Does priming influence word association and semantic fields?

This paper investigates how priming may affect people's semantic fields. A semantic field is a set of words related in meaning (Collins & Quillian, 1969), and this study is motivated by curiosity towards how people are affected by their surrounding environment and how their semantic networks function with regards to knowledge representation (Meyer et al.,1971). Semantic priming is the observed effect in response to a target word when preceded by a semantically related priming word, compared to an unrelated word. (Jong-Sun Lee et al., 2014).

Furthermore, this paper will contribute to the knowledge about how humans are influenced by priming, specifically which words people associate with the word "Garden", and how two different priming conditions: Botanical and Crops, may prime them in a direction to a specific semantic network(Katharina Sass et al., 2009). Based on this curiosity, the two following hypotheses have been made:

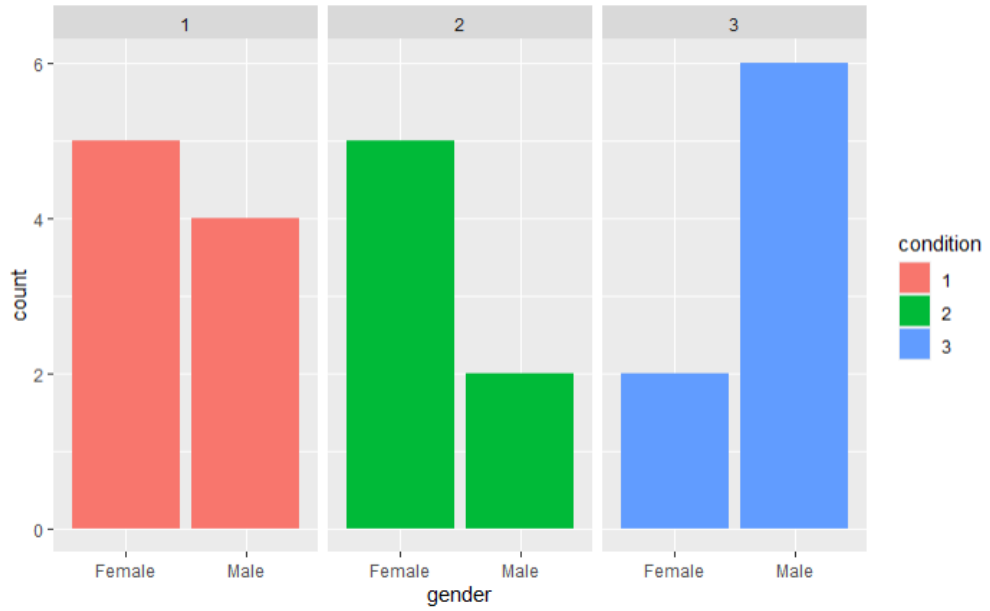
The Null-Hypothesis (H0): there is no significant difference in the semantic fields between the control condition and either the botanical priming condition or the crops priming condition

The Alternative Hypothesis (H1): there is a significant difference in the semantic fields between the control condition and either the botanical priming condition or the crops priming condition

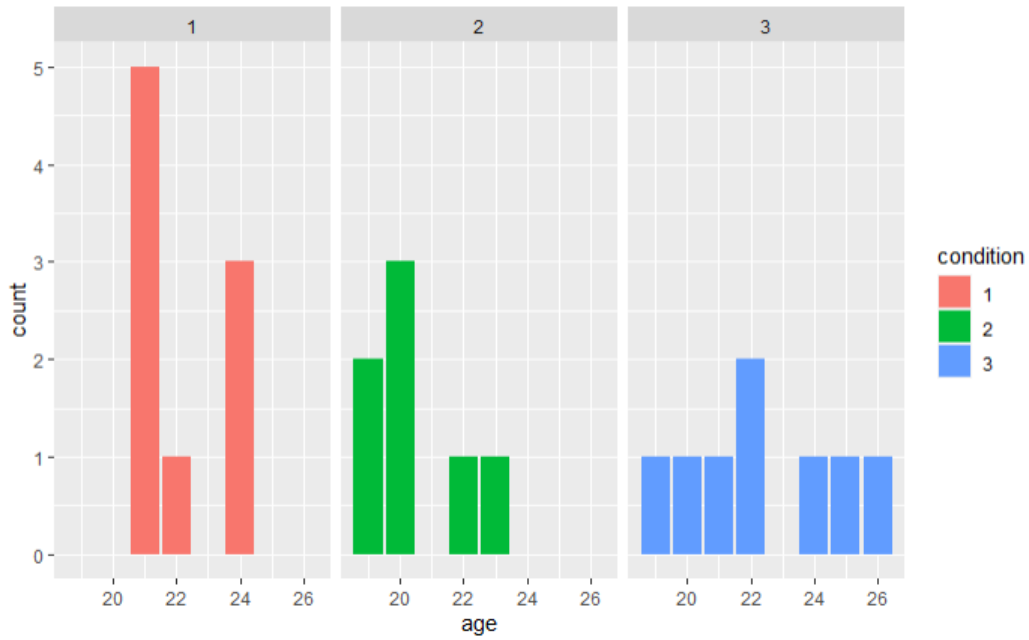
The experimental procedure to test the hypotheses, will be for participants to be exposed to either the control condition or one of the two priming conditions, and then they will have to write down as many words as possible that they associate with the stimuli presented to them.

Methods

***Participants*:** The experiment had 24 participants that were a mix between native and non-native Danish speakers, both female and male. All participants were university students, whose age varied between 19 and 25 years, with a mean age of 22 and a SD of 1.97.



Plot 1: Distribution of gender in the three conditions



Plot 2: Distribution of age in the three conditions

***Materials/Stimuli*:** The experiment had three conditions:

1. No priming; the only word that the participant saw was "Garden". This was the control condition.
2. Botanical priming; here the participants were given the hint "e.g. tree & flower".
3. Crops priming; here the participants were given the hint "e.g. carrot & tomato".

The priming categories, "Botanical" and "Crops" were chosen as conditions as they vary in their respective semantic distance from the primary semantic field "Garden". The chosen botanical example words had a semantic distance to "Garden" of M = 99.7, whereas the crops example words had a semantic distance to "Garden" of M = 144.2. These semantic distances were estimated using the Euclidean distance method.

Thus, to accept the alternative hypothesis H1, condition 2 and 3 must show a priming effect on the participants' semantic fields, for there to be a significant semantic distance compared to the control condition.

Procedure: The experiment was conducted in PsychoPy and the script was in English.

First, the participants were presented with a dialogue box, where they had to write their participant ID, age, gender and native language. In the dialogue box, the conductors of the experiment also chose one of the three conditions. Then the participants were presented with an introduction text.

Next, the participants were presented with an instruction text, which stated:

If condition = 1: *"Please write words you associate with the word: "Garden""*

If condition = 2: *"Please write words you associate with the word: "Garden" e.g., tree & flower"*

If condition = 3: *"Please write words you associate with the word: "Garden" e.g., carrot & tomato"*

The instruction text above was shown briefly for five seconds, which ensured that the participants only had limited time to come up with words associated with "Garden" before the word-association task began. Then the participants had 20 seconds to write as many words as possible that they

associate with the word "Garden". As this was a between-subjects independent measures experiment, each participant only went through one of the three conditions one time.

Analysis and results

The spread of the participants' semantic fields was measured as follows: For each participant the mean semantic distance of all the possible two-word combinations of the words they wrote was calculated. E.g. if the participant wrote the words "dirt", "tree" and "grass", the value would be: *mean(s("dirt", "tree"), s("dirt", "grass"), s("tree", "grass"))*, where *s(<word1>, <word2>)* represents the Euclidean distance of the vector mappings of word1 and word2 from the *EN_100k* database. This was chosen to be the outcome variable and the predictor variable was the different conditions of the experiment.

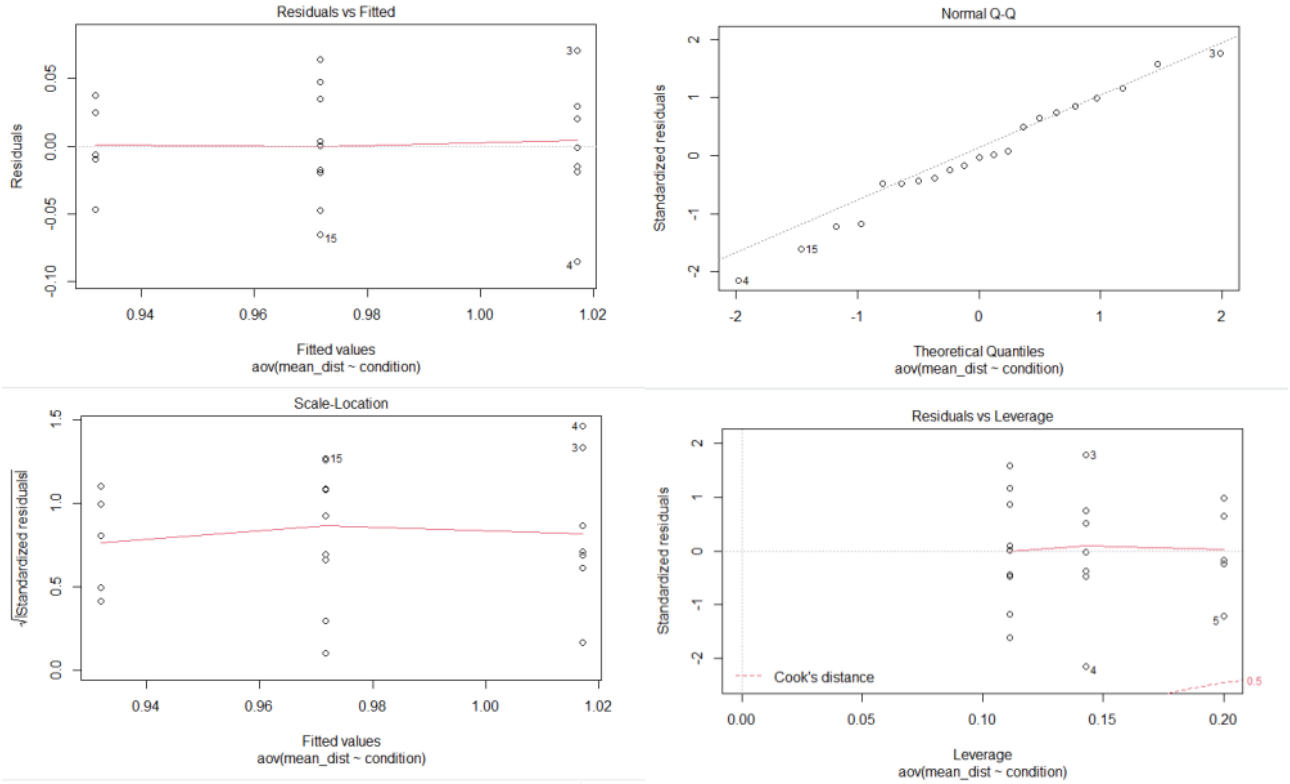
As the experiment was based on the participants manually writing the words, it inevitably led to misspellings and plural forms, which made it difficult to measure the semantic distance, as the calculations were limited to entries in the chosen semantic distance database. Therefore, misspellings and pluralism were removed from the dataset.

The data did not meet the assumptions of ANOVA. Therefore *log(), sqrt()* and inverse transformations were conducted on the data, but this proved to be ineffective. Three highly influential data points were removed from the data set. The influence metric used for removal of the aforementioned points was their Cook's distance with a threshold of 4/N, where N is the number of observations. All the data points removed were from condition 3, which reduced the number of participants in condition 3 from eight to five participants. Condition 1 had nine participants and condition 2 had seven participants.

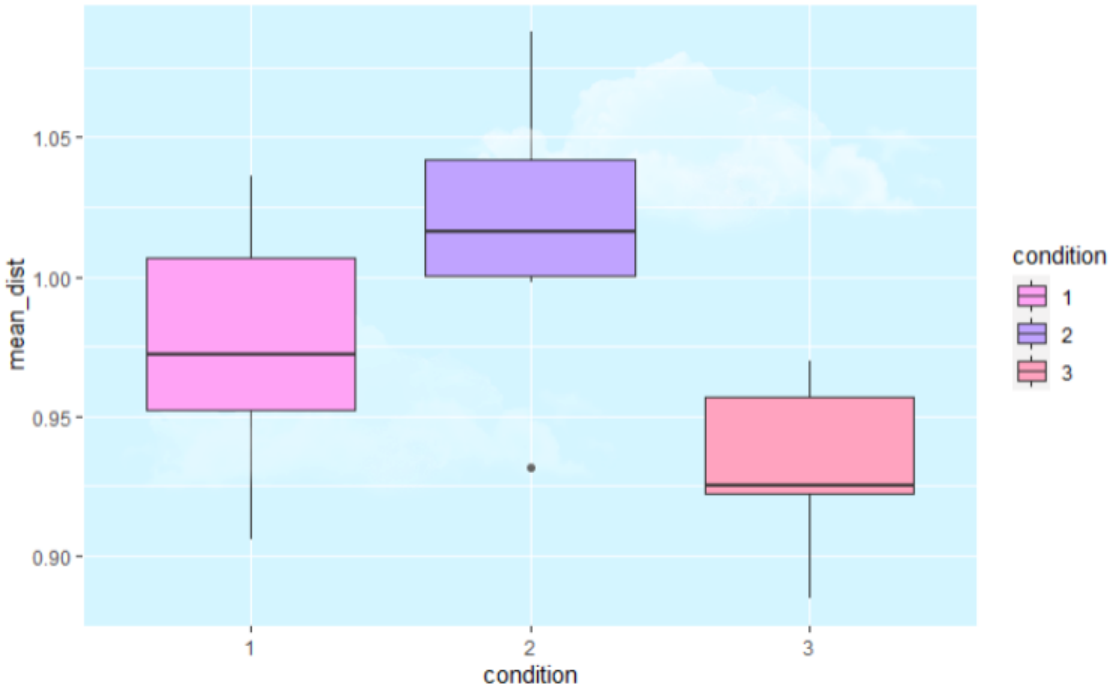
Then, the ANOVA test was performed, and the results showed that there was a significant difference between the three conditions of the experiment, $F(2) = 5.833$, $p = .0111$. Individual pairwise effects were tested with a post-hoc Bonferroni-corrected t-test. Only the difference between condition 2 and condition 3 was statistically significant ($p = .01$).

The mean and SD of semantic spread for the three conditions:

	Condition 1	**Condition 2**	**Condition 3**
Mean	0.97	1.02	0.93
SD	0.04	0.05	0.03



Plot 3: Plots over assumptions



Plot 4: Boxplot of the semantic distance from "garden" in the three conditions

Discussion

The results showed that there was no statistically significant difference in the semantic fields between the control condition and the priming conditions, meaning that the Null-Hypothesis H0 must be accepted and therefore the Alternative Hypothesis H1 was rejected. However, the results showed a statistically significant difference between priming condition 2 and 3, which was not expected. This could be explained by how the *EN_100k* database lists semantic distances, as it was not specifically made for "Garden" contexts. In the experiment, many of the participants submitted the word "flower", which could be interpreted as "flower" being highly associated with the word "Garden" in the participants' semantic fields. However, in the database the semantic distance for "flower" and "Garden" is rather large. This indicates that even if the words submitted by the participants are associated with the word "Garden" in their semantic field, the semantic similarity between them would be disregarded in the database which might have had an influence on the result. In a further study, it would be wise to generate a semantic distance database specifically made for garden words.

Also, "in general the more priming stimuli that the participant is presented to, the stronger the obtained priming effects" (Harry Reis, 2000). Hence, two priming words were chosen and not only one. One could argue that there should be more than two priming words in order for the priming effect to be stronger. However, the priming words were limited to two to ensure that the participants came up with new words associated with "garden" themselves, and not just repeated the presented priming words. This would be problematic for the analysis; were the participants primed or did they just remember the priming words and wrote them down.

Another choice that was made was having two priming conditions and not just one. This choice was made, as there were explicit hypotheses about the two conditions having differential effects. However, this was not analyzed, which could be a limitation to this study, since it could have had an significant impact on the results.

Conclusion

This study found that, in accordance with the Null-Hypothesis, H0, priming did not have a statistically significant effect on semantic fields compared to the control condition. However, there was a statistically significant difference between the two priming conditions, "Botanical" and "Crops", which indicated that priming might have an influence on word association and the structure of semantic fields, just not in the way that the hypotheses of this study anticipated.

References:

Harry T. Reis, Charles M. (2000), Handbook of Research Methods in Social and Personality Psychology - p. 264

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 90(2), 227–234. https://doi.org/10.1037/h0031564

Allan M. Collins, M. Ross Quillian (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior. 8(2), 240-247. 10.1016/S0022-5371(69)80069-1

Jong-Sun Lee et al (2014). The effect of word imagery on priming effect under a preconscious condition: An fMRI study. https://doi.org/10.1002/hbm.22512

Katharina Sass, Sören Krach, Olga Sachs, Tilo Kircher (2009). Lion – tiger – stripes: Neural correlates of indirect semantic priming across processing modalities, NeuroImage, 45(1), 224-236, https://doi.org/10.1016/j.neuroimage.2008.10.014.