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: HDHS.txt

Project/Data Description: (200-500 words brief description of the theory/background for the data): The data come from a study reported in Heyman, De Deyne, Hutchison, & Storms (2015, Behavior Research Methods; henceforth HDHS). More specifically, the study involved a continuous lexical decision task intended to measure (item-level) semantic priming effects (i.e., Experiment 3 of HDHS). It is similar to the SPAML set-up, but with fewer items and participants. The study had several goals, but principally we wanted to examine how a different/new paradigm called the speeded word fragment completion task would compare against a more common, well-established paradigm like lexical decision in terms of semantic priming (i.e., magnitude of the effect, reliability of item-level priming, cross-task correlation of item-level priming effects,…). Experiment 3 only involved a continuous lexical decision task, so the datafile contains no data from the speeded word fragment completion task (I can share those as well, if useful).

Methods Description: Participants were 40 students of the University of Leuven, Belgium (10 men, 30 women, mean age 20 years). A total of 576 pairs were used in a continuous lexical decision task (so participants don’t perceive them as pairs): 144 word–word pairs, 144 word–pseudoword pairs, 144 pseudoword–word pairs, and 144 pseudoword–pseudoword pairs. Of the 144 word-word pairs 72 were fillers, and 72 were critical pairs, half of which were related, the other half unrelated (this was counterbalanced across participants). The dataset only contains data for the critical pairs. Participants were informed that they would see a letter string on each trial and that they had to indicate whether the letter string formed an existing Dutch word or not by pressing the arrow keys. Half of the participants had to press the left arrow for word and the right arrow for nonword, and vice versa for the other half.

Data Location: <https://osf.io/frxpd/> [also contains R code with some AIPE implementation]

Date Published: 2022-02-04

Dataset Citation: Heyman, T. (2022, February 4). Dataset AIPE. Retrieved from osf.io/frxpd [based on Heyman, T., De Deyne, S., Hutchison, K. A., & Storms, G. (2015). Using the speeded word fragment completion task to examine semantic priming. Behavior Research Methods, 47(2), 580-606.]

Keywords: Semantic priming; continuous lexical decision task

Use License: CC-By Attribution 4.0 International

Geographic Description - City/State/Country of Participants: Belgium

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.) |
| --- | --- | --- |
| RT | Response time to the target in seconds | Numeric |
| zRT | Z-transformed target response times per participant | Numeric |
| Pp | Participant identifier (1 to 40) | Integer |
| Type | Whether target was preceded by a related prime (“R”) or an unrelated prime (“U”) | Character |
| Prime | Prime stimulus (in Dutch) | Character |
| Target | Target stimulus (in Dutch) | Character |
| accTarget | Whether response to target was correct (1) or not (0) | Integer |
| accPrime | Whether response to the preceding prime was correct (1) or not (0) | Integer |

What columns should we use to simulate the data?

* Item labels are found: Target
* Variable(s) of interest are found: zRT (and RT)

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.