Dataset Requirements: We are looking for datasets wherein participants see multiple cognitive or social stimuli (loosely defined), and the data is preserved with each original item-answer. For example, you may have participants rate 25 items on their pleasantness. If the data contains each rated item for each participant (i.e., not averaged across items), this data would be an appropriate dataset for our project. Note that it does not have to be your data, but you may know an appropriate dataset that is open source that we can use.

Project/Data Title: Cue-word valence

Project/Data Description: (200-500 words brief description of the theory/background for the data): Participants participated in a voluntary memory task, where they were provided with word-cue in response to which they were about to recall an autobiographical memory. After the task completion they classified the word-cue as either: positive, negative or neutral. The data set consists of 30 word-cues that were rated/classified by 142 participants.

Methods Description: (brief description of how the data was collected): Data was collected using the computer program where they were provided with a cue, one-by-one on the screen and rated its valence by clicking on the given button (positive, negative, neutral).

Data Location: (URL or upload on Canvas): uploaded on Canvas

Date Published: (YYYY-MM-DD): not published

Dataset Citation: (please include author information): The cues were used in study published here: Barzykowski, K., Niedźwieńska, A., & Mazzoni, G. (2019). How intention to retrieve a memory and expectation that it will happen influence retrieval of autobiographical memories. Consciousness and Cognition, 72, 31-48. DOI: <https://doi.org/10.1016/j.concog.2019.03.011>

Keywords: cue-word, valence, memory retrieval

Use License: open access with reference to original paper.

Geographic Description - City/State/Country of Participants: Poland, Kraków

Column Metadata: Fill in the chart below for each column of data in the dataset. Please note you can filter out columns that are not useful for this project.

|  |  |  |
| --- | --- | --- |
| Variable Name | Variable Description | Type (numeric, character, logical, etc.) |
| ID | Participants’ identification number | numeric |
| Word\_cue\_valence\_1  Word\_cue\_valence\_2  Word\_cue\_valence\_...  Word\_cue\_valence\_30 | Word-cue classified/rated as: 1 = positive  2 = negative  3= neutral | Numeric/label |

What columns should we use to simulate the data?

* Item labels are found: Columns from B (Valence\_1) to AE (Valence\_30)
* Variable(s) of interest are found: only variables of interest are in the file.

Goals: we will use this data to provide examples of our simulation process on how to determine sample size for a project based on item rather than participant. You can read about this idea here: <https://github.com/SemanticPriming/SPAML/blob/master/02_Power/power_aipe.pdf> We will use the example provided in this link as the main portion of the paper and then add your data as a vignette example to supplement the paper. You will be considered an author for completing this template worksheet (no coding skills necessary, we will do that part), and reviewing/commenting on the draft of the paper. Please email [007spaml@gmail.com](mailto:007spaml@gmail.com) if you have questions.