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:

Typicality, goodness, imageability, and familiarity of stimuli across 16 categories

Project/Data Description: (200-500 words brief description of the theory/background for the data)

As part of a larger project, we collected, among others, typicality, goodness, imageability, and familiarity ratings for several items for 16 semantic concepts (with each 5-33 items). (Additionally, we collected feature applicability ratings for these items, as well as inter-item pairwise similarity ratings, but these constructs seem of limited interest for the goal of this project.)

Methods Description: (brief description of how the data was collected)

The data collection took place in a large classroom where all the participants were present at the same time. The participants received a booklet with instructions on the first page, followed by four sheets with a semantic category label printed in bold on top. Each of the category labels was followed by a list of 5–33 items, referring to exemplars. The participants were asked to indicate, for every item in the list, how (e.g.) typical it was for the category printed on top of the page. They used a Likert-type rating scale, ranging from 1 for very atypical items to 20 for very typical items. If they encountered an exemplar they did not know, they were asked to circle it. Every participant completed (e.g.) typicality ratings for four different categories. The assignment of categories to participants was randomized. For every category, four different random permutations of the exemplars were used, and each of these permutations was distributed with an equal frequency among the participants. All the exemplars of a category were rated by 28 different participants.

Data Location: (URL or upload on Canvas)

<https://static-content.springer.com/esm/art%3A10.3758%2FBRM.40.4.1030/MediaObjects/DeDeyne-BRM-2008b.zip>

Date Published: (YYYY-MM-DD)

2008-11-??

Dataset Citation: (please include author information)

De Deyne, S., Verheyen, S., Ameel, E. *et al.* Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behavior Research Methods* **40,**1030–1048 (2008). https://doi.org/10.3758/BRM.40.4.1030

Keywords: Typicality, goodness, imageability, familiarity

Use License: Unknown. I do not know the correct name of the legal license, but the data can (and have) been used previously for academic purposes by researchers which were not the authors.

Geographic Description - City/State/Country of Participants: Psychology students at the University of Leuven, Belgium. Their ages ranged from 18 to 63 years.

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.) |
| exemplarTypicalityRatings.xls | Typicality: how typical is the item for the category? | Numeric |
| exemplarGoodnessRatings.xls | Goodness: how good an example is the item for the category? | Numeric |
| exemplarImageabilityRatings.xls | Imageability: how imageable is the item? | Numeric |
| exemplarFamiliarityRatings.xls | Familiarity: how familiar is the item? | Numeric |

What columns should we use to simulate the data?

* Item labels are found: see the .xls files in the exceldata\exemplar judgments\ folder
* Variable(s) of interest are found: see the .xls files in the exceldata\exemplar judgments\ folder

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.