Processing Pipeline

1. Search (February 9th, 2023)

* Search parameters: “predict semantic priming” OR “continuous lexical decision” “semantic priming”
* Google scholar search using: <https://harzing.com/resources/publish-or-perish/> software [google](https://docs.google.com/spreadsheets/d/1tFAJXWwwMTTd3r0BKhfsvW3hBQ_QyrQW47BVB1kM4V4/edit#gid=2057925199)
* Ebsco host using same search parameters [ebsco](https://docs.google.com/spreadsheets/d/1N_1SQW5XzdLnR5daxfEm3wqKXkyNE37vnZrKWjmjCR4/edit#gid=383501986)
* Duplicates were marked in individual files
* Should merge together and assign an ID for further coding [combined\_results](https://docs.google.com/spreadsheets/d/1lArGtRhLDW35SZ7-AFTvyucXM-B0ln6ngmb6j-x_eVw/edit?usp=sharing)

1. Test the coding of the literature form and the pipeline form (erin + tom) - code 10 to ensure that the forms make sense

<https://forms.gle/7yGbB1MGzrsN5siy5>

<https://forms.gle/MAy3rcmun5BfFb9d7>

Summary:

* Agreement was the same for the final papers to be coded
* Added an option on the second coding area that it was about semantic priming but further investigation reveals that it doesn’t have a CLDT or not predicting priming at the item level
* Clarified that results should be on neurotypical adults

1. Inclusion coding (Tom) - a google form to organize and mark information about inclusion of the article. This document is formatted like the PRISMA image to make it easier to fill in.

<https://forms.gle/7yGbB1MGzrsN5siy5>

1. Information extraction (erin/Katja) - a google form to organize information about the processing pipelines. Only the files marked above will be coded.

<https://forms.gle/MAy3rcmun5BfFb9d7>

1. Information compilation

First set of coded answers:

| Pipeline Part | Options |
| --- | --- |
| Individual Participants | * Exclude due to missing data * Non-native speakers * Experiment errors |
| Individual Stimuli | * Missing data on the items * Missing data on matched variables |
| Individual Trials | * Only correct targets * Both prime and target are correct |
| Outlier Participants | * RTs are long (2/2.5SD above) the condition mean * RTs are short (2SD below) the mean * High error rate (10%/20%) or low accuracy (<50/64%) * LogRT did not correlate with average LogRTs |
| Outlier Stimuli | * Target had a RT 2.5SD above rest * Low accuracy (<50%) * People didn’t know the words * Random pairing caused the unrelated pair to have higher similarity than related pair |
| Outlier Trials   * Exclude * Replace with highest value | * RTs too short (150, 200, 250, bottom 5%) * RTs too long (1500, 2000, 3000, 4000, top 5%) * RTs by individual 2/2.5/3SDs above/below mean * RTs by item 2/3SDs above/below mean * Van Selst and Jolicoeur (1994) |
| Missing data | No common rule, generally excluded when the data is missing due to not being able to match to another database |
| Transformations | * Log * Z-score transform * Both log and Z together * None |
| Order | * Generally: removed incorrect trials, looked at accuracy to exclude people, looked at RTs for people or trial exclusion |
| “Priming Scores” | * ANOVA/Regression with condition predictor * MLM * Subtract by item |

1. Expert Survey