ORGANISATION OF SEMANTIC KNOWLEDGE:

A COMPARISON BETWEEN THE SEMANTIC SYSTEM AND THE MENTAL MODEL

Julian Merten

Psychology

Supervisor: Prof. Dr. Frank van der Velde
2nd Supervisor: Dr. Martin Schmettow

University of Twente

Table of Contents

1. Abstract

Semantic knowledge refers to the storing, passing on and saving of information about concepts and the meaning of words. Since it is not yet clear how humans process and store information about concepts and semantics this paper focuses on the fit between activation in the brain and the people's mental model of semantic knowledge. Using card sorting and a questionnaire, we measured the people's mental model of semantical connections between concepts. The results of those were then compared to a voxel-wise model by Huth et al. (2016) that mapped out the semantic system in the brain. Our results indicate a clear semantical connection and a great overlap for two specific categories namely, *Time* and *Social*, whereas a little overlap and a weak cohesiveness for categories of *Visual* and *Outdoor*. Taken together, the outcome suggests there is an overlap between the people's mental model and the established semantic map for specific categories, however, they are not conclusive for all semantic areas researched.

2. Introduction

The meaning of language and how humans organise it is a crucial element in brain research. In 1972, when Endel Tulving established the theoretical notion that there is a difference between the semantic and episodic memory, which drew the focus on the research of semantic knowledge. His definition is that semantic memory is not connected to the memory of events. Instead, it is the storing, passing on and saving of information about concepts and the meaning of words (Binder et al., 2009; Tulving, 1972). For instance, knowing the names of colours or how many legs a dog has. More so, when we are trying to make a sandwich, it is necessary to know the consistency of the shape and colour of the peanut butter glass, the consistency of the peanut butter itself, the form and hardness of a knife and so on.

Recent developments of fMRI scans ensured the growth in this field of research and made it possible to get a detailed view of the processing and, more specifically, the categorisation of knowledge in the human brain. Three lines of research are relevant to this question. First, earlier studies suggest that a multimodal organisation of semantic knowledge in the brain might be possible. That means the organisation of semantic knowledge is split up into the brain areas that are responsible for the specific modality (Quillian, 1988; Warrington & McCarthy, 1983; Warrington & Shallice, 1984). That means, this body of research consists of the notion that semantics might be organised by the senses of the human body. Second, studies pointed into the direction to categorise according to a higher level organisation and semantic relatedness (Hart & Gordon, 1992). It is thought that the human brain processes and stores information according to its similarity in meaning. Moreover, more recent studies have shown a greater overlap in findings that conclude a semantic-based representation of knowledge in the human brain (Caramazza & Mahon, 2003; Huth et al., 2016; Huth et al., 2012; Just et al., 2010). Third, it is thought of as a combination of the two different organisational systems in the brain using the *hub* and *spoke* hypothesis. This theory states that information processing and knowledge storing happens in the designated *spokes* for the particular modality. However, it also suggests a modality independent knowledge storing. Both of which processes are being mediated through the *hub*(Chio et al., 2018; Ralph et al., 2017). Nevertheless, these three lines of research are not yet conclusive in their results. For now, most studies analysed information processing through means of fMRI scans and inferred the categories from the data to deduce a map of the semantics in the brain.

A new approach to this topic is to start with the mental model of individuals and to compare it with the physical regions of information processing and knowledge storing. This can enrich the current research to the extent that the semantic organisation is seen from the perspective of the active use of this knowledge by an individual and can be properly compared to brain imaging. Because as shown by (Unsworth, Sears, & Pexman, 2005) the development of the semantic system is considerably influenced by, for example, the culture and the individual experience (Patterson, Nestor, & Rogers, 2007). So, in order to analyse the semantic system, according to their processing location and the semantic relatedness, it is necessary to establish a ground. We address this problem by using card sorting to map out an average mental model of individuals. This is then to be compared with the categories found in the investigation of the semantic system by Huth et al. (2016). In there, a method was used, called voxel-wise modelling to inspect the location of the semantic information. This method helps to create a relatively reliable spatial layout of the activity in the human brain based on the blood-oxygen-level-dependent (BOLD) screening (Çelik et al., 2019; Naselaris et al., 2015).

1.1 Outline of the paper

This research conducted by Huth et al. (2016) is used as a point of comparison because it made a step forward towards the location of the human semantic knowledge in great detail and laid out a possible map of the semantic processing and storage (See the items, their categories and corresponding location us in this research in appendix C). The research inferred the map of the human information processing with the use of blood-oxygen-level-dependent (BOLD) screening, which locates the activation points of the brain while it is processing the semantics through language. However, there might be a discrepancy between the pattern based on the activity and the actual mental model, the semantic connection, of the participant. Thus, the first research question is:

“Is there a correlation between the mental model of participants and the categories established in the voxel model?”.

Additionally, the categories have been established by Huth et al. based on mathematical analyses and labelled by the researchers. This leaves room to question whether the selected categories would be similar to the ones the average person has.

Thus, the second research question is:

“To what extent are the groups, established by Huth et al., in line with people’s semantic relation?”

The benefit of this research might be a significant increase in the understanding of the semantic organisation and through giving a new perspective on the spatial processing of meaning in the brain.

To answer this question, we conducted a study consisting of a card sorting task and a follow-up questionnaire. This was to see how participants categorise words according to their meaning. After that, participants were asked to indicate in the questionnaire how closely, the by Huth et al. established category names, are related to the words used in the card sort. The experimental outcomes have been analysed and discussed at the end of this paper.

1.2 Card sorting

During the card sorting experiment, the participants are asked to sort paper cards into groups. On each of those cards was one word. The goal of the researchers was to find out how the participants see the relation of those words and how they group the terms given by Huth et al. (2016).

There are different kinds of card sorting methods, depending on the result that is needed for the research. For example, there is a difference between open and closed card sorting. For open card sorting, the participants are allowed to put the cards in any formation and call the groups according to their best fit. However, in closed card sorting, there are predefined groups that are provided by the researcher, and the participants have to sort these cards into one of the given groups. In that process, the participant is asked to do that in a semantic manner so that each group has an internal semantic connection (Schmettow & Sommer, 2016). For the current study, open card sorting has been used.

Furthermore, to achieve a greater level of detail during the card sort, a specific form of card sorting has been used in this study, called hierarchical card sorting. The reason for using it is that the semantical connection between two items can be measured more in depth so that in the end the results can be more reliable.

1.2.1 Hierarchical card sorting

Hierarchical card sorting is a form of card sorting and shows greater detail than the simple card sorting method. In a hierarchical card sort, the participant is asked to divide the groups into smaller subgroups. This, however, is not a necessity but rather a possibility for the participants if the they want to do that. The result of this is that the groups are sorted in nested sub-groups and, therefore, might show a finer level of distinction within the overall group. Furthermore, through this method of card sorting, it is possible to calculate the *distance* between each card, also called the Jaccard coefficient (Faiks & Hyland, 2000).

The Jaccard coefficient is calculated as follows. The participants were given the opportunity to sort the cards in hierarchies. In this study, the maximum number of allowed levels was three. This can result in different scores for the Jaccard coefficient for each hierarchical level. In the following example, *a* and *b* are the items to be measured, then one has to divide the number of groups *a* and *b* have in common (A AND B), by the number of groups that at least one of them is part of (A OR B) (Figure 1).

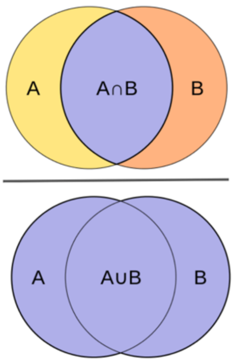

Figure 1: Graphical representation of the calculation for the *Jaccard coefficient*

Starting with the first level, participants sorted the cards in as many main groups as needed. Here are two possible outcomes, either the items are in the same group or in different groups. This results in two different *Jaccard coefficients*, {0, 1}(Real & Vargas, 1996; Schmettow & Sommer, 2016).

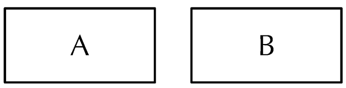

Figure 2: First level of the hierarchical card sorting.

In the second level of card sorting are additional outcomes possible. For instance, when one calculates the *Jaccard coefficient* for *a* and *b,* the calculation is as follows. The groups they have in common (A AND B) / the groups at least one of them is part of (A OR B). For this example, the result would be ⅓. Other possible outcomes are {0, 1} (Figure 3).

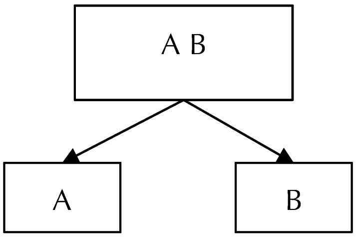

Figure 3: Second level of the hierarchical card sorting.

In the third level of the card sorting task are even more options possible. Aside from the previously mentioned {0, ⅓, 1}, {½, ⅕, ¼,} are possible outcomes. One example of the Jaccard coefficient of ½ is when one calculates the distance between the items *a* and *c* as they are in the same groups in the first level and in the second level. That means in four groups is at least one of them included. That results in an equation of 2/4, thus ½. The score of ¼ can happen on occasions when the distance between items such as *C* and *D* have to be calculated. Lastly, ⅕ occurs when the relation between the items *C* and *E* have to be calculated, as shown in figure 4.

1.
2.
3.

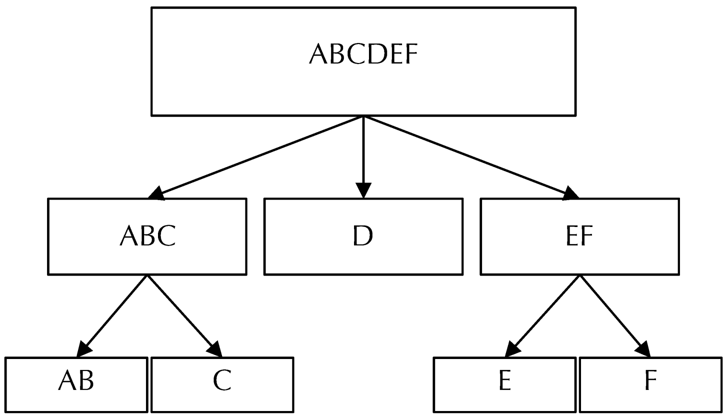

Figure 4: Third level of the hierarchical card sorting.

3. Methods

2.1 Participants

Thirty participants took part in this study, twenty-nine of them were students from the University of Twente. Eleven participants were male and nineteen participants were female, ranging from 19-27 in age, with an average age of 22.17 years old (SD = 1.599). Nineteen of the participants were Germans, nine were from the Netherlands, one person from Latvia, and one person from Norway. All participants signed the informed consent sheet with the right to withdraw, however, none of them withdrew from the study. The participants were recruited through the university's internal *Sona-Systems*. *Sona* is a system in which students can describe their research to find participants for their data collection. Students from the Behavioural, Management and Social sciences (BMS) can then login and take part in those studies to receive *Sona Credits*. Furthermore, it is necessary for students of this department to have 15 *Sona Credits* (1 Credit per 1-hour participation) in order to get the bachelor diploma at the end of their study. The estimated time for a participant to take for this card sorting task was about thirty minutes. Therefore, participants could earn 0.5 *Sona Credits* by taking part in this study. Next to the recruitment through *Sona-Systems,* more participants were gathered through convenience sampling.

2.2 Materials

Firstly, card sorting was conducted using the cards, which were prepared beforehand. The preparation consisted of selecting the terms appropriate for the card sorting study in the 3D voxel-model of Huth et al. (2016) (See Table 1). The criteria for the selection were reliability prediction from the model and had to be at least "Not bad, pretty reliable". Other sufficient reliability descriptions were "Good, very reliable" and "Excellent, extremely reliable", which were given on the website "https://gallantlab.org/huth2016/". The current study focused on a limited amount of five categories mentioned on the website, the categories used are "Social" (red), "Time" (brown-purple), "Outdoor" (blue), "Tactile" (blue-green) and, "Visual" (green). For each category, two voxels have been selected, one from the right and the other on from the left-brain hemisphere (See Appendix C). Within this voxel five words have been selected according to their colour which indicated the category they belong to. That made five words per voxel, two voxels per category and five categories, which is a total of fifty words.

Table 1

Items selected from the Huth et al. study sorted by their semantic category

Social	Time	Outdoor	Tactile	Visual
Son	Wednesday	Anticipation	Touches	Pieces
Wife	Monday	Tedious	Flame	Plastic
Father	College	Focus	Crushed	Gloves
Mother	Week	Relax	Thicker	Cardboard
Sister	Saturday	Deeper	Grip	Pairs
Daughter	Month	Emptiness	Grinding	Meters
Children	Date	Atmosphere	Mixing	Wheels
Relatives	Years	Cosmos	Smooth	Inches
Married	Since	Planets	Soft	Leather
Parents	Last	Objective	Solid	Cap

Secondly, a questionnaire was used in the study (See Appendix A). This questionnaire consists of two columns which contained in one column an item and in the other column the corresponding semantic category. Next to those two columns is a Likert scale, ranging from 1-5, whereas 1 means “highly related” and 5 means “highly unrelated”. This scale is often used in social sciences and ranges from 1-5, with 3 displaying a neutral or undecided attitude (Croasmun & Ostrom, 2011). The previously described fifty words were used in the questionnaire again with the addition of twenty filler words. Those filler words were selected on the same criteria as mentioned above except that they were from three different categories. The categories for the filler words are “Body parts”, “Violence”, “Numbers” (Table 2). Those words were then randomly assigned to the five groups in the card sorting. This resulted in a questionnaire of seventy items with fifty of which have a correctly assigned semantic category and twenty are incorrectly assigned according to Huth et al. (2016).

Table 2

Filler items selected from the Huth et al. study for the use in the questionnaire.

Numbers	Body parts	Violence
Cents	T-shirt	Deadly
Half	Tops	Severe
Million	Brand	Extreme
Six	Regular	Fatal
Plus	Dress	Infection
Score		
Points		
Quarter		
Extra		
Counting		

2.3 Procedure

2.3.1 Briefing

After the participants were willing to participate, the intention of this study was explained in order to understand the experiment.

The procedure was explained to each participant, with the option to ask any questions. Additionally, it was explained to each participant that they have the right to withdraw from the study at any given point in time and the chance to ask any question within and after the experiment. The prepared informed consent was given to each participant with all rights and information for the experiment provided, which were additionally explained for clarification by the researcher (See Appendix B). Furthermore, the participants were instructed to lay out the fifty paper cards in groups according to their opinion. However, it was made clear to each participant that the grouping should be according to the semantics of the words rather than any syntax. After that, the participants started to sort the cards until their satisfaction with the sorting task. When the participants finished, they were asked to subdivide the established groups into smaller subgroups within the main groups on a voluntary basis. After the second round of card sorting, they were asked to subdivide them even further into smaller sub-groups, also voluntarily. To capture the results, between the rounds, pictures have been taken with a common mobile device for further analysis. Thereafter, having satisfactorily established groups during the card sorting task, each participant was asked to fill in the questionnaire.

2.3.2 Debriefing

At the end of both tasks, the card sorting and the questionnaire, the researcher explained the detailed purpose of the study and asked for any other question. Additionally, for those students that signed up through Sona-Systems, the available credits were distributed to each student. Finally, it is repeated that the student has the right to withdraw from this study at any time without providing any reasons for the withdrawal.

2.3.3 Data Analysis

2.3.3.1 Card sorting.

The resulting data from the card sort were entered in spreadsheets, using a 50x50 table to clearly show the relationship between each item. The relationship is shown through a distance score called the *Jaccard coefficient*.

The *Jaccard coefficients* are put into the earlier mentioned spreadsheet, and thus, for each item is a corresponding number entered in the crossing cell. This is done for each participant.

After that, all cells get accumulated with the corresponding cells in all tables, resulting in a cumulative table. This table includes the accumulated *Jaccard coefficient* for each item of all participants. With the accumulated *Jaccard coefficient* table, the researchers can continue by using the *Vector Analysis*. This increased the preciseness of the analysis and has a more complex result than the standard card sorting analysis (Hierarchical Cluster Analysis) would have. In the *Vector Analysis* not only the highest *Jaccard score* between two items is relevant, but all the *Jaccard scores* the two items have with all the other times in the card sort. These scores can be used to calculate a distance between two items in terms of all *Jaccard scores* they have. The two items with the shortest distance are then clustered together.

The *Vector* Analysis was conducted with a tailored R script (See Appendix D) and resulted in a new cross-table, which is called a heatmap. This heatmap indicates the relation between two words by colouring the cell in a spectrum. This spectrum reached from yellow to red whereas the red colour indicates high relation and yellow low relation. After that has been done, the researcher can analyse the data and find clusters with enough justification based on the warmth (redness) of the cluster and the justification through logic and reasoning by the researcher.

2.3.3.2 Questionnaire.

The questionnaire focussed on the analysis of the previously established categories. This was done by comparing the scoring results of the participants with the assumed score based on the study by Huth et al. The mean of each item of the questionnaire has been calculated before taking the mean of all items split up in *filler items* and *regular items* to investigate whether there is a difference between those. Additionally, the standard deviation and variances have been calculated. Moreover, the means for each item has been calculated as well. Finally, the means, standard deviation and variances have been calculated between each category.

4. Results

3.1 Card Sorting

The analysis of the card sorting resulted in a heatmap showing the cells with a higher *Jaccard coefficient* (towards 1, redder) and the cells with a lower semantic connection (towards 0, yellower).

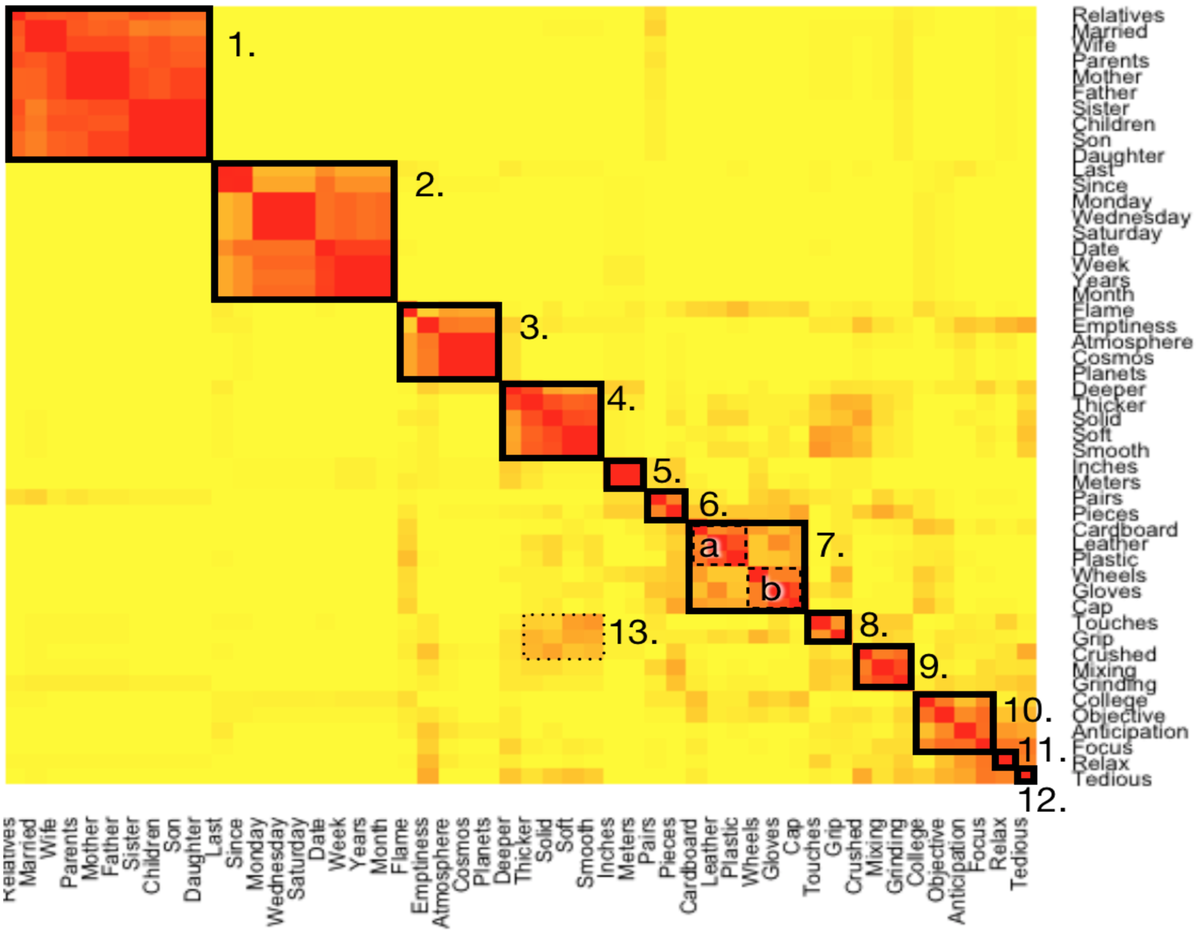

Figure 5: Heatmap showing the cluster from the card sorting based on vector analysis. It includes 12 clusters and one bleeding spot.

In figure 5 are 12 main clusters and 2 sub-clusters noted. Furthermore, within those 12 main clusters is one triad, two dyads and two single items. These clusters were created as follows. The first criterion of the creation of clusters happened through the visual proximity of the red cells seen in figure 5. The second criterion is based on the researchers semantic reasoning for the best fit.

Table 3

The clusters from 1-6 with the corresponding items based on the heatmap above.

1. Cluster	**2. Cluster**	**3. Cluster**	**4. Cluster**	**5. Cluster**	**6. Cluster**
Son	Wednesday	Flame	Deeper	Inches	Pairs
Wife	Monday	Emptiness	Thicker	Meters	Pieces
Father	Week	Atmosphere	Solid		
Mother	Saturday	Cosmos	Soft		
Sister	Month	Planets	Smooth		
Daughter	Date	Flame			
Children	Years				
Relatives	Since				
Married	Last				
Parents					

Table 4

The clusters from 7-12 and the bleeding spot with the corresponding items based on the heatmap above.

7. Cluster	**8. Cluster**	**9. Cluster**	**10. Cluster**	**11. Cluster**	**12. Cluster**	**Bleeding spot**
a	Touches	Crushed	College	Focus	Relax	Touches
Cardboard	Grip	Mixing	Objective			Grip
Leather		Grinding	Anticipation			Crushed
Plastic			Focus			Thicker
b						Solid
Wheels						Soft
Gloves						Smooth
Cap						

Additionally, there has been found a bleeding spot and single item clusters in the heatmap. First, bleeding spots seem to be clusters as well, however, their location is off the main diagonal where the other clusters are arranged. Secondly, those clusters are off the main diagonal because even though the items have a correlation with each other, it is not as strong as the ones on the main diagonal.

The first cluster includes ten items that correlate highly with each other, which can be seen due to the redness of the cluster (see Table 3). Furthermore, it also shows a semantic connection since all items describe statuses or persons within a family. These are for example items such as "*mother, father, son* and *married*". All of the items sorted into this cluster are the same items which have been put into the group *Social* in the research by Huth et al. (2016) (see Table 1).

The second cluster has a value of nine items, which include words such as "*last, since, monday, wednesday, month* and *years*" (see Table 3). These items generally resemble a high correlation throughout the whole cluster even though there might be differences. The highest cohesiveness within the group have the words "*week, month and year*" as well as *"monday*, *wednesday* and *saturday*" and similar to "*last* and *since"*. All items of this cluster are put into the same semantic group in the Huth et al. (2016) study, except the word "*college*" (See Table 1).

The third cluster consists of five items and contains the words "flame, *emptiness, atmosphere, cosmos* and *planets".* Although *flame and emptiness* have a weak relation, either of them is sufficiently correlated with the other three items of the group. In this case, "*emptiness, atmosphere, cosmos and planets"* share the same group according to Huth et al. (2016), which is called "*Outdoor"* with each other (See Table 1 and Table 3).

The next cluster, cluster number four contains five words, which are all relatively highly correlated with each other except the item *"deeper".* However, in the heatmap, it has a high connection to the item *"thicker"* and therefore can be counted to that cluster with *"solid, smooth* and *soft".* Even though "*deeper"* is in the category "*Outdoor*" based on the Huth et al. (2016) in the heatmap it shows a weaker connection to other items from the group "*Outdoor"* than it does for the items in cluster four (See Table 1 and Table 3).

Cluster five is a dyad and therefore consists of two items with a high correlation. These items are "*meter* and *inches"*. Since they do not have any other item nearly as highly correlated as these, they cannot be sorted into a bigger overarching group (See Table 3).

The sixth cluster is another dyad and includes the items "*pieces* and *pairs*". They are not as highly correlated as the dyad above, however, the semantic similarity of these two words is sufficient to sort them into a group (See Table 3).

Group number seven contains six items in total which are split into two sub-groups. The first sub-group 7a includes the items "*cardboard, leather and plastic"*. Whereas, the sub-group 7b combines the items "*wheels, gloves* and *cap"* (see Table 4). Due to the semantic relation of "*cardboard, leather* and *plastic"* with materials or surfaces and the redness within this triad, the subgroup was established (See Figure 5). Taken the whole group seven together, according to Huth et al. (2016), all of the mentioned items belong to the same semantic group, namely "*Visual*" (Table 1).

The eighth cluster is a dyad, which consists of the words "*touches* and *grip".* These two items seem to have a sufficient semantic relation and the necessary redness to group them together into a cluster (see Table 4 and Figure 5).

Cluster number nine is a triad and therefore consists of the three items "*crushed, mixing* and *grinding*". These items have a relatively high correlation. Additionally, the corresponding meaning within this triad is sufficient to cluster them together (see Table 4).

The 10th group contains four items in total. These items do not have such a high correlation as some other clusters have, however, putting the words into context, one can argue that there is sufficient overlap in its semantics so that it can be sorted into the same cluster. The corresponding items are "*college, objective, anticipation* and *focus*". This is partly supported by the categories established by Huth et al. (2016) in terms that "*objective, anticipation* and *focus*" belong to the group called "*Outdoor*" (see Table 1 and Table 4).

To those above-mentioned clusters are two more clusters to name. However, these are two single item clusters which do not seem to belong to another one. On the heatmap, these items are marked as cluster eleven and cluster twelve. The words corresponding to it are "*relax*" and "*tedious*" (See Table 4).

Cluster thirteen is apart from the main diagonal of the heatmap. This cluster is called a bleeding spot and includes "*soft, smooth, solid, thicker, touches, grip* and *crushed*". These words show a relatively high correlation toward each other even though they are off the main diagonal. Furthermore, the semantic relationship between those items is sufficient so that this can be evaluated as a bleeding spot. Furthermore, the items that are mentioned in this cluster also correspond to the category "*Tactile*" in the Huth et al. study (2016) (See Table 1).

3.2 Questionnaire

With the outlook to the research question: "Is there a correlation between the mental model of participants and the categories established in the voxel model?" the analysis focused on the commonalities and differences of the pre-established categories in the voxel model by Huth (2016).

According to this question, the following analyses have been conducted. The 5-point Likert scale was coded with the numbers 1-5 from *highly related* to *highly not related*, respectively (1 = highly related, 2= related, 3 = neutral, 4 = not related, 5 = highly not related). That means on the one hand, that the higher the score, the higher is the semantic difference between the item and the category. On the other hand, the lower the score, the lower is the semantic difference between the item and the category. First, the responses were sorted to whether it was a filler item or not. This has been done to compare the overall results between the filler and the regular items. This analysis resulted in a mean score for the *filler items* of *3.500,* a standard deviation of *0.756* and a 95% *Confidence Interval* between *3.169* and *3.831*.

Whereas the results for the regular items showed a mean of 2.329, a standard deviation of *0.720* and a *95% Confidence Interval* between *2.13* and *2.529*.

As next, each group within the regular items was analysed on the same statistics as before (Table 5). Here to point out are the results of the category "*Time*" which scored the lowest of all categories with a mean of 1.49, a standard deviation of 0.325 and a *Confidence Interval* of *1.04* and *1.94*. This shows a high relation of the items to the pre-established categories and internal consistency of the answers due to the lowest standard deviation. In contrast, the category "*Outdoor*" scored the highest in this questionnaire. Here, the results are a mean of *2.950*, a standard deviation of *0.826* and a 95% *Confidence Interval of 2.5* to *3.4 95%*. Therefore, together with the category "*Visual", a 95% Confidence Interval* of *2.4* to *3.29*, either of them lies within the *Confidence Interval* of the *Filler items*. That indicates the means of the categories *Social, Tactile and Time* are with a probability of 95% or more distinct from the ones of the *Filler items*. Whereas, the means of the categories *Outdoor* and *Visual* are not clearly distinct with a probability of 95% or more from random *Filler items* (Table 5).

Table 5

The words and their corresponding semantic categories used for the card sorting task selected from the study by Huth et al. (2016).

Categories	SD	M	95% Confidence Interval (CI)	
			Lower CI	Upper CI
Outdoor	0.826	2.950	2.5	3.4
Social	0.686	2.157	1.71	2.6
Tactile	0.588	2.220	1.77	2.67
Time	0.325	1.49	1.04	1.94
Visual	0.339	2.843	2.4	3.29
Total	0.720	2.329	2.13	2.529
Filler Items	0.756	3.500	3.169	3.831

The results from the cards sort and the questionnaire have some overlap but also differences in their results. Comparing the items of the category *Social* from the heatmap with the outcomes of the questionnaire, it seems that they are similar in their results. The heatmap shows that participants group the items together and similarly use the label *Social* to label them. Almost the same can be said for the category *Time* except that within that category the word *college* is put into a different group in the card sorting task by the participants. Nevertheless, the SD *0.335* within the category *Time* combined with the mean of *1.49* showed that similar to the card sorting task also the questionnaire had clear results.

However, when looking at the results of the *Tactile* groups, they do not show the same similarity as the two groups above. The heatmap does not have a single group representing the exact *Tactile* category. Yet, taken the bleeding spot and cluster four together one gets indications for the overall group and can compare it to the significant results from the questionnaire. The two other categories *Outdoor and Visual* both, do not show a cluster in the heatmap that is comparable to the category nor a significant result in labelling of the groups through the questionnaire.

Table 6

The items in each category with the corresponding mean per item within this category.

Social	**M**	**Time**	**M**	**Outdoor**	**M**	**Tactile**	**M**	**Visual**	**M**
son	2.567	wednesday	1.633	anticipation	3.367	touches	1.600	pieces	2.800
children	2.167	college	2.300	relax	2.000	crushed	2.800	gloves	3.600
mother	2.100	month	1.367	atmosphere	1.833	grinding	2.767	inches	2.600
wife	1.933	years	1.300	deeper	3.767	smooth	1.733	pairs	2.900
parents	2.200	since	1.500	planets	2.267	thicker	2.267	plastic	2.933
sister	2.267	week	1.233	tedious	3.800	solid	2.067	meters	2.667
married	1.833	saturday	1.333	focus	3.300	grip	1.667	leather	2.833
relatives	1.867	last	1.667	emptiness	3.500	flame	2.700	cardboard	2.400
daughter	2.300	monday	1.367	objective	3.700	mixing	3.100	wheels	3.133
father	2.333	date	1.200	cosmos	1.967	soft	1.500	cap	2.567

With the help of table 6, it is possible to see the differences within a category to have a closer look at the findings. With this view of the data, it is possible to find particular items that either do not properly fit the group and might find others that suit particularly well.

In the category *Social*, the item with the highest mean within the group is *"son"* with a score of *2.567*. On the opposite side, *"married"* has the lowest score within the category with a mean of *1.833*. Hence, participants semantically connected *married* more closely to the category *Social* than the word *son*.

Furthermore, the category *Time* includes the item "*date*", which has the lowest score of all items across categories, *1.2*. That means it has a high correlation to the category *Time*. As a comparison, the score that indicates perfect fit into the category would be 1.0, thus, the score of the item "*date*" comes very close to that. The highest score in the category *Time* is 2.3. Whereas, it is also to note that it has the highest score by far because the next lower score is 1.667.

The category *Outdoor* has a range from *1.833* to *3.8*. The items that represent this sore are *atmosphere* and *tedious* respectively.

The item "*soft*" has the score *1.5*, thus it has the lowest score in the category *Tactile*. The highest score within this category has "*mixing*" with a score of *3.1*, which seems to be relatively high in comparison to the other items since it has a *0.3* difference to the next item.

The last category is called *Visual* and includes the lowest score of 2.4, which belongs to the item "*cardboard*". The highest score of this category has the item "*gloves*" with *3.6* which could be regarded as an outlier since the next closest value has a difference of 0.467.

5. Discussion

4.1 Summary

In this research paper, we aimed at providing an additional perspective on the organisation of semantic knowledge in the human brain. This was done by comparing the mathematically established voxel-model with the mental model of participants. Additionally, we evaluated the labelling of the semantic groups in the voxel-model with means of a questionnaire. First, participants grouped the items into categories of their choice during the card sorting task. Second, Participants filled in a questionnaire in which they rated the relatedness of items to the categories, which were found by Huth et al. (2016). The groups resulting out of the card sort study were to a certain extent overlapping with the categories established in the study by Huth et al. Additionally, the questionnaire supported the category labelling with three out of five groups. Even though there are considerable differences, it can still be said that the overall commonalities indicate a common ground between the specific findings by Huth et al. to our study.

4.2 Interpretation

More in detail, these results show that there is a very close relatedness between some groups that were established by Huth and the three groups found in heatmap analysis as well as in the questionnaire. These groups are "*Social, Time and Tactile*". This shows that those groups are the most clearly defined categories for the individuals in the card sorting task and the questionnaire. However, the remaining categories were not clearly defined during the card sorting task, instead, they were grouped into small clusters with little indicators for an overarching cluster. Furthermore, the outcome of the questionnaire did not show a significant difference between the remaining categories and *Filler items*. This can have multiple causes. For example, the word "*plastic*" in the *Visual* category can have two distinct meanings when taken out of context. It can mean the visual aspect of something being plastic in terms of a 3D model or something is made out of plastic when the physical material is meant. Thus, this item could have resulted in misunderstandings that deviated from the findings of the study by Huth and his colleagues (2016) because they used narrative stories in which these words were embedded into context.

Next, the categories of *Visual and Outdoor* can be described as rather open and loosely defined categories. This can lead to overlap between categories. Here an example can be "*gloves*". A common reason to wear gloves is when it is cold outside and thus it could be put into the category *Outdoors* even though it belongs to *Visual* according to Huth et al. (2016). Furthermore, materials like "*cardboard, leather* and *plastic*" could be thought of in terms of their surface, thus the haptic and the feel of it might be important. This leads to the category of *Tactile* rather than *Visual* and shows that the interpretation without context and the loosely defined categories can have had an impact on the final results. This might also be a reason why *Time* and *Social* are well described. They seem to have lower semantic overlap with other categories from this research than *Visual* and *Outdoors* have.

However, since some of the words inside the categories have large differences between each other, this can also result in a seemingly greater Confidence Interval than it would have otherwise.

4.3 Integration

The results of the current research showed that there are multiple ways to integrate the findings of this paper into contemporary research. First, our results support the research by Huth and his colleagues to a certain extent. For two categories, the results showed clearly that they are in line with Huth's expectations. Additionally, a study done by Just et al. (2010) showed that there is a remarkable overlap between individuals' brain activity. They concluded a semantic representation across individual based on nouns. Contrarily, for the remaining groups, it was contradicting the results of the semantic representation of knowledge in the human brain. This means that almost half of the data could not provide evidence for the theory. Nevertheless, the two groups, *Time* and *Social*, were found to be very clearly related to the expected findings of Huth et al. (2016), thus it is suggested to research further and in greater detail to deny or support this connection.

Second, previously it has been found that concrete meanings are easier to categorise in the human brain. This is usually connected to modalities that process the direct information from the environment (Ghio, Vaghi, & Tettamanti, 2013). However, this is not in line with the findings of our study. Our results seem to support a modality independent categorisation most clearly because the categories connected to modalities such as *Visual* and *Tactile* were not conclusive. Additionally, in comparison to more abstract categories such as *Time* and *Social* the findings do not give new evidence towards the categorisation of knowledge according to modalities. Also, in 2003 researchers could not clearly identify the organisation of semantic knowledge between modality-specificity, attribute-specificity, and category-specificity (Thompson-Schill, 2003). Thus, this is not an unexpected finding in this field of research but it leaves room for a conclusive explanation in the future.

4.4 Limitations

A weak point in this research is the words used in the card sort and the questionnaire were words taken out of narrative speech in the study by Huth et al. (2016). However, in this paper, the context was not given to the participants and they had to make the categorisation and labelling without further information. This might have led to a different interpretation of ambiguous items when multiple meanings were possible. Yet, the research had thirty participants, of which the average result has been taken to have sufficient ground for the analysis.

Additionally, the participants of the study are coming from a non-English speaking country and usually do not have English as their mother tongue. Even though, the participants were either studying in English or were using it on a daily it might have deviated from a sample of native English speakers.

Another point that could have influenced the outcome of the research was the selection of the items used in the questionnaire and the in the card sorting task. The selection was done by the researchers which may be a point of potential biases. Although, this might be a source of potential limitation this was tried to counteract with the selection process and the reliability score. So that the researcher chose the items for each category out of the right and left hemisphere as well as using a minimum reliability score of “Not bad, pretty reliable”.

4.5 Conclusion

In the present article, we investigated the relationship between the mental model of participants and categories established in the voxel-model as well as the relation between the labelling of these categories and the people’s semantic relation. The findings of this study could support the creation and labelling of the categories of Huth et al. (2016) study only to a certain extent. This means, the findings were consistent for specific categories, *Time* and *Social* of the study by Huth et al. (2016), whereas other categories *Outdoor* and *Visual* showed a considerable difference between the expected model and the outcome of this study. Important aspects to consider in the future research may be the focus on native English-speaking participants and the involvement of the context of items these points can be of significance.

Since our findings are not fully in line with the study by Huth et al. (2016), neither with other contemporary research it is important to validate these findings again but also to open up the opportunity to draw the research in a new direction.

6. References

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: The evidence from category-specific semantic deficits. *Trends in Cognitive Sciences*, *7*(8), 354–361. https://doi.org/10.1016/S1364-6613(03)00159-1

Çelik, E., Dar, S. U. H., Yılmaz, Ö., Keleş, Ü., & Çukur, T. (2019). Spatially informed voxelwise modeling for naturalistic fMRI experiments. *NeuroImage*, *186*(October 2018), 741–757. https://doi.org/10.1016/j.neuroimage.2018.11.044

Chiou, R., Humphreys, G. F., Jung, J. Y., & Lambon Ralph, M. A. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex*, *103*, 100–116. https://doi.org/10.1016/j.cortex.2018.02.018

Croasmun, J. T., & Ostrom, L. (2011). Using Likert-Type Scales in the Social Sciences. *Environmental and Resource Economics*, *40*(1), 19–22.

Faiks, A., & Hyland, N. (2000). Gaining User Insight: A Case Study Illustrating the Card Sort Technique. *College & Research Libraries*, *61*(4), 349–357. https://doi.org/10.5860/crl.61.4.349

Ghio, M., Vaghi, M. M. S., & Tettamanti, M. (2013). Fine-Grained Semantic Categorization across the Abstract and Concrete Domains. *PLoS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0067090

Hart, J., & Gordon, B. (1992). Neural subsystems for object knowledge. *Nature*, *359*(6390), 60–64. https://doi.org/10.1038/359060a0

Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210–1224. https://doi.org/10.1016/j.neuron.2012.10.014

Just, M. A., Cherkassky, V. L., Aryal, S., & Mitchell, T. M. (2010). A neurosemantic theory of concrete noun representation based on the underlying brain codes. *PLoS ONE*, *5*(1).

https://doi.org/10.1371/journal.pone.0008622

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, *105*, 215–228. https://doi.org/10.1016/j.neuroimage.2014.10.018

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987. https://doi.org/10.1038/nrn2277

Quillian, M. R. (1988). Semantic Memory. *Readings in Cognitive Science*, (May 2014), 80–101. https://doi.org/10.1016/B978-1-4832-1446-7.50013-3

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42.

Real, R., & Vargas, J. M. (1996). The probabilistic basis of Jaccard's index of similarity. *Systematic Biology*, *45*(3), 380–385. https://doi.org/10.1093/sysbio/45.3.380

Schmettow, M., & Sommer, J. (2016). Linking card sorting to browsing performance – are congruent municipal websites more efficient to use? *Behaviour and Information Technology*, *35*(6), 452–470. https://doi.org/10.1080/0144929X.2016.1157207

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia*, *41*(3), 280–292. https://doi.org/10.1016/S0028-3932(02)00161-6

Tulving, E. (1972). Episodic and Semantic Memory. In *Organiz. Mem. Lon.* (Vol. 381, pp. 381–403).

Unsworth, S. J., Sears, C. R., & Pexman, P. M. (2005). Cultural influences on categorization processes. *Journal of Cross-Cultural Psychology*, *36*(6), 662–688. https://doi.org/10.1177/0022022105280509

Warrington, E. K., & McCarthy, R. (1983). CATEGORY SPECIFIC ACCESS DYSPHASIA. *Brain*, *106*(4), 859–878. https://doi.org/10.1093/brain/106.4.859

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*(3), 829–853. https://doi.org/10.1093/brain/107.3.829

7. Appendices

Appendix A

The survey used in the questionnaire study with fifty regular items and twenty filler items

	Word 1	Word 2	1 Highly related	2 Related	3 Neutral	4 Not related	5 Highly unrelated
1	son	social	◯	◯	◯	◯	◯
2	wednesday	time	◯	◯	◯	◯	◯
3	anticipation	outdoor	◯	◯	◯	◯	◯
4	touches	tactile	◯	◯	◯	◯	◯
5	cents	social	◯	◯	◯	◯	◯
6	relax	outdoor	◯	◯	◯	◯	◯
7	deadly	time	◯	◯	◯	◯	◯
8	crushed	tactile	◯	◯	◯	◯	◯
9	extreme	visual	◯	◯	◯	◯	◯
10	pieces	visual	◯	◯	◯	◯	◯
11	college	time	◯	◯	◯	◯	◯
12	children	social	◯	◯	◯	◯	◯
13	t-shirt	outdoor	◯	◯	◯	◯	◯
14	grinding	tactile	◯	◯	◯	◯	◯
15	gloves	visual	◯	◯	◯	◯	◯
16	month	time	◯	◯	◯	◯	◯
17	million	tactile	◯	◯	◯	◯	◯
18	atmosphere	outdoor	◯	◯	◯	◯	◯
19	mother	social	◯	◯	◯	◯	◯
20	score	tactile	◯	◯	◯	◯	◯
21	years	time	◯	◯	◯	◯	◯
22	wife	social	◯	◯	◯	◯	◯
23	since	time	◯	◯	◯	◯	◯
24	half	outdoor	◯	◯	◯	◯	◯
25	deeper	outdoor	◯	◯	◯	◯	◯
26	inches	visual	◯	◯	◯	◯	◯
27	smooth	tactile	◯	◯	◯	◯	◯
28	pairs	visual	◯	◯	◯	◯	◯
29	planets	outdoor	◯	◯	◯	◯	◯
30	parents	social	◯	◯	◯	◯	◯
31	tedious	outdoor	◯	◯	◯	◯	◯
32	thicker	tactile	◯	◯	◯	◯	◯
33	plastic	visual	◯	◯	◯	◯	◯
34	plus	visual	◯	◯	◯	◯	◯
35	week	time	◯	◯	◯	◯	◯
36	focus	outdoor	◯	◯	◯	◯	◯
37	sister	social	◯	◯	◯	◯	◯

38	saturday	time	○	○	○	○	○
39	brand	visual	○	○	○	○	○
40	fatal	time	○	○	○	○	○
41	meters	visual	○	○	○	○	○
42	married	social	○	○	○	○	○
43	emptiness	outdoor	○	○	○	○	○
44	regular	social	○	○	○	○	○
45	last	time	○	○	○	○	○
46	solid	tactile	○	○	○	○	○
47	quarter	tactile	○	○	○	○	○
48	leather	visual	○	○	○	○	○
49	relatives	social	○	○	○	○	○
50	grip	tactile	○	○	○	○	○
51	six	outdoor	○	○	○	○	○
52	extra	social	○	○	○	○	○
53	cardboard	visual	○	○	○	○	○
54	monday	time	○	○	○	○	○
55	severe	outdoor	○	○	○	○	○
56	wheels	visual	○	○	○	○	○
57	daughter	social	○	○	○	○	○
58	objective	outdoor	○	○	○	○	○
59	counting	social	○	○	○	○	○
60	flame	tactile	○	○	○	○	○
61	mixing	tactile	○	○	○	○	○
62	infection	time	○	○	○	○	○
63	cosmos	outdoor	○	○	○	○	○
64	tops	tactile	○	○	○	○	○
65	father	social	○	○	○	○	○
66	date	time	○	○	○	○	○
67	dress	visual	○	○	○	○	○
68	points	time	○	○	○	○	○
69	soft	tactile	○	○	○	○	○
70	cap	visual	○	○	○	○	○

Age:

Gender:

Highest Education:

Appendix B

Informed consent used for card sorting and questionnaire study

INFORMED CONSENT FORM

Project Title
A study of conceptual organisation

Purpose of the Study
This research is being conducted by Julian Merten, bachelor student at the University of Twente. I am inviting you to participate in this research project about Conceptual Learning.

Procedures
You will participate in a card sorting lasting approximately 30 minutes.

In this study you are requested to place the words (cards) together in groups, if you think they belong together in a group. You can make as many groups as you want.
After that you are requested to fill in a questionnaire (about 10 minutes).

You must be at least 16 years old.

Potential Risks and Discomforts
You do not have to answer any questions or do task you do not wish to answer or do. Your participation is voluntary and you are free to discontinue your participation at any time.

Confidentiality
Your privacy will be protected to the maximum extent allowable by law. No personally identifiable information will be reported in any research product. Moreover, only trained research staff will have access to your responses. Within these restrictions, results of this study will be made available to you upon request.
At the start of the research your name will be coded.
Only participating researchers have access to the provided data.

Right to Withdraw and Questions
Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify. The data you provided before you stopped participating however will be processed in this research; no new data will be collected or used.

If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact the primary investigator:
[Julian Merten, j.merten@student.utwente.nl]

Statement of Consent
Your signature indicates that you are at least 16 years of age; you have read this consent form or have had it read to you; your questions have been answered to your satisfaction and you voluntarily agree that you will participate in this research study. You will receive a copy of this signed consent form.

I agree to participate in a research project led by Julian Merten. The purpose of this document is to specify the terms of my participation in the project through being interviewed.

1. I have been given sufficient information about this research project. The purpose of my participation as an interviewee in this project has been explained to me and is clear.

2. My participation as an interviewee in this project is voluntary. There is no explicit or implicit coercion whatsoever to participate.

______________________ ______________________ ________

Name Participant Signature Date

______________________ ______________________ ________

Name Researcher Signature Date

Appendix C

Items and their category and location selected from the Huth et al. 2016 study.

Word	Category	Voxel	Location
Emptiness	Outdoor (blue)	9,90,64	LH occipital l.
Atmosphere	Outdoor (blue)	9,90,64	LH occipital l.
Cosmos	Outdoor (blue)	9,90,64	LH occipital l.
Planets	Outdoor (blue)	9,90,64	LH occipital l.
Objective	Outdoor (blue)	9,90,64	LH occipital l.
Touches	Tactile (blue-green)	9,79,68	LH occipital l.
Flame	Tactile (blue-green)	9,79,68	LH occipital l.
Crushed	Tactile (blue-green)	9,79,68	LH occipital l.
Thicker	Tactile (blue-green)	9,79,68	LH occipital l.
Grip	Tactile (blue-green)	9,79,68	LH occipital l.
Daughter	Social (red)	18,81,66	LH parietal l.
Children	Social (red)	18,81,66	LH parietal l.
Relatives	Social (red)	18,81,66	LH parietal l.
Married	Social (red)	18,81,66	LH parietal l.
Parents	Social (red)	18,81,66	LH parietal l.
Wednesday	Time (brown-purple)	14,87,65	LH parietal l.
Monday	Time (brown-purple)	14,87,65	LH parietal l.
College	Time (brown-purple)	14,87,65	LH parietal l.
Week	Time (brown-purple)	14,87,65	LH parietal l.
Saturday	Time (brown-purple)	14,87,65	LH parietal l.
Meters	Visual (green)	15,82,59	LH parietal l.
Wheels	Visual (green)	15,82,59	LH parietal l.
Inches	Visual (green)	15,82,59	LH parietal l.
Leather	Visual (green)	15,82,59	LH parietal l.
Cap	Visual (green)	15,82,59	LH parietal l.
Anticipation	Outdoor - Mental	13,18,36	RH frontal l.
Tedious	Outdoor - Mental	13,18,36	RH frontal l.
Focus	Outdoor (blue)	13,18,36	RH frontal l.
Relax	Outdoor (blue)	13,18,36	RH frontal l.

Deeper	Outdoor (blue)	13,18,36	RH frontal l.
Son	Social (red)	16,73,24	RH parietal l.
Wife	Social (red)	16,73,24	RH parietal l.
Father	Social (red)	16,73,24	RH parietal l.
Mother	Social (red)	16,73,24	RH parietal l.
Sister	Social (red)	16,73,24	RH parietal l.
Month	Time (brown-purple)	22,79,36	RH parietal l.
Date	Time (brown-purple)	22,79,36	RH parietal l.
Years	Time (brown-purple)	22,79,36	RH parietal l.
Since	Time (brown-purple)	22,79,36	RH parietal l.
Last	Time (brown-purple)	22,79,36	RH parietal l.
Grinding	Tactile (blue-green)	21,67,25	RH parietal l.
Mixing	Tactile (blue-green)	21,67,25	RH parietal l.
Smooth	Tactile (blue-green)	21,67,25	RH parietal l.
Soft	Tactile (blue-green)	21,67,25	RH parietal l.
Solid	Tactile (blue-green)	21,67,25	RH parietal l.
Pieces	Visual (green)	9,74,26	RH temporal l.
Plastic	Visual (green)	9,74,26	RH temporal l.
Gloves	Visual (green)	9,74,26	RH temporal l.
Cardboard	Visual (green)	9,74,26	RH temporal l.
Pairs	Visual (green)	9,74,26	RH temporal l.

Appendix D

Tailored R script for the card sort analysis

```
------------------------------------------------------------------------------------------------------------------------
# R script to generate a heatmap based on concepts from Huth et al. (2016)

# Call these libraries. They need to be installed as packages
library(gplots)
library(RColorBrewer)

data <- read.csv("Card sorting data.csv")

# Transform data in numerical format
mat_data <- data.matrix(data[,1:ncol(data)])

# Define colors of heatmap: red for high numbers
my_palette <- colorRampPalette(c("yellow","red"))(n = 299)

# Call heatmap function (from gplots), with these arguments
# See: https://www.rdocumentation.org/packages/gplots/versions/3.0.1/topics/heatmap.2
# Note: argument 'main=' gives name of plot

heatmap.2(mat_data, col = my_palette, density.info="none", trace="none",
        revC = TRUE, main="Name")
------------------------------------------------------------------------------------------------------------------------
```