (responses retrieved from Social Media), and Raw_Data/Data_UvA_Lab_AN.csv (responses retrieved from first year psychology students at the University of Amsterdam).

Participants that received the survey through social media, or participated via the university lab, could choose to take the survey either in Dutch or English. Here, we explain the data structure of the "Raw_Data/Data_UvA_Lab_AN.csv".

The structure of the data are as follows. Each row corresponds to the responses of one participant. The columns are explained below.

- columns 1 -- 17: standard Qualtrics output

General information:

- column 18: "**Q1.2**"
Informed consent. 1 = participant agrees, 2 = participant disagrees
- column 19: "**phd**"
Does participant completed a PhD in psychology. 1 = yes, 2 = no
- column 20: "**language**"
Language the survey will be displayed in. 1 = English, 2 = Dutch
- column 357: "**psych**"
Has the participant studies psychology. 1 = yes, 2 = no. Note that this column was relevant for a thesis project from one of our students.
- column 358: "**knowledge**"
Was the participant familiar with the ML2 and/or the SSRP project. 4 = yes, 5 = no
- column 361 & 362: "**preorder**" and "**sequence**"
Displays the order in which the study descriptions were presented to the participant
- column 365: "**Conditie**"
Displays the condition the participant was assigned to. 1 = Description Only, 2 = Description Plus Bayes Factor

Data columns (column 20 -- 672).

The data structure for the following columns is as follows. Each study (labeled X1 - X28; see DescriptionAndStatisticsPerStudy.csv to link the item number to the corresponding study) is represented by three columns in the data set. The columns include identifiers for the study conditions, that is either 'des.' if participants were assigned to the Description Only condition,

or 'bf.' if participants were assigned to the Description Plus Bayes Factor condition. This identifier is further denoted as 'conditionlabel'.

The three columns represent:

- the participants' understanding of the study description. These columns are labeled as Xstudynumber_conditionlabel_0. They take the values 'NA' if a participant understood the study description, or '1' if a participant did not understand the study description.
- the participants' binary prediction of the replicability of the study. These columns are labeled as Xstudynumber_conditionlabel_belief. They take the values '1' if a participant believes that the study will replicate, or '2' if a participant believes that a study will not replicate.
- the participants' confidence in their prediction. These columns are labeled as Xstudynumber_conditionlabel_conf_1. Values in these column can range from 0 (participant was uncertain) to 100 (participant was 100% certain).

In total, the participants could receive four versions of the survey (that are denoted as blocks), depending on the assigned condition and the language participants chose:

- column 21 to 104: Condition = Description Only, Language = English
- column 105 to 188: Condition = Description Only , Language = Dutch
- column 189 to 272: Condition = Description Plus Bayes Factor, Language = English
- column 273 to 356: Condition = Description Plus Bayes Factor, Language = Dutch

Note that the column names that correspond to the Dutch version of the survey include 'NL'.

Information relevant for further correspondence with participants:

- column 359 -- 360; 364, 365

Information displayed to participants at the end of the survey:

- column 363: "**TotalScore**"
Number of correctly predicted studies. This number was presented to the participants