

Codebook

Directories

- **raw:** Contains the raw experiment json data that has had the Prolific IDs removed (see `anonymise-raw-data` for how we did that)
- **clean:** Contains the cleaned experiment data in long format, where each row pertains to a trial of the experiment. Files are separated by participant, but there is also a file with all of the data files combined.
- **derived:** Contains various transformations of the clean data generated from the **analysis** folder.

Instructions

To get the clean data from the raw data, run `import-cleaning-script.R`. After the clean data is generated, switch to the **analysis** folder.

Data Dictionary

Below is the dictionary for all of the clean data files. The variables are the same regardless of if whether it is an individual participant data file or the `all_data` file.

Variable	Description	Units	Coding Notes	Other Notes
participant	Participant number	integer	shortened version of full anon ID	
ANON_PID	Anonymized Participant ID	string	UUID	Used to anonymize participants
session_number	Session number	integer	Numeric	Participants did the experiment across two sessions
claim_set	Claim set category	string	Factor	All of the claims were divided into two “sets”, one for each session, which was counterbalanced.
claim	Specific claim	string		The claim being evaluated in that trial
claim_type	Claim type	string	Factor	The specific type of claim (four types total)
broad_claim_type	Broad claim category	string	Factor	The broad claim type (whether the claim is knowable or unknowable)
source	Claim source	string	Factor	Origin of the claim (e.g., media, university)
pre	Pre-test belief rating	integer	Scale from 0-100	Rating before exposure to information
post	Post-test belief rating	integer	Scale from 0-100	Rating after exposure to information
original_stance_against_consensus	Stance against consensus before	boolean	TRUE/FALSE	Whether initial stance was against consensus

Variable	Description	Units	Coding Notes	Other Notes
original_stance_against_claim	Stance against claim before	boolean	TRUE/FALSE	Whether initial stance was against the claim itself
original_stance_claim	Stance on claim	string	Factor ('for'/'against')	Initial stance before exposure to other perspectives
changed_mind	Whether the participant changed their mind	boolean	TRUE/FALSE	If stance shifted post-test
pre_adjusted	Adjusted pre rating	integer	Scale from 0-100	Adjusted pre-test belief so that consensus' arguing against and consensus' arguing for the claim resulted in the same direction of belief updating
post_adjusted	Adjusted post rating	integer	Scale from 0-100	Adjusted post-test belief (as above)
update	Change in stance	integer	Difference between post-adjusted and pre-adjusted	
consensus	Type of consensus around claim	string	Factor	Indicates consensus type ('dependent', 'independent', 'contested')
side_A	Stance of consensus	string	Factor ('pro'/'con'/NA)	One side of the argument
side_B	Stance of any opposing sources	string	Factor ('pro'/'con'/NA)	Other side of the argument
nSources_A	Number of sources for Side A	integer	Count	Number of sources supporting Side A
nSources_B	Number of sources for Side B	integer	Count	Number of sources supporting Side B
prop_pro	Proportion supporting 'pro' stance	float	Range: 0-1	Fraction of sources supporting the pro stance
trial_accuracy	Trial accuracy	integer	Binary (0/1)	Extent to which the participant correctly identified the stance of the consensus
total_duration	Total trial duration	float	Minutes	Time taken to complete the experiment session
stances	Stances of the tweets	string	Coded string	Sequence of stances of the tweets
tweetOrder	Order of tweet exposure	string	Coded string	Order in which tweets were seen
index	Trial index	integer	Numeric	Index number of the trial