‘Drifting Away’ or ‘Drifting a Car’:
The Categorization of Semantic Concepts Across Languages

Bachelor's thesis

Luisa C. Reiter

July 05, 2021

First Supervisor: Prof. Dr. F. van der Velde

Second Supervisor: Dr. S. Borsci

BMS Faculty
Department of Psychology
University of Twente

Abstract

Conceptual knowledge is an essential aspect of our communication, actions, and thinking. Despite the longstanding interest in the topic, it has not been settled how concepts are processed and how this knowledge is represented within the brain. A recent fMRI study by Huth et al. (2016) was able to shed more light on this representation and created a brain map that compromised eleven concept categories. However, it was questioned whether this representation also aligns with the active categorization of words. Additionally, since previous research has indicated that language might affect categorization, it was challenged whether the correspondence to the categories from the brain map is cross-lingual. Therefore, a card sorting study was conducted that made use of 50 words out of six categories from the brain map by Huth et al. (2016). Further, a questionnaire was administered to assess the perceived relation of the words to their category names. To assess whether the translation of the semantic concepts from English to German has an influence, a between-subjects design was applied. Twenty-two participants conducted the card sorting and questionnaire with the original English words, while 19 participants received translations in German. The findings revealed that the categories could only be partially recreated, and some categories were better represented than others. Therefore, the representation of concepts in the cortex could not fully account for the use of conceptual knowledge. The categorization differed to an extent between the languages. However, these differences did not lead to significant differences regarding the correspondence to the categories from the brain map. Hence, this representation might apply to different languages.

Keywords: categorization, language, card sorting, conceptual knowledge, semantic knowledge, hub-and-spoke theory, semantic control

Table of Content

1. Introduction

"What is the nature of the bee? and you answer that there are many kinds of bees, and I reply: But do bees differ as bees, because there are many and different kinds of them; or are they not rather to be distinguished by some other quality, as for example beauty, size, or shape? How would you answer me?" (Plato, ca. 380 B.C.E./1949).

The question of how we conceptualize the world around us has prevailed since the antique in Plato's Meno. Conceptual knowledge also called semantic knowledge, concerns the knowledge we have of a word's meaning, a person, or any object we encounter (Patterson et al., 2007). It is crucial for our understanding of language, as well as the recognition of objects around us, and the interaction with them (Lambon Ralph et al., 2017). When we encounter and learn about a concept, we also gain knowledge about the category it belongs to (Tyler & Moss, 2001). This enables us to recognize that an object we have never seen before belongs to a specific category, such as a chair, pullover, or bee. Based on this knowledge we can also infer information, which can guide our actions, for instance, sitting on the chair, putting on the pullover, or avoiding the bee (Lambon Ralph et al., 2017; Patterson et al., 2007). Consequently, concepts build a necessary basis for our communication, thinking, and actions.

There are varying theories that seek to explain and give insight into how concepts are processed and represented within the brain. One of them is the 'distributed-only view', which frames the semantic memory as a 'widely distributed neural network' (Patterson et al., 2007, p.976). It states semantic memory comes about by the activation of information within regions that are specific to one modality, such as colour or sound (Lambon Ralph & Patterson, 2008). Subsequently, categories, such as tools, are seen as not distinctly located within the brain but as emerging through the modality-specific regions (Patterson et al., 2007).

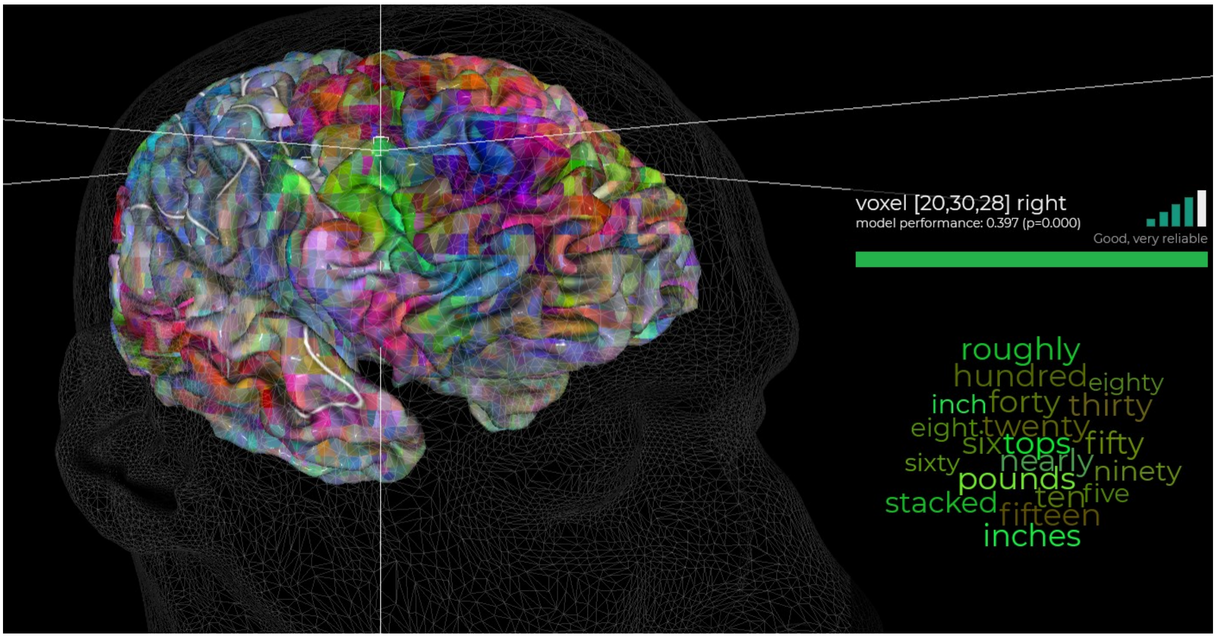
A variation of this view is the 'hub-and-spoke theory' or 'distributed plus-hub view'. It aims to explain how semantic memory can generate concepts that have similar semantic meaning, despite having different characteristics like colour or shape (Patterson et al., 2007). According to the theory, there are modality-specific areas in the cortical regions that deliver and retrieve information specific to one modality, like in the 'distributed-only view' (Patterson & Lambon Ralph, 2016). However, to generate higher-order categories we require more than these areas. Patterson et al. (2007) use the example of scallops and prawns. Even though they differ in shape, colour, and taste, they can be seen as belonging to the same category of seafood, as they are associated with similar scenarios. To explain this, the theory hypothesises that the modality-specific areas are like 'spokes' that are connected to a 'hub'. The different representations in the spokes are jointly connected or associated in the hub, and thus mediate the formation of a concept (Patterson et al, 2007). Borghesani et al. (2018) state, in line with the hub-and-spoke theory, that semantic representation is retrieved, at similar points of time, through a reciprocal relationship between the hub and spokes. When individuals are reading a word, its perceptual dimensions within the semantic system are retrieved in parallel to conceptual dimensions. For instance, when reading the word 'bee', it both comes to mind concurrently that bees have yellow and black stripes and are insects. The spokes are proposed to be located in modality-specific cortices across the cortex (Lambon Ralph et al., 2017; Martin, 2016). Findings, related to the illness semantic dementia (SD), led to the suggestion that the hub is located in the anterior temporal regions (ATR) (Lambon Ralph et al., 2017). Patients with SD were generally found to suffer from atrophy of the ATR (Silveri et al., 2018). This causes a decline in the individual's performance in semantic tasks, regardless of modality and concept (Lambon Ralph et al., 2017). The patient's errors are proposed to stem from the impairment of semantic generalizations and the ability to discriminate between concepts (Lambon Ralph & Patterson, 2008). For instance, a patient with SD might describe a picture of a zebra as a horse (Patterson et al., 2007). This supports the view that this region is responsible for the mediation of information from the cortices, or 'spokes', to generate concepts.

In addition to this, Lambon Ralph et al. (2017) propose that to explain the use of conceptual knowledge in varying contexts, the impact of cognitive control and semantic control processes needs to be considered. Thus, they proposed the 'controlled semantic cognition' (CSC) framework which describes the interaction between semantic representation and control. According to CSC, the 'hub' and 'spokes', or generally said the semantic representation processes, interact with areas that are important for semantic control (Lambon Ralph et al., 2017). These areas are hypothesised to be localized in the inferior parietal cortices, left inferior frontal gyrus, and posterior middle temporal gyrus, as supported by findings based on patients affected by semantic aphasia (SA) (Jefferies, 2013). In contrast to semantic dementia, the impaired performance of patients with SA depends on the degree of control it requires, such as inhibiting a response or discriminating between items (Rogers et al., 2015). Therefore, CSC proposes that conceptual knowledge is accessed and utilized through a system encompassing the control, as well as the representation of concepts.

To gain further insights into how the semantic system, or specifically the cerebral cortex represents words, Huth et al. (2016) conducted an fMRI study. They asked participants to listen to narrative stories while measuring their BOLD responses. Huth et al. (2016) were able to identify 965 common words which were used within the narrative stories and mapped them across the cerebral cortex. This led to the development of a 'semantic atlas' which represents voxel-wise information about specific semantic domains. This atlas was also made available online (Huth, n.d). Figure 1 presents a screenshot of the resulting semantic atlas on their website. The colours of each voxel in the atlas indicate the predicted activation in the cortex for each category. By clicking on a voxel, the words that are associated with it are displayed in the colour of their belonging category.

Figure 1

Screenshot of the Semantic Atlas as Presented in the Online Brain Viewer (Huth, n.d.)



Their findings indicate that the semantic system is to a large extent domain selective. Based on the brain activity observed, Huth et al. (2016) identified eleven different concept categories. Namely, they differentiated between the categories: 'visual', 'tactile', 'outdoor', 'number', 'bodypart', 'place', 'violence', 'person', 'mental', 'time', and 'social'.

It is unclear if a person would actively retrieve these categories and would categorize words like they seem to be represented within the brain. For instance, would a person actively categorize 'melting' and 'flame' into one category, as suggested by Huth et al. (2016)? The present thesis will be able to seek answers to these questions by studying whether participants categorize concepts as suggested by Huth et al. (2016).

2. Current Research

The current research will study how people active categorize concepts using a card sorting task and thereby, investigate whether participants retrieve and reproduce the categories that Huth et al. (2016) identified in their semantic atlas. To do so, it will focus on six of the eleven categories outlined by them. Specifically, the study will address the categories: ‘mental’, ‘number’, ‘social’, ‘tactile’, ‘outdoor’, and ‘place’. Thereby, the first research question is posed: *‘Does the categorization of words by participants correspond to the representation of semantic concepts in the cerebral cortex?’.*

To make general claims about the categorization of words and theories that seek to apply to different cultures and languages, it will also be investigated whether language has an influence. While Dehghani et al. (2017) show in an fMRI study that the representation of the meaning of stories seems to be remarkably cross-lingual, this might prove to be different for concepts expressed through an isolated word. For instance, Kim et al. (2021) illustrate that the categorization of verbs across Korean and German was only partially cross-lingual. Specifically, gustatory verbs (e.g., ‘eat’ or ‘chew’) were grouped similarly in both languages, but other verbs, such as ‘mean’, ‘practice’, and ‘examine’, were categorized differently across the languages. They hypothesise that these differences stem from the culturally different experiences of respondents, such as the education system. Therefore, it might be beneficial to address how language differences could affect word categorization and ultimately, conceptual knowledge.

Following this, the study will use the original English words, as well as German translations of them to test whether this might affect categorization and thereby, the compatibility with the semantic atlas by Huth et al. (2016). This will address the second research question: *‘Does language affect the participants’ categorization and consequently, the correspondence with concepts in the cerebral cortex?’.*

To find answers to the research questions, a one-level open card sorting task will be conducted. Card sorting is used to study how participants categorize their knowledge (Wood, & Wood, 2008). It enables the research of the organizations of conceptual knowledge and mental models in different domains (e.g. Doran et al., 2018; Smith et al., 2013). When conducting a card sorting task, participants are provided with a set of words or items on cards and asked to group them into categories. In single-level card sorting, participants are asked to group the cards one time, whereas, in hierarchical card sorting, they are asked to separate their initial categories into smaller ones within several rounds (Schmettow & Sommer, 2016). Due to the Covid-19 restrictions at the time of the data collection the card sorting is conducted online. Due to the limited options for online card sorting a one level card sorting is used. Further, within open card sorting, the names and number of the categories are not given beforehand by the researcher but freely selected by the participant (Hudson, 2005). Using the brain viewer by Huth (n.d.) fifty words for the card sorting were selected and translated into German. They are presented in Table 1. For a more thorough list of all items which indicates the originating voxel within the map and brain area, as well as its reliability, see Appendix A.

Table 1

Words Used in the Card Sorting Task from the Selected Six Categories by Huth et al. (2016)

mental		**number**		**social**	
English	German	English	German	English	German
thoughts	Gedanken	each	jedes	daughter	Tochter
overwhelmed	überwältigt	ten	zehn	family	Familie
consciousness	Bewusstsein	six	sechs	relatives	Verwandte
anxiety	Angst	quarter	Viertel	police	Polizei
loneliness	Einsamkeit	per	je	wife	Ehefrau
laughing	lachen	set	Set	husband	Ehemann
calmly	ruhig	plus	plus	refused	verweigert
loudly	laut	collect	sammeln	died	gestorben
startled	aufgeschreckt			parent	Elternteil

tactile		**outdoor**		**place**	
English	German	English	German	English	German
brittle	spröde	clouds	Wolken	driveway	Einfahrt
melting	schmelzen	waves	Wellen	unlocked	unverschlossen
fluid	flüssig	drifting	treiben	bus	Bus
flame	Flamme	flooded	überflutet	parking	parken
swallow	schlucken	breathe	atmen	neighborhood	Wohngegend
grip	Griff	atmosphere	Atmosphäre	attendance	Anwesenheit
touches	berühren	heavens	Himmel	university	Universität
squeeze	quetschen	explosions	Explosionen	students	Schüler

Further, to test whether the words are associated with their originating category, a questionnaire with a 5-point-Likert scale will be used. It will assess to which degree the participants see the items as relating to the names of the category. Likert scales were introduced in 1932 by Rensis Likert and were originally used to quantitatively assess the level of agreement or approval of a statement on a metric scale (Joshi et al., 2015; Likert, 1932). A 5-point Likert scale presents a symmetric scale, as the neutral point of assessment lies in between the two extremes and participants can make a balanced decision between them (Joshi et al., 2015). In this study, the participants will be asked to which extent they perceive the items to be related to their category name, ranging from 'highly related' to 'highly not related'. In addition to the experimental items that are displayed above, control/filler items will be used. These control items stem from the remaining categories identified by Huth et al. (2016) that are not being investigated (Table 2). In the questionnaire, each filler item will be presented with one of the six categories under inspection. The category names in the German condition were translated, as well: 'mental' to 'mental', 'number' to 'Nummer', 'social' to 'sozial', 'tactile' to 'haptisch', 'outdoor' to 'draußen', and 'place' to 'Ort'.

Table 2

Filler Items Used in the Questionnaire from the Remaining Categories by Huth et al. (2016)

visual		violence		person		time	
English	German	English	German	English	German	English	German
coloured[4]	farbig	evidence[1]	Beweis	cop[1]	Polizist	monday[3]	Montag
sleeve[5]	Ärmel	harm[5]	Schaden	sheriff[2]	Sheriff	week[6]	Woche
brown[6]	braun	punished[6]	bestraft	robbery[4]	Raub	college[5]	Hochschule
boots[2]	Stiefel	victim[2]	Opfer	sons[4]	Söhne	school[4]	Schule
coat[3]	Mantel	innocent[1]	unschuldig			spent[3]	verbracht
olive[2]	oliv						

Note. The superscripts indicate the category names the filler item was presented with. 1 = mental, 2 = number, 3 = social, 4 = tactile, 5 = outdoor, 6 = place.

3. Method

3.1 Design

The study employed a one-level open card sorting task[1] and a questionnaire. A between-subjects design was used with language (English vs. German) as the independent variable.

3.2 Participants

The study included 42 participants across the two conditions. They were either directly contacted by the researcher or were students that participated via the SONA system of the University of Twente. Thus, a convenience sample was drawn. Thirty-one of the participants were female (73.81%) and 11 were male (26.19%). Their age ranged from 18 to 52 (*M* = 22.78, *SD* = 7.28). Further, six participants were Dutch (14.29%), 33 German (78.57%), and three from other countries (7.14%). To be able to participate in the study it was required to have sufficient language skills in either English or German. Depending on the participants' language skills, they were assigned to one of the conditions. If participants indicated a level above B1 in both languages, they were randomly assigned to one of the conditions. Twenty-three of them took part in the English condition and 19 in the German condition. The participants' demographics differed slightly between the two conditions (see Appendix B). The participants were provided with a consent form that informed them about the aim of the study, the procedure, and their rights (see Appendix C). Additionally, if participants signed up via the SONA system, they received 0.25 credits for taking part in the study after completion. The research was reviewed and approved on the 12th of March 2021 by the BMS Ethics Committee of the University of Twente.

3.3 Material

3.3.1 Card Sorting

Within the open one-level card sorting task participants were provided with the 50 words taken from the semantic atlas of Huth et al. (2016) (see Table 1). For each of the investigated categories (mental, number, social, tactile, outdoor, place), at least two voxels were selected, out of which eight to nine words were then included. The used voxels had to have at least a reliability criterion that was labelled 'Not bad, pretty reliable' within the semantic atlas. If respondents conducted the English condition, they were presented with the original words, while in the German condition they were presented with a translation. The

[1] Since the study was conducted remotely a one-level card sorting had to be employed, as the used survey platform 'Qualtrics' did not enable hierarchical card sorting.

words were translated by meaning by a native German and fluid English speaker. These translations were checked by a second native German and fluid English speaker. Additionally, they were compared and possibly adapted to the most frequent translation indicated by the platform Google Translate. However, if the meaning of the word could be described better by a less frequent word this was chosen. The selected 50 words were shown to the participants in a randomly ordered list on the left side of the computer screen within the 'Qualtrics' survey. On the right, 16 boxes were displayed in which they could drag the words to group them. They were informed that they would not have to make use of all boxes.

3.3.2 Questionnaire

The participants were asked to rate the strength of the relation between the items and category names in a questionnaire using a 5-point Likert scale. The scale ranged from '1=highly related' to '5=highly not related'. The first word always was an experimental or filler item and the second word the name of one of the six categories. The filler items were taken from the remaining categories described by Huth et al. (2016) (see Table 2). They were also presented with one of the investigated category names so that they could be used as control items. In total, the participants had to judge the relation of 70 word pairs; 50 of them included the experimental items and 20 the filler items. The questionnaire is presented in English in Appendix E and German in Appendix F. All items were presented to the participants in the same order in both conditions.

3.4 Procedure

The study was conducted using the online survey platform 'Qualtrics' aimed at gathering responses from participants remotely. The participants were provided with a link to the survey via mail or via the SONA system. They were asked beforehand to conduct it on their personal computer, instead of their mobile phone, to ensure that it was displayed correctly. The time needed to conduct the survey was approximately 15 to 30 minutes. At the beginning of the survey, the participants were presented with a consent form (Appendix B). It informed them about the aim of the study, the procedure, and their rights. They were informed that they would have the right to withdraw from the study at any moment without having to give a reason. Additionally, it stated that their data would be collected and stored anonymously. If they agreed, they proceeded to questions concerning their demographic information (Appendix D). These included their age, gender, nationality, level of education, and level of English and German skills. Depending on their language skills, they were either continuing with the one-level card sorting task in German or English. They were asked to sort the 50 words into groups. The maximum number of groups was 16, however, they were instructed that they would not have to make use of all boxes. After this, they were asked to fill out the questionnaire with the 70 questions regarding the relationship between the item words and category names. Upon completion, they were thanked for their participation and provided with an email address, if they had any questions or would like to receive the report.

3.5 Data Analysis

The data was prepared and screened in Excel. Specifically, variables such as the date or progress of participation were deleted. Additionally, the time people used to complete the survey was looked at to identify if people potentially did not participate seriously. Moreover, the descriptive statistics of the demographic data were calculated generally and for each condition separately. This encompassed the mean, standard deviation, and range, for the participants' age, and frequencies for gender, education, nationality, and language skills.

The data from the card sorting was transformed into a 'proximity matrix' using Python. Specifically, within Python, the Jaccard scores for each word combination were calculated. Since it was a one-level card sorting, this calculation was done by counting how often a specific word combination occurred across all participants. The Jaccard score was then written into the corresponding field within the matrix. The matrix was exported into a csv file that could be processed by the statistics software 'R studio'. For the complete Python script, see Appendix G. Manually the scores for the 'combination' of each word with itself were set to the participant number in the condition and the first column (including the item names) was deleted. Then, the data file was processed in 'R studio' and the heatmaps were created by using the 'heatmap.2' function (see Appendix H). To be able to compare the six categories by Huth et al. (2016) to the heatmaps, six clusters were selected based on the dendrograms.

Further, to assess whether the card sorting differed between the conditions, the Jaccard scores that corresponded to the categories in the brain map of Huth et al. (2016) were identified. The scores were noted down in a list for each condition (English or German). The Jaccard scores were divided by the number of participants in each condition to arrive at a score between 0 and 1, to make them comparable. For each category and the combined categories, descriptive statistics, including the mean and standard deviation, were calculated. Following, in 'R studio', q-q plots and Shapiro-Wilk-tests were utilized to see whether the Jaccard scores were normally distributed. After this, a Mann Whitney U test was performed to see whether the Jaccard scores differed significantly between the conditions.

Following, the questionnaire assessing the perceived relationship between the item words and category words was analysed. For this, the mean and standard deviation for each item was calculated. The possible scores ranged from 1 to 5, with 1 indicating that the words were ‘highly related’ and 5 ‘highly not related’. Afterwards, new variables were introduced that entailed the average for each category and one variable that described the average across all categories. The variables were then screened for normality using the Shapiro-Wilk-test and q-q plot. Afterwards, an assessment was made to see whether the experimental items differed significantly from the control items in both conditions, using a paired t-test. Lastly, it was assessed whether the category rating in the questionnaire differed significantly between the conditions with an independent samples t-test.

4. Results

The data included 41 participants across the English (*n*=22) and German conditions (*n*=19). In the English condition, one respondent was excluded because it was questionable whether he/she responded genuinely.[2] First, the heatmaps of the conditions (English or German) are described separately. Following, the results of the analysis of the Jaccard scores are presented. Lastly, the results of the analysis of the questionnaire are reported.

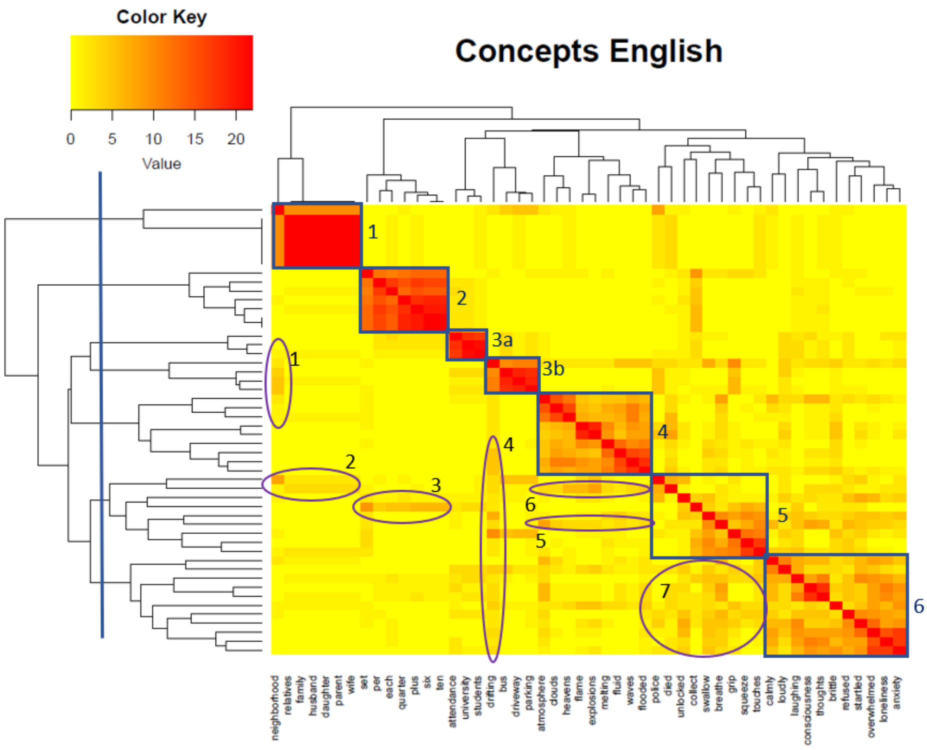
4.1 Card Sorting

4.1.1 Heatmaps

The data acquired in the English one-level open card sorting task was used to produce a heatmap with 6 clusters (Figure 2). The clusters were chosen based on the dendrogram (blue line). With the dendrogram, either 5 or 7 clusters could have been selected. Due to the similarity concerning their originating category, the words in the third cluster will be approached together. In the following, each cluster was analysed and compared to the clusters of Huth et al. (2016).

Figure 2

Heatmap based on the Card Sorting Task in the English Condition (n=22)



Note. The blue rectangles indicate the selected clusters, and the purple ovals highlight ‘warm spots’ (signifying overlap of words with words outside their cluster).

[2] The participant was excluded since they only used 4.57 min to finish the survey and only formed two groups, while on average participants needed 19.40 min and when pilot testing it was estimated to take at least 10 min.

Table 3 provides an overview of the words in each cluster (in order of the heatmap) and their original categories in Huth et al. (2016).

The first cluster resembles the category of *social* to a large extent. Nearly all words in the formed cluster stem from the *social* category, such as ‘husband’ or ‘parent’. Notably, they are all nouns and are used to describe family members. An exception to this is ‘neighborhood’ which is the only word that does not stem from the *social* domain but *place*. It is visible in the heatmap that there is some ambiguity regarding this word (oval 1). Thereby, ‘neighborhood’ shares some overlap with cluster 3b, which mainly entails words from the category *place*. Furthermore, three of the social words, ‘police’, ‘died’, and ‘refused’, were grouped with other clusters. ‘Police’ and ‘died’ both are found in cluster 5 but show some ambiguity regarding the first cluster, as well (oval 2). Consequently, some participants associated these words with the other words in the *social* category. Lastly, the word ‘refused’ fell into cluster 6 and does not show any overlap with cluster 1.

The second cluster aligns with the category *number*. Seven of the 8 words in the category are also presented in the heatmap together. The words in the cluster are very tightly associated. Merely the word ‘collect’ that stems from this category was found in cluster 5 instead. However, it also presents some overlap with the second cluster, as visible in the third oval. This shows that it was frequently grouped with the other *number* items, particularly ‘set’.

The third cluster covers most of the words from the category *place*. Despite having some overlap, clusters 3a and 3b seem to have been frequently grouped apart. Notably, the words in cluster 3a, ‘attendance’, ‘university’, and ‘students’, all originate from the same voxel in the parietal cortex (Appendix A). In cluster 3b, the word ‘drifting’ originally stemmed from the category *outdoor*. It also shows some overlap with the other clusters (oval 4). Moreover, there are several words from the domain *place* that are not present in cluster 3. Specifically, ‘neighborhood’ as discussed above, and ‘unlocked’. The word ‘unlocked’ does not show a particular overlap with the other words from the third cluster.

The items in the fourth cluster mainly resemble the category *outdoor* and to some extent *tactile*. Particularly, the words ‘atmosphere’, ‘clouds’, and ‘heavens’ were often categorized together. Similarly, ‘flame’ and ‘explosions’, as well as, ‘fluid’, ‘waves’, and ‘flooded’ were frequently grouped. Six of the 8 words of the category outdoor are present in the cluster. Merely the words ‘drifting’ and ‘breathe’ do not fall into the cluster. The word ‘drifting’ in cluster 3 was to an extent associated with the words: ‘flooded’, ‘waves’, and ‘fluid’ (oval 4). The word ‘breathe’ was grouped into cluster 5. However, it also shows some ambiguity towards the third cluster. The word ‘died’ from the social category also shows overlap with cluster 4 (oval 6).

The fifth and sixth clusters show a lot of overlap and ambiguity with each other (oval 7) and are less dense in terms of their colour values. The fifth cluster does not resemble any of the categories of Huth et al. (2016) fully. Two items from the category *social*, 1 from *place*, 1 from *number*, 1 from *outdoor*, and comparatively the largest amount, 4 from *tactile* fall into cluster 5. The four tactile words, 'grip', 'squeeze', 'touches', 'swallow', and the word 'breathe' from the category *outdoor*, are most clearly associated within this cluster in the heatmap.

Lastly, the words in the sixth cluster closely align with the category *mental* in the study by Huth et al. (2016). Despite the numerous associations of its words with others in cluster 5 (oval 7), all words of the category *mental* fell into the sixth cluster. Within it, the strongest associations are formed by the words 'overwhelmed', 'loneliness', and 'anxiety'. Furthermore, the words 'consciousness' and 'thoughts' were frequently combined. The words 'brittle' from the *tactile* category and 'refused' from the *social* category were also grouped into the sixth cluster. As stated above, 'refused' did not show any overlap with the first cluster with the other social words.

Table 3

Clusters in the English Heatmap

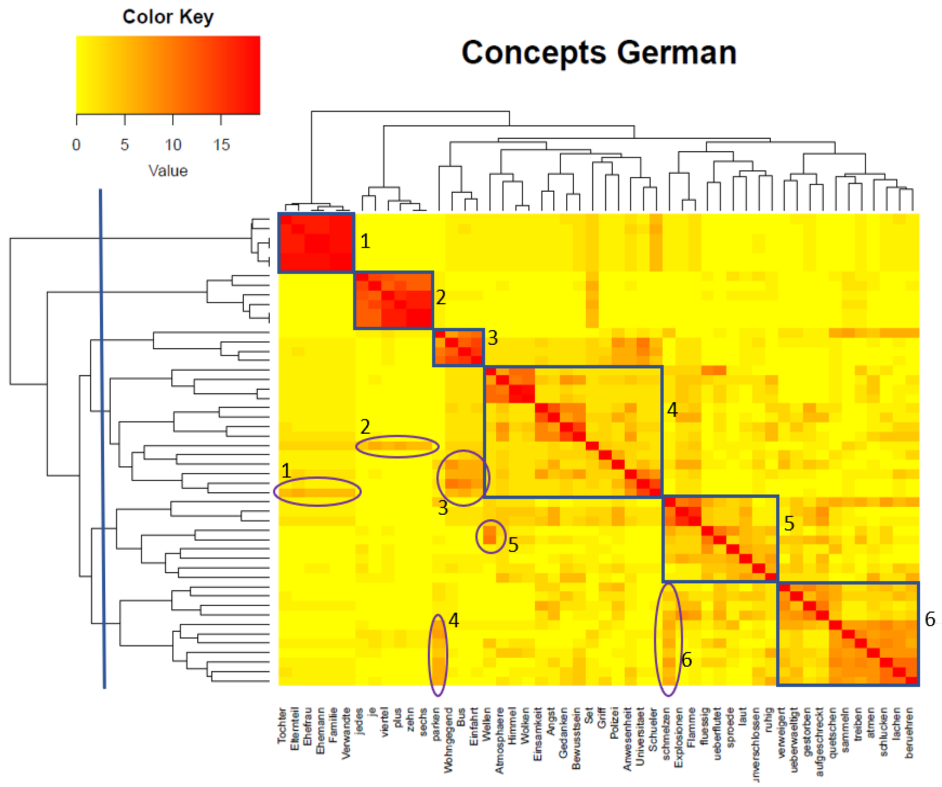
Cluster 1	Cluster 2	Cluster 3 (a&b)	Cluster 4	Cluster 5	Cluster 6
neighborhood (p)	set (n)	attendance (p)	atmosphere (o)	police (s)	calmly (m)
relatives (s)	per (n)	university (p)	clouds (o)	died (s)	loudly (m)
family (s)	each (n)	students (p)	heavens (o)	unlocked (p)	laughing (m)
husband (s)	quarter (n)	drifting (o)	flame (t)	collect (n)	consciousness (m)
daughter (s)	plus (n)	bus (p)	explosions (o)	swallow (t)	thoughts (m)
parent (s)	six (n)	driveway (p)	melting (t)	breathe (o)	brittle (t)
wife (s)	ten (n)	parking (p)	fluid (t)	grip (t)	refused (s)
			waves (o)	squeeze (t)	startled (m)
			flooded (o)	touches (t)	overwhelmed (m)
					loneliness (m)
					anxiety (m)

Note. The letters in the parenthesis indicate the originating categories in the brain map (see Huth, n.d.). p = place; s = social; n = number; o = outdoor; t = tactile; m = mental.

The heatmap for the German card sorting task was also separated into six clusters by utilizing its dendrogram (Figure 3). Table 4 presents the corresponding clusters in the German heatmap and the English translations of the words in the presented order of the map.

Figure 3

Heatmap in the German Condition Based on the Card Sorting Task (n=19)



Note. The blue rectangles indicate the selected clusters, and the purple ovals highlight ‘warm spots’ (signifying overlap of words with words outside their cluster).

The first cluster closely resembles the *social* category in Huth et al. (2016) and the first cluster in the English heatmap. All the words from the *social* category that were present in the English heatmap are presented in the first cluster in the German heatmap, as well. One difference is that the word ‘neighborhood’ from the category *place* is not grouped into the cluster and does not share any overlap with the words in the first cluster, as presented in the heatmap. Moreover, there is some ambiguity regarding the word ‘students’ (‘Schüler’) from the category *place*, which is present in cluster 4 (oval 1).

The second cluster aligns with the domain *number* and the second cluster in the English condition. All words that are grouped in cluster 2 stem from the category *number* and were highly associated with each other. However, the word ‘set’ (‘Set’) which was a part of the second cluster in the English heatmap is grouped in the fourth cluster within the German heatmap, despite showing some association with the other words in the second cluster (oval 2). Similar to the English heatmap, the word ‘collect’ (‘sammeln’) is not presented in the second cluster but in the sixth with words from several other categories.

The third cluster solely constitutes words from the category *place* and cluster 3b in the English heatmap. The word ‘neighborhood’ is also present in this cluster. The cluster shows some overlap with three of the other words from its original category in cluster 4 (oval 3). Namely, ‘attendance’ (‘Anwesenheit’), ‘university’ (‘Universität’), and ‘students’ (‘Schüler’), which are all from the same voxel and make up cluster 3a in the English heatmap. Furthermore, the word ‘parking’ (‘parken’) in cluster 3 shows some warm spots regarding the sixth cluster (oval 4).

The fourth cluster is considerably less conclusive than the previous clusters and is made up of 14 words. Four of the words are from the category *outdoor*, 4 from *mental*, 3 from *place*, and 1 each from *number*, *tactile*, and *social*. All words in the cluster are nouns. The 4 words from the category *outdoor* overlap the most with each other and only to an extent with the other words in the cluster. The word ‘waves’ (‘Wellen’) was also associated with the words ‘fluid’ (‘flüssig’) and ‘flooded’ (‘überflutet’) in the fifth cluster (oval 5). This association also occurred in the fourth cluster in the English condition. Apart from this, the words from the category *mental* are distinctly visible within the cluster. Unlike in the English condition, the nouns do not seem to have been associated with the other words from their originating category. Moreover, the three words from the category *place* are most clearly associated with each other and show some association with the words in cluster 3 (oval 3).

The fifth cluster does not represent any domain but shares some similarities with the fourth cluster in the English condition. For instance, ‘explosions’ (‘Explosionen’) from the category *outdoor* and ‘flame’ from *tactile* are most highly related in the cluster. This combination also occurred within the English heatmap. The word ‘melting’ (‘schmelzen’) was associated with them as well, but also shows some ambiguity towards the sixth cluster (oval 6). Further, other words from the *tactile* category, like ‘brittle’ (‘spröde’) and ‘fluid’ (‘flüssig’), as well as words from the category *mental*, *place*, and *outdoor*, were grouped into the fifth cluster.

Finally, the sixth cluster entailed words from several categories such as *outdoor*, *social*, *mental*, *number*, and *tactile*, and does not overlap with any of the 6 English clusters. All the words in the sixth cluster present verbs or adjectives, while in the fourth cluster only nouns are present, which is a distinction that did not occur in the English condition.

Table 4

Clusters in the German Heatmap

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
daughter (s)	each (n)	parking (p)	waves (o)	melting (t)	refused (s)
parent (s)	per (n)	neighborhood (p)	atmosphere (o)	explosions (o)	overwhelmed (m)
wife (s)	quarter (n)	bus (p)	heavens (o)	flame (t)	died (s)
husband (s)	plus (n)	driveway (p)	clouds (o)	fluid (t)	startled (m)
family (s)	ten (n)		loneliness (m)	flooded (o)	squeeze (t)
relatives (s)	six (n)		fear (m)	brittle (t)	collect (n)
			thoughts (m)	loudly (m)	drifting (o)
			consciousness (m)	unlocked (p)	breathe (o)
			set (n)	calmly (m)	swallow (t)
			grip (t)		laughing (m)
			police (s)		touches (t)
			attendance (p)		
			university (p)		
			students (p)		

Note. The letters in the parenthesis indicate the originating categories in the brain map (see Huth, n.d.). p = place; s = social; n = number; o = outdoor; t = tactile; m = mental.

4.1.2 Jaccard scores

The Jaccard scores ranged between 0 and 1, whereby 1 indicated that all participants sorted the words into one group and 0 that no participant combined them. On average, the word combinations from the category *number* had the highest Jaccard score (M = .60, SD = .23) in the English condition. This indicates that the words from the category number were sorted more frequently with each other than the words from the other categories. In contrast, the category *tactile* had on average the lowest Jaccard score (M = .22, SD = .27). Thus, especially the words from the *tactile* category were less frequently combined. Similarly, in the German condition, the highest average Jaccard score was found for the category *number* (M = .49, SD = .31) and the lowest score for the category *tactile* (M = .20, SD = .17). The mean and standard deviation of each category's Jaccard score across the English and German condition are displayed in Table 5.

Table 5

Average Jaccard Scores per Category in the English (n=22) and German (n=19) Condition

	English		German	
	Mean	SD	Mean	SD
mental	.37	.18	.21	.18
number	.60	.23	.49	.31
social	.47	.37	.43	.43
tactile	.22	.27	.20	.17
outdoor	.30	.20	.26	.22
place	.25	.32	.32	.24
combined	.37	.32	.32	.30

The q-q plot and Shapiro-Wilcoxon test were used to assess the normality of the combined Jaccard score in both conditions and indicated that the Jaccard scores were not normally distributed ($p < .05$). Therefore, the Mann-Whitney U test was used and indicated that the Jaccard scores did not differ significantly between the English ($Md = .27$) and German ($Md = .21$) condition ($U = 15141$, $z = -1.75$, $p = .08$).

4.2 Questionnaire

The questionnaire scores could vary between 1 and 5, whereby 1 indicated that the item and category were rated as 'highly related' and 5 that they were rated as 'highly not related'. For the mean scores and standard deviation of each item, see Appendix I. The lowest mean, indicating the closest relationship between the items and their category name, was found for the category *mental* ($M = 1.73$, $SD = 0.37$). The category *tactile* received the highest mean ($M = 2.38$, $SD = 0.48$) (see Table 6). The scores of the experimental items were compared to the control items in the English condition. The q-q plot and Shapiro-Wilcoxon test were performed and showed that the experimental item score, as well as control item score, were normally distributed ($p > .05$). The paired t-test showed that the experimental items ($M = 1.96$, $SD = 0.31$) and control items ($M = 3.65$, $SD = 0.65$) differed significantly ($t(21) = -13.51$, $p < .01$). Hence, the experimental words received significantly lower ratings, indicating that the words were more related to the category names than the control words.

Table 6

Average Questionnaire Scores per Category in the English (n=22) and German (n=19) Condition

	English		**German**	
	Mean	SD	Mean	SD
mental	1.73	0.37	1.80	0.41
number	1.74	0.52	2.03	0.55
social	1.82	0.43	2.02	0.49
tactile	2.38	0.48	2.51	0.90
outdoor	1.97	0.42	2.03	0.53
place	2.10	0.59	2.23	0.54
combined	1.96	0.31	2.10	0.32

Furthermore, the categories of the two conditions were compared. All means of the categories were slightly higher in the German condition than in the English condition (see Table 4). However, like in the English condition, the items in the category *mental* were assessed as most related to the category name (M = 1.80, SD = 0.41). The highest mean was found for the category *tactile* (M = 2.51, SD = 0.90). The collapsed item score in the German condition was normally distributed ($p > .05$). The score of the filler items was normally distributed, as well ($p > .05$). The paired t-test revealed that the experimental items (M = 2.10, SD = 0.32) were rated as significantly more related to the category names than to the filler items (M = 3.94, SD = 0.48; t(39) = -13.97, $p < .01$). Lastly, an independent t-test was used to test whether the assessment differed significantly between the conditions (English versus German). The test indicated that the items were not rated significantly different in both conditions (t(39) = 1.47, p = .15).

5. Discussion

The purpose of the study was to assess whether the categorization of words aligns with the representation of concepts in the cerebral cortex, as reported by Huth (2016). Thus, the following research question was posed: '*Does the categorization of words by participants correspond to the representation of semantic concepts in the cerebral cortex?'*. Additionally, it was questioned whether the concepts are transferable to other languages or if they would be affected by translating the English words to another language, in this case, German. Following, the second research question was formulated: *'Does language affect the participants' categorization and consequently, the correspondence with concepts in the cerebral cortex?'*.

5.1 Research Question 1

Concerning the first research question, it can be said that the categorization resulting from the card sorting partially aligned with the domains that Huth et al. (2016) identified. Especially the categories *number*, *social*, and to an extent *place* and *mental* were well represented. The categories *outdoor*, and particularly *tactile*, were less conclusive. In addition, the words from the categories were generally rated as significantly more related to the category name than control items in the questionnaire. However, similarly to the result of the card sorting, the words of the categories *mental*, *social*, and particularly *number* were viewed as more related to the categories' name. Especially the category *tactile* was rated as less related. Therefore, the words of the categories *outdoor*, *tactile* and *place* do not seem to form as definite categories as expected.

Within the card sorting, also in the more conclusive categories, several words did not fall into their original category. For instance, the word 'neighborhood' was grouped with the *social* words in the first cluster while it stemmed from *place*. Possibly this occurred because it encapsulates the word 'neighbour', which labels the relationship towards a person, like the other words in cluster 1. Interestingly, this did not occur in the German condition and the word also did not show an overlap with the words from the *social* category. Presumably, since the word 'neighborhood' ('Wohngegend') means literally translated as 'living area', it might be less associated with the *social* words. This would indicate that not only the complete meaning of the word but also the meaning of its components plays a role in categorization. Another example is the word 'drifting'. It was grouped into cluster 3b with words from the category *place,* while it stemmed from the category *outdoor*. This might have occurred because 'drifting' is a homonym, as it can have multiple different meanings. For instance, when interpreting the word as 'drifting with a car' it seems more readily associated with words such as 'bus' or 'parking' in cluster 3b, while ‘drifting away’ or ‘drifting on water’ might allude to words in its original category, such as ‘waves’ or ‘flooded’. Thereby, the interpretation of the word is largely dependent on its context (Rice et al., 2018). Notably, this context is not given in the card sorting task, while it was given in the study by Huth et al. (2016), as participants listened to narrative stories. Another example of this is the association between ‘died’ and ‘police’, which alludes to the more situational context the words are used in. Moreover, one of the clusters in the heatmap mixed the words from the categories *tactile* and *outdoor*. The words ‘flame’ and ‘explosions’, as well as, ‘fluid’, ‘waves’, and ‘flooded’ were frequently combined. This might point towards categorization of the words, not due to their relation to *tactile* or *outdoor*, but due to their association with the elements, such as air and water. Especially since the words in the *tactile* category were rated as the least related to their category name, this might show that it does not form a definite category.

These results suggest that the semantic representation in the cerebral cortex, as reported by Huth et al. (2016), is not fully conclusive to indicate the active categorization of concepts. Only a few of the researched categories, such as *number*, *social*, and *mental* were well represented. This could indicate that the domains that Huth et al. (2016) outline should be questioned. However, their findings could be largely replicated even when participants read the narrative stories instead of listening to them (Deniz et al., 2019). Therefore, another possibility is that the semantic representation is not sufficient to explain the categorization of concepts within card sorting. This supports the view that in addition to the hub and spokes, the impact of semantic control is crucial to account for the use of conceptual knowledge (Lambon Ralph et al., 2017). This might be especially important in this study, as individual words, as seen by the word ‘drifting’, can be interpreted ambiguously. Hence, participants had to selectively decide between associations of the different meanings of the word. This requires considerable demands of semantic control (Hoffmann, 2018).

5.2 Research Question 2

Concerning the second research question, the results indicate that there was not a significant difference between the categorization in English and German concerning the resemblance to the categories of Huth et al. (2016). Similarly, the relation of the words with their category names given by Huth et al. (2016) did not differ significantly between the English and German ratings. Nevertheless, within the card sorting, several differences between the formed clusters were observed. For instance, there was a distinction between word types in the clusters in the German condition. While in the fourth cluster only nouns were present, in the fifth and sixth only verbs and adjectives were found, which was not the case in the English condition. Perhaps, an influence on this could have been that nouns are visually more distinct from other word types in German than in English because the first letter of a noun is capitalized. Therefore, the clusters within the heatmap showed some differences generally and regarding specific words, such as ‘neighborhood’, as discussed above. Consequently, the findings align with the conclusion of the study by Kim et al. (2021) that categorization is partially cross-lingual. They also point out that differing contexts of a word’s use might lead to differences between the languages. Interestingly, while Kim et al. (2021) assessed Korean and German words, which stem from distinct language families, this finding seems to prevail between Germanic languages, such as English and German. In line with this, Majid et al. (2007) found in their study of semantic categories for cutting and breaking events that even within Germanic languages differences can occur. Nevertheless, in the present study, these differences did not significantly affect the correspondence with the concepts in the cerebral cortex, as found by Huth et al. (2016).

Hence, as previous research has indicated, some differences occur in the categorization between languages, possibly due to slightly different meanings or contexts of the translated word. However, these differences did not lead to a significant difference in the applicability of the findings by Huth et al. (2016). Therefore, their findings might be equally transferable to other languages. Notably, German and English are part of the same language family and therefore, the applicability to languages that share less overlap might be more affected.

5.3 Limitations of the Study

Although the study was able to yield answers to the research questions, it must be noted that there were a few limitations that might have had an impact on the findings. Firstly, the card sorting was conducted online due to the current Covid-19 restrictions. Because of this, the words were given in a list in the survey rather than physical cards, as traditionally in card sorting. It was noted that the card sorting, especially on smaller monitors, was not easy to navigate. For instance, to sort a word that was presented lower in the list, it had to be moved several times to reach the higher boxes. To counteract this, the participants were instructed to use their computer, instead of a phone. Further, the words were presented in a random order to prevent certain words from being frequently grouped into lower boxes and others into higher boxes. Despite this, the possibility that the online format might have impacted the findings cannot be fully excluded.

Secondly, especially for the comparison between the English and German sorting, it might have had an impact that most participants were bilingual and possessed skills in both languages. It has been found that proficiency in several languages does not lead to separate but to shared semantic representations of concepts (Francis, 2020). This also presents a difference from the original study by Huth et al. (2016), as almost all participants in the present study had another mother tongue than English.

Lastly, due to the scope of the present study, only a part of the categories was looked at and only (comparatively) few concepts could be selected for the card sorting task. Therefore, it remains open whether the combination of other domains found by Huth et al. (2016) and the selection of other words would yield different results.

5.4 Future Research

Following the above-mentioned limitations, several recommendations for future research can be made. For instance, it is recommended to test whether the differences between the categorization across English and German remain non-significant when participants are less proficient in the other language under investigation. Further, since the proficiency of other languages can affect the semantic representation of words, it is advised to test whether native monolingual English speakers produce fewer differences to the domains found in English participants by Huth et al. (2016). Furthermore, in a similar vein, the compared languages stemmed from the Germanic language family. Thus, it is recommended to test whether languages from other language families also share overlap with the domains that are reported for English. Lastly, since the scope of the study was limited to six of the eleven categories reported by Huth et al. (2016), it is advisable to test the remaining categories and whether different combinations affect the categorization. For instance, in the present study, the categories of *outdoor* and *tactile* were mixed. Therefore, presenting different combinations might also affect the outcome of correspondence with the semantic atlas. Generally, it is advised to place greater emphasis on the context and associations of concepts when researching conceptual knowledge. In addition, the findings in the German condition suggest that nouns were largely sorted separately from verbs and adjectives because of their different appearance. Research could further investigate whether differences in word appearance affect categorization.

6. Conclusion

To summarize, the categorization of words by participants aligns to a certain extent to the semantic representation in the cerebral cortex, as reported by Huth et al. (2016). However, not all of the six researched categories could be equally well reproduced. Only the categorization of the words from the social, number, place, and mental categories resembled the categories of Huth et al. (2016) to a large extent. Especially the category of tactile could not be recreated in the card sorting task. Rather, there seemed to be other topics that unified the words, such as a similar association with specific contexts. This indicates that the semantic representation within the brain cannot fully account for the categorization of concepts. In addition to this, it was found that the correspondence of the formed categories to the concepts in the cerebral cortex is not significantly affected by translating the words to German. This suggests that, at least within Germanic languages, the categories of social, number, place, outdoor, tactile, and mental seem to prevail and correspond to the semantic representation in the brain to an equal extent.

References

Borghesani, V., Buiatti, M., Eger, E., & Piazza, M. (2018). Conceptual and perceptual dimensions of word meaning are recovered rapidly and in parallel during reading. *Journal of Cognitive Neuroscience, 31*(3), 1-14. doi: 10.1162/jocn_a_01328

Dehghani, M., Boghrati, R., Man, K., Hoover, J., Gimbel, S. I., Vaswani, A., Zevin, J. D., Immordino-Yang, M. H., Gordon, A. S., Damasio, A., & Kaplan, J. T. (2017). Decoding the neural representation of story meanings across languages. *Human Brain Mapping, 38*(12), 6096-6106. doi: 10.1002/hbm.23814

Deniz, F., Nunez-Elizalde, A. O, Huth, A. G., & Gallant, J. L. (2019). The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *Journal of Neuroscience, 39*(39), 7722-7736. doi: 10.1523/JNEUROSCI.0675-19.2019

Doran, R., Böhm, G., & Hanss, D. (2018). Using card sorting to explore the mental representation of energy transition pathways among laypeople. *Frontiers in Psychology, 9*(2322). doi: 10.3389/fpsyg.2018.02322

Francis, W. S. (2020). Shared core meanings and shared associations in bilingual semantic memory: evidence from research on implicit memory. *International Journal of Bilingualism, 24*(3), 464-477. doi: 0.1177/1367006918814375

Hoffmann, P. (2018). An individual differences approach to semantic cognition: divergent effects of age on representation, retrieval and selection. *Scientific Reports, 8*(1). doi: 10.1038/s41598-018-26569-0

Hudson, W. (2005). Playing your cards right: getting the most from card sorting for navigation design. *Interactions, 12*(5), 56-58. doi: 10.1145/1082369.1082410

Huth, A. G. (n.d.). *Brain viewer.* Retrieved from https://gallantlab.org/huth2016/

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*, 453-458. doi: 10.1038/nature17637

Jefferies, E. (2013). The neural basis of semantic cognition: converging evidence from neuropsychology, neuroimaging and TMS. *Cortex, 49*(3), 611-625. doi: 10.1016/j.cortex.2012.10.008

Joshi, A., Kale, S., Chandel, S., & Kal, D. K. (2015). Likert scale: explored and explained. *British Journal of Applied Science & Technology, 7*(4), 396-403. doi: 10.9734/BJAST/2015/14975

Kim, M., Müller, H. M., & Weiss, S. (2021). What you “mean” is not what I “mean”: categorization of verbs by Germans and Koreans using the semantic differential. *Lingua, 252*. doi: 10.1016/j.lingua.2020.103012

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*(1), 42-55. doi: 10.1038/nrn.2016.150

Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory. *Annals of the New York Academy of Sciences, 1124*, 61-76. doi: 10.1196/annals.1440.006

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*, 5-55.

Majid, A., Gullberg, M., van Staden, M., Bowerman, M. (2007). How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four germanic languages. *Cognitive Linguistics, 18*(2), 179-194. doi: 10.1515/COG.2007.007

Martin, A. (2016). GRAPES – Grounding representations in action, perception, and emotion systems: How object properties and categories are represented in the human brain. *Psychonomic Bulletin & Review, 23*, 979-990. doi: 10.3758/s13423-015-0842-3

Patterson, K., & Lambon Ralph, M. A. (2016). The hub-and-spoke hypothesis of semantic memory. In G. Hickok & S. Small (Eds.). *Neurobiology of Language* (pp. 765-775). Cambridge, Massachusetts: Academic Press.

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. doi: 10.1038/nrn2277

Plato. (1948). *Plato's Meno* (B. Jowett, Trans.). Retrieved from http://classics.mit.edu/Plato/meno.html (Original work published ca. 380 B.C.E.)

Rice, C. A., Beekhuisen, B., Dubrovsky, V., Stevenson, S., & Armstrong, B. C. (2019). A comparison of homonym meaning frequency estimates derived from movie and television subtitles, free association, and explicit ratings. *Behaviour Research Methods, 51*, 1399-1425. doi: 10.3758/s13428-018-1107-7

Rogers, T. T., Patterson, K., Jefferies, E., & Lambon Ralph, M. A. (2015). Disorders of representation and control in semantic cognition: effects of familiarity, typicality, and specificity. *Neuropsychologia, 76*, 220-239. doi: 10.1016/j.neuropsychologia.2015.04.0

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), 1-19. doi: 10.1080/0144929X.2016.1157207

Silveri, M. C., Brita, A. C., Liperoti, R., Piludu, F., & Colosimo, C. (2018). What is semantic in semantic dementia? The decay of knowledge of physical entities but not of verbs, numbers and body parts. *Aphasiology, 32*(9), 989-1009. doi: 10.1080/02687038.2017.1387227

Smith, J. I., Combs, E. D., Nagami, P. H., Alto, V. M., Goh, H. G., Gourdet, M. A. A., Hough, C. M., Nickell, A. E., Peer, A. G., Coley, J. D., & Tanner, K. D. (2013). Development of the biology card sorting task to measure conceptual expertise in biology. *CBE Life Sciences Education, 12*(4), 628-644. doi: 10.1187/cbe.13-05-0096

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences, 5*(6), 244-252. doi: 10.1016/s1364-6613(00)01651-x

Wood, J. R., & Wood, L. E. (2008). Card sorting: current practices and beyond. *Journal of Usability Studies*, *4*(1), 1–6.

Appendices

Appendix A

Item Lists

Experimental Items

Item no.	English word (original)	German word (translation)	Category	Voxel Numbers	Location	Reliability score
1	thoughts	Gedanken	mental	12,91,60	left, occipital cortex	Not bad, pretty reliable (0.213)
2	overwhelmed	überwältigt	mental	12,91,60	left, occipital cortex	Not bad, pretty reliable (0.213)
3	consciousness	Bewusstsein	mental	12,91,60	left, occipital cortex	Not bad, pretty reliable (0.213)
4	anxiety	Angst	mental	12,91,60	left, occipital cortex	Not bad, pretty reliable (0.213)
5	loneliness	Einsamkeit	mental	12,91,60	left, occipital cortex	Not bad, pretty reliable (0.213)
6	laughing	lachen	mental	22,20,42	right, prefrontal cortex	Good, very reliable (0.340)
7	calmly	ruhig	mental	22,20,42	right, prefrontal cortex	Good, very reliable (0.340)
8	loudly	laut	mental	22,20,42	right, prefrontal cortex	Good, very reliable (0.340)
9	startled	aufgeschreckt	mental	22,20,42	right, prefrontal cortex	Good, very reliable (0.340)
10	each	jedes	number	20,80,38	right, parietal cortex	Good, very reliable (0.399)
11	ten	zehn	number	20,80,38	right, parietal cortex	Good, very reliable (0.399)
12	six	sechs	number	20,80,38	right, parietal cortex	Good, very reliable (0.399)
13	quarter	Viertel	number	20,80,38	right, parietal cortex	Good, very reliable (0.399)
14	per	je	number	24,34,62	left, frontal cortex	Good, very reliable (0.376)
15	set	Set	number	24,34,62	left, frontal cortex	Good, very reliable (0.376)
16	plus	plus	number	24,34,62	left, frontal cortex	Good, very reliable (0.376)
17	collect	sammeln	number	24,34,62	left, frontal cortex	Good, very reliable (0.376)
18	daughter	Tochter	social	22,32,58	left, frontal cortex	Excellent, extremely reliable (0.421)
19	family	Familie	social	22,32,58	left, frontal cortex	Excellent, extremely reliable (0.421)
20	relatives	Verwandte	social	22,32,58	left, frontal cortex	Excellent, extremely reliable (0.421)
21	police	Polizei	social	22,32,58	left, frontal cortex	Excellent, extremely reliable (0.421)
22	wife	Ehefrau	social	15,80,27	right, parietal cortex	Good, very reliable (0.370)
23	husband	Ehemann	social	15,80,27	right, parietal cortex	Good, very reliable (0.370)
24	refused	verweigert	social	15,80,27	right, parietal cortex	Good, very reliable (0.370)
25	died	gestorben	social	15,80,27	right, parietal cortex	Good, very reliable (0.370)
26	parent	Elternteil	social	15,80,27	right, parietal cortex	Good, very reliable (0.370)
27	brittle	spröde	tactile	21,82,54	left, lateral parietal cortex	Not bad, pretty reliable (0.264)
28	melting	schmelzen	tactile	21,82,54	left, lateral parietal cortex	Not bad, pretty reliable (0.264)
29	fluid	flüssig	tactile	21,82,54	left, lateral parietal cortex	Not bad, pretty reliable (0.264)

30	flame	Flamme	tactile	21,82,54	left, lateral parietal cortex	Not bad, pretty reliable (0.264)
31	swallow	schlucken	tactile	18,66,21	right, parietal cortex	Not bad, pretty reliable (0.204)
32	grip	Griff	tactile	18,66,21	right, parietal cortex	Not bad, pretty reliable (0.204)
33	touches	berühren	tactile	18,66,21	right, parietal cortex	Not bad, pretty reliable (0.204)
34	squeeze	quetschen	tactile	18,66,21	right, parietal cortex	Not bad, pretty reliable (0.204)
35	clouds	Wolken	outdoor	15,17,29	right, prefrontal cortex	Good, very reliable (0.304)
36	waves	Wellen	outdoor	15,17,29	right, prefrontal cortex	Good, very reliable (0.304)
37	drifting	treiben	outdoor	15,17,29	right, prefrontal cortex	Good, very reliable (0.304)
38	flooded	überflutet	outdoor	15,17,29	right, prefrontal cortex	Good, very reliable (0.304)
39	breathe	atmen	outdoor	14,16,30	right, prefrontal cortex	Not bad, pretty reliable (0.2)
40	atmosphere	Atmosphäre	outdoor	14,16,30	right, prefrontal cortex	Not bad, pretty reliable (0.2)
41	heavens	Himmel	outdoor	14,16,30	right, prefrontal cortex	Not bad, pretty reliable (0.2)
42	explosions	Explosionen	outdoor	14,16,30	right, prefrontal cortex	Not bad, pretty reliable (0.2)
43	driveway	Einfahrt	place	16,85,33	right, occipital cortex	Good, very reliable (0.312)
44	unlocked	unverschlossen	place	16,85,33	right, occipital cortex	Good, very reliable (0.312)
45	bus	Bus	place	16,85,33	right, occipital cortex	Good, very reliable (0.312)
46	parking	parken	place	12,76, 55	left, occipital cortex	Good, very reliable (0.357)
47	neighborhood	Wohngegend	place	12,76, 55	left, occipital cortex	Good, very reliable (0.357)
48	attendance	Anwesenheit	place	18,60,52	left, parietal cortex	Not bad, pretty reliable (0.28)
49	university	Universität	place	18,60,52	left, parietal cortex	Not bad, pretty reliable (0.28)
50	students	Schüler	place	18,60,52	left, parietal cortex	Not bad, pretty reliable (0.28)

Filler Items

Item no.	English word (original)	German word (translation)	Category	Voxel Numbers	Location	Reliability score
51	coloured	farbig	visual	19,32,72	left, frontal cortex	Excellent, extremely reliable (0.466)
52	sleeve	Ärmel	visual	19,32,72	left, frontal cortex	Excellent, extremely reliable (0.466)
53	brown	braun	visual	19,32,72	left, frontal cortex	Excellent, extremely reliable (0.466)
54	boots	Stiefel	visual	16,27,26	right, frontal cortex	Good, very reliable (0.372)
55	coat	Mantel	visual	16,27,26	right, frontal cortex	Good, very reliable (0.372)
56	olive	oliv	visual	16,27,26	right, frontal cortex	Good, very reliable (0.372)
57	evidence	Beweis	violence	14,32,74	left, frontal cortex	Not bad, pretty reliable (0.294)
58	harm	Schaden	violence	14,32,74	left, frontal cortex	Not bad, pretty reliable (0.294)
59	punished	bestraft	violence	14,32,74	left, frontal cortex	Not bad, pretty reliable (0.294)
60	victim	Opfer	violence	14,32,74	left, frontal cortex	Not bad, pretty reliable (0.294)
61	innocent	unschuldig	violence	14,32,74	left, frontal cortex	Not bad, pretty reliable (0.294)
62	cop	Polizist	person	16,13,40	right, prefrontal cortex	Good, very reliable (0.322)
63	sheriff	Sheriff	person	16,13,40	right, prefrontal cortex	Good, very reliable (0.322)
64	robbery	Raub	person	16,13,40	right, prefrontal cortex	Good, very reliable (0.322)
65	sons	Söhne	person	16,13,40	right, prefrontal cortex	Good, very reliable (0.322)
66	monday	Montag	time	14,87,65	left, extrastriate cortex	Excellent, extremely reliable (0.46)
67	week	Woche	time	14,87,65	left, extrastriate cortex	Excellent, extremely reliable (0.46)
68	college	Hochschule	time	14,87,65	left, extrastriate cortex	Excellent, extremely reliable (0.46)
69	school	Schule	time	14,87,65	left, extrastriate cortex	Excellent, extremely reliable (0.46)
70	spent	verbracht	time	14,87,65	left, extrastriate cortex	Excellent, extremely reliable (0.46)

Appendix B

Demographics of Participants Across the Two Conditions (N=42)

Demographics	English Items *(n=23)*	German Items *(n=19)*
Gender		
Female	18 (78.26%)	13 (68.42%)
Male	5 (21.74%)	6 (31.58%)
Nationality		
German	13 (60.87%)	19 (100%)
Dutch	6 (26.09%)	-
other	3 (13.04%)	-
Age		
Mean	20.17	25.95
SD	1.81	9.60
Range	18-24	20-53
Level of Education		
High school diploma	20 (86.96%)	12 (63.16%)
Bachelor's degree	3 (13.04%)	4 (21.05%)
Master's degree	-	1 (5.26%)
Other	-	2 (10.53%)
Level of Language skills (English/German)		
A1/A2	-	-
B1	-	-
B2	8 (34.78%)	-
C1	8 (34.78%)	-
C2	6 (26.09%)	1 (5.26%)
Native	1 (4.35%)	18 (94.74%)

Appendix C

Informed Consent

Introduction

You are participating in the study '**Conceptual learning**'. The research is conducted to study potential relations between words. Please use a laptop or computer to fill out the survey. The survey will take about 30 minutes to complete but you might be finished earlier.
Thank you for participating!

Informed consent

Purpose of the research
The research aims towards studying the potential relations between items. You will be presented with a set of words/items and asked to group them together. Further, you will be asked to fill out a questionnaire in which you have to judge the relation between two words. There are no right or wrong answers, just choose the option that you find most suitable.

All data you provide within the study is **anonymous** and used for a Bachelor thesis within the Psychology programme at the University of Twente. After completion the overall results will be stored and used for future projects.
If you are interested in retrieving the report of the study, you can send a mail to: l.c.reiter@student.utwente.nl

Risks & Benefits
There are no know risks or benefits associated with participating in the study. The research project has been reviewed and approved by the BMS Ethics Committee of the University of Twente.

Your rights
Participation in the study is **voluntarily** and you can withdraw from the study at any time, without having to give a reason.

If you have **questions or remarks** about the study, you can contact Luisa Reiter via mail: l.c.reiter@student.utwente.nl

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher, please contact the Secretary of the Ethics Committee of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-bms@utwente.nl

○ I have read the abovementioned information and consent to participating in this study

Appendix D

Demographic Questions

Please answer a few questions regarding your demographic data.

Please indicate your Age:

____

Your Gender

- ○ Female
- ○ Male
- ○ Other
- ○ Prefer not to say

Highest level of Education

- ○ High school diploma
- ○ Bachelor's degree
- ○ Master's degree
- ○ Doctorate
- ○ Other ____
- ○ Prefer not to say

Nationality

- ○ Dutch
- ○ German
- ○ Other ____
- ○ Prefer not to say

Level of English

- ○ Basic communication skills (A1/A2)
- ○ Good command (B1)
- ○ Very good command (B2)
- ○ Excellent command (C1)
- ○ Near native/fluid (C2)
- ○ Native

Level of German

- ○ I don't speak German
- ○ Basic communication skills (A1/A2)
- ○ Good command (B1)
- ○ Very good command (B2)
- ○ Excellent command (C1)
- ○ Near native/fluid (C2)
- ○ Native

Appendix E
Questionnaire English

You will now be presented with a pair of words and judge to which extent they are related.

How do you judge the relation between this pair of words:

answer options:

- ○ highly related
- ○ related
- ○ neutral
- ○ not related
- ○ highly not related

family - social

harm - outdoor

quarter - number

coloured - tactile

punished - place

heavens - outdoor

calmly - mental

six - number

husband - social

startled - mental

sheriff - number

relatives - social

parent - social

squeeze - tactile

clouds – outdoor

ten – number

sons – tactile

victim – number

neighborhood – place

anxiety – mental

evidence – mental

flame – tactile

unlocked – place

atmosphere – outdoor

each – number

robbery – tactile

overwhelmed – mental

explosions – outdoor

spent – social

university – place

touches – tactile

students – place

innocent – mental

grip – tactile

daughter – social

wife – social

laughing – mental

coat – social

driveway – place

drifting – outdoor

police – social

cop – mental

waves – outdoor

thoughts – mental

consciousness – mental

bus – place

per – number

set – number

died - social

olive - number

melting - tactile

attendance - place

school – tactile

loudly - mental

parking - place

plus - number

fluid - tactile

brittle - tactile

refused - social

collect - number

week - place

breathe - outdoor

monday - social

boots - number

flooded - outdoor

swallow - tactile

college - outdoor

loneliness - mental

brown - place

sleeve – outdoor

Appendix F

Questionnaire German

You will now be presented with a pair of words and judge to which extent they are related.

How do you judge the relation between this pair of words:

answer options:

○ highly related

○ related

○ neutral

○ not related

○ highly not related

Familie - sozial

Schaden – draußen

Viertel – Nummer

farbig – haptisch

bestraft – Ort

Himmel – draußen

ruhig – mental

sechs – Nummer

Ehemann – sozial

aufgeschreckt – mental

Sheriff – Nummer

Verwandte – sozial

Elternteil – sozial

quetschen – haptisch

Wolken – draußen

zehn – Nummer

Söhne – haptisch

Opfer – Nummer

Wohngegend – Ort

Angst – mental

Beweis – mental

Flamme – haptisch

unverschlossen - Ort

Atmosphäre – draußen

jedes – Nummer

Raub – haptisch

überwältigt – mental

Explosionen – draußen

verbracht – sozial

Universität – Ort

berühren – haptisch

Schüler – Ort

unschuldig – mental

Griff – haptisch

Tochter – sozial

Ehefrau – sozial

lachen – mental

Mantel – sozial

Einfahrt – Ort

treiben – draußen

Polizei – sozial

Polizist – mental

Wellen – draußen

Gedanken – mental

Bewusstsein – mental

Bus – Ort

je – Nummer

Set – Nummer

gestorben – sozial

oliv – Nummer

schmelzen – haptisch

Anwesenheit – Ort

Schule – haptisch

laut – mental

parken – Ort

plus – Nummer

flüssig – haptisch

spröde – haptisch

verweigert – sozial

sammeln – Nummer

Woche – Ort

atmen – draußen

Montag – sozial

Stiefel – Nummer

überflutet – draußen

schlucken – haptisch

Hochschule – draußen

Einsamkeit – mental

braun – Ort

Ärmel – draußen

Appendix G

Python Script to Transform the Card Sorting Data into Jaccard Scores

```
#open csv file with all participants card sorting data (incl. demographics & item-category ratings)
import csv

with open ('Data_English_ed.csv') as file:
    reader = csv.reader(file)
    array =[]

    for row in reader:
        array.append(row)

    result = {}
    header =['Items']
    #indicate no. of participants+1 (or lines including heading) in Excel file
    lines=23
    #indicate no. of items in card sorting task
    NumberItems=50

    for index in range(1,lines):
        print(index)
        groups = [[],[],[],[],[],[],[],[],[],[],[],[],[],[],[],[],[]]
        groupCount = 0

        for i in range(len(array[index])):

            if i == 0: #deleting the demographic data of the participants
                firstString = array[index][0].split(";")
                groups[groupCount].append(firstString[len(firstString)-1])

            elif groupCount<16: #when words are in cvs file seperated with ; they are read as
                                groups
                splitString = array[index][i].split(";")
                count = 0
                for piece in range(len(splitString)):
                    # increase the group index if the word is part of the next group
                    # if the loop loops multiple times and there are not yet 16 groups exist, a new group is created
                    if count >= 1 and groupCount<16:
                        groupCount+=1
                    count+=1
                    groups[groupCount].append(splitString[piece]) #Word is sorted into current
                                                                  group
        for i in range(len(groups)):
            if i <= 15 and len(groups[i]) > 1: #are there min. 2 words in group? --> otherwise
                                               ignore if only one; ignore word-category ratings
                for j in range(len(groups[i])):
                    if j+1 < len(groups[i]):
                        for after in range(j+1,len(groups[i])):
                            #print(str(groups[i][j]) + str(groups[i][after]))
                            pair1 = str(groups[i][j]) + "_" +str(groups[i][after])
                            pair2 = str(groups[i][after]) + "_" +str(groups[i][j])
```

```
                if pair1 not in result:
                    result[pair1] = 1
                else:
                    result[pair1] += 1

                if pair2 not in result:
                    result[pair2] = 1
                else:
                    result[pair2] += 1

# create a csv file to create final table and open it
with open('map2_english.csv', mode='w') as csvfile:
    writer = csv.writer(csvfile, delimiter=';', lineterminator='\n', )

# create header for the final table
    for key in result:
        firstWord = key.split("_")[0]
        secondWord = key.split("_")[1]

        if key.split("_")[0] not in header:
            header.append(key.split("_")[0])

    writer.writerow(header)

# extracting the values of the word combinations into the corresponding column x row
    for currentRow in header:
        if currentRow != "Items":
            numbers = [0]*NumberItems
            for key in result:
                firstWord = key.split("_")[0]
                secondWord = key.split("_")[1]
```

Appendix H

R Script to Create Heatmaps & Draw Comparisons Between the Conditions

```
##Create heatmaps
library(gplots)
library(RColorBrewer)

#open data file
data <-read.csv2("map_english.csv")
View (data)
#transforming English data into numerical format
data <- data.matrix(data[,1:ncol(data)])
# define colors of heatmap: red for high numbers
my_palette <- colorRampPalette(c("yellow","red"))(n = 299)
#create heatmap
heatmap.2(data, col = my_palette, density.info="none", trace="none",
        revC = TRUE, cexRow = 0.4, cexCol = 0.6, main="Concepts English")

#open 2nd data file
data2 <-read.csv2("map_german.csv")
View (data2)
#transforming German data into numerical format
data2 <- data.matrix(data2[,1:ncol(data2)])
#create heatmap
heatmap.2(data2, col = my_palette, density.info="none", trace="none",
        revC = TRUE, cexRow = 0.4, cexCol = 0.6, main="Concepts German")

##Comparison between Conditions
install.packages("ggpubr")
install.packages("psych")
library(ggpubr)
library (psych)

data_ENGLISH <-read.csv2("Data_English_R2.csv") #questionnaire data
View(data_ENGLISH)
data_GERMAN <-read.csv2("Data_German_R.csv") #questionnaire data
View(data_GERMAN)
data_Jaccard <-read.csv2("Mappe1.csv") #list of Jaccard scores of both conditions
View(data_Jaccard)

#1st the analysis of the Jaccard scores from the Card Sorting:
# the added Jaccard scores of each participant divided by the possible total score/combinations
data_Jaccard$English <- data_Jaccard$English/22
data_Jaccard$German <-data_Jaccard$German/19

shapiro.test(data_Jaccard$English)
ggqqplot(data_Jaccard$English)
shapiro.test(data_Jaccard$German)
ggqqplot(data_Jaccard$German)

#checking whether the Jaccard scores differ sign. between the conditions
test <- wilcox.test(data_Jaccard$English, data_Jaccard$German, exact=FALSE)

#if z-score and effect size are required
test
Zstat <- qnorm(test$p.value/2)
Zstat
abs(Zstat)/sqrt(184)
```

```
#2nd the analysis of the questionnaire
#compute variable for mean category scores
data_ENGLISH$Qu_Mental <- (data_ENGLISH$thoughts + data_ENGLISH$overwhelmed +
data_ENGLISH$consciousness + data_ENGLISH$anxiety + data_ENGLISH$loneliness +
data_ENGLISH$laughing + data_ENGLISH$calmly + data_ENGLISH$loudly + data_ENGLISH$startled)/9
data_ENGLISH$Qu_Number <- (data_ENGLISH$each + data_ENGLISH$ten + data_ENGLISH$six +
data_ENGLISH$quarter + data_ENGLISH$per + data_ENGLISH$set + data_ENGLISH$plus+
data_ENGLISH$collect)/8
data_ENGLISH$Qu_Social <- (data_ENGLISH$daughter + data_ENGLISH$family +
data_ENGLISH$relatives + data_ENGLISH$police + data_ENGLISH$wife + data_ENGLISH$husband +
data_ENGLISH$refused + data_ENGLISH$died + data_ENGLISH$parent)/9
data_ENGLISH$Qu_Tactile <- (data_ENGLISH$brittle + data_ENGLISH$melting + data_ENGLISH$fluid
+ data_ENGLISH$flame + data_ENGLISH$swallow + data_ENGLISH$grip + data_ENGLISH$touches +
data_ENGLISH$squeeze)/8
data_ENGLISH$Qu_Outdoor <- (data_ENGLISH$clouds + data_ENGLISH$waves +
data_ENGLISH$drifting + data_ENGLISH$flooded + data_ENGLISH$breathe +
data_ENGLISH$atmosphere + data_ENGLISH$heavens + data_ENGLISH$explosions)/8
data_ENGLISH$Qu_Place <- (data_ENGLISH$driveway + data_ENGLISH$unlocked +
data_ENGLISH$bus + data_ENGLISH$parking + data_ENGLISH$neighborhood +
data_ENGLISH$attendance + data_ENGLISH$university + data_ENGLISH$students)/8
data_ENGLISH$Qu_Collapsed <- (data_ENGLISH$Qu_Mental + data_ENGLISH$Qu_Number +
data_ENGLISH$Qu_Social + data_ENGLISH$Qu_Tactile + data_ENGLISH$Qu_Outdoor +
data_ENGLISH$Qu_Place)/6

#calulate descriptive statistics for categories in questionnaire
describe(data_ENGLISH$Qu_Mental)
describe(data_ENGLISH$Qu_Number)
describe(data_ENGLISH$Qu_Social)
describe(data_ENGLISH$Qu_Tactile)
describe(data_ENGLISH$Qu_Outdoor)
describe(data_ENGLISH$Qu_Place)
describe(data_ENGLISH$Qu_Collapsed)

# check for normality of collapsed variable
shapiro.test(data_ENGLISH$Qu_Collapsed)
ggqqplot(data_ENGLISH$Qu_Collapsed)

#comparing the scores to the control items
data_ENGLISH$Qu_Me_Control <- (data_ENGLISH$cop + data_ENGLISH$innocent +
data_ENGLISH$evidence)/3
data_ENGLISH$Qu_Nu_Control <- (data_ENGLISH$sheriff + data_ENGLISH$victim +
data_ENGLISH$olive + data_ENGLISH$boots)/4
data_ENGLISH$Qu_So_Control <- (data_ENGLISH$spent + data_ENGLISH$Monday +
data_ENGLISH$coat)/3
data_ENGLISH$Qu_Ta_Control <- (data_ENGLISH$school + data_ENGLISH$robbery +
data_ENGLISH$sons + data_ENGLISH$coloured)/4
data_ENGLISH$Qu_Ou_Control <- (data_ENGLISH$college + data_ENGLISH$harm +
data_ENGLISH$sleeve)/3
data_ENGLISH$Qu_Pl_Control <- (data_ENGLISH$week + data_ENGLISH$punished +
data_ENGLISH$brown)/3
data_ENGLISH$Qu_allControl <- (data_ENGLISH$Qu_Me_Control + data_ENGLISH$Qu_Nu_Control +
data_ENGLISH$Qu_So_Control + data_ENGLISH$Qu_Ta_Control + data_ENGLISH$Qu_Ou_Control +
data_ENGLISH$Qu_Pl_Control)/6

#desriptive statistics
describe(data_ENGLISH$Qu_allControl)

# check for normality of control variable
shapiro.test(data_ENGLISH$Qu_allControl)
ggqqplot(data_ENGLISH$Qu_allControl)
```

```
# checking whether Control Items scores differ significantly from Exiperimental items
t.test(data_ENGLISH$Qu_Collapsed, data_ENGLISH$Qu_allControl, paired = TRUE, alternative =
"two.sided")

data_GERMAN$Qu_Mental <- (data_GERMAN$Gedanken + data_GERMAN$überwäötigt +
data_GERMAN$Bewusstsein + data_GERMAN$Angst + data_GERMAN$Einsamkeit +
data_GERMAN$lachen + data_GERMAN$ruhig + data_GERMAN$laut + data_GERMAN$aufgeschreckt)/9
data_GERMAN$Qu_Number <- (data_GERMAN$jedes + data_GERMAN$zehn + data_GERMAN$sechs +
data_GERMAN$viertel + data_GERMAN$je + data_GERMAN$Set + data_GERMAN$plus+
data_GERMAN$sammeln)/8
data_GERMAN$Qu_Social <- (data_GERMAN$Tochter + data_GERMAN$Familie +
data_GERMAN$Verwandte + data_GERMAN$Polizei + data_GERMAN$Ehefrau +
data_GERMAN$Ehemann + data_GERMAN$verweigert + data_GERMAN$gestorben +
data_GERMAN$Elternteil)/9
data_GERMAN$Qu_Tactile <- (data_GERMAN$spröde + data_GERMAN$schmelzen +
data_GERMAN$flüssig + data_GERMAN$Flamme + data_GERMAN$schlucken + data_GERMAN$Griff +
data_GERMAN$berühren + data_GERMAN$quetschen)/8
data_GERMAN$Qu_Outdoor <- (data_GERMAN$Wolken + data_GERMAN$Wellen +
data_GERMAN$treiben + data_GERMAN$überflutet + data_GERMAN$atmen +
data_GERMAN$Atmosphäre + data_GERMAN$Himmel + data_GERMAN$Explosionen)/8
data_GERMAN$Qu_Place <- (data_GERMAN$Einfahrt + data_GERMAN$unverschlossen +
data_GERMAN$Bus + data_GERMAN$parken + data_GERMAN$Wohngegend +
data_GERMAN$Anwesenheit + data_GERMAN$Universität + data_GERMAN$Schüler)/8
data_GERMAN$Qu_Collapsed <- (data_GERMAN$Qu_Mental + data_GERMAN$Qu_Number +
data_GERMAN$Qu_Social + data_GERMAN$Qu_Tactile + data_GERMAN$Qu_Outdoor +
data_GERMAN$Qu_Place)/6
data_GERMAN$Control <-(data_GERMAN$farbig + data_GERMAN$Ärmel + data_GERMAN$braun +
data_GERMAN$Stiefel + data_GERMAN$Mantel + data_GERMAN$oliv + data_GERMAN$Beweis +
data_GERMAN$Schaden + data_GERMAN$bestraft + data_GERMAN$Opfer +
data_GERMAN$unschuldig + data_GERMAN$Polizist + data_GERMAN$Sheriff + data_GERMAN$Raub
+ data_GERMAN$Söhne + data_GERMAN$Montag + data_GERMAN$Woche +
data_GERMAN$Hochschule + data_GERMAN$Schule + data_GERMAN$verbracht)/20

#calulate descriptive statistics for categories in questionnaire
describe(data_GERMAN$Qu_Mental)
describe(data_GERMAN$Qu_Number)
describe(data_GERMAN$Qu_Social)
describe(data_GERMAN$Qu_Tactile)
describe(data_GERMAN$Qu_Outdoor)
describe(data_GERMAN$Qu_Place)
describe(data_GERMAN$Qu_Collapsed)

# check for normality
shapiro.test(data_GERMAN$Qu_Collapsed)
ggqqplot(data_GERMAN$Qu_Collapsed)

shapiro.test(data_GERMAN$Control)
ggqqplot(data_GERMAN$Control)

#check if German experimental items better than control
describe (data_GERMAN$Control)
var.test(data_GERMAN$Qu_Collapsed, data_GERMAN$Control)
t.test(data_GERMAN$Qu_Collapsed, data_GERMAN$Control, var.equal = TRUE)

#compare questionnaire English - German
var.test(data_GERMAN$Qu_Collapsed, data_ENGLISH$Qu_Collapsed)
t.test(data_GERMAN$Qu_Collapsed, data_ENGLISH$Qu_Collapsed, var.equal = TRUE)
```

Appendix I
Questionnaire Scores for Each Item in the English (n=22) and German (n=19) Condition

	English		German	
Item	Mean	SD	Mean	SD
Mental				
thoughts	1.23	0.42	1.16	0.36
overwhelmed	1.41	0.49	1.79	0.69
consciousness	1.14	0.34	1.32	0.73
anxiety	1.27	0.45	1.42	0.59
loneliness	1.50	0.58	1.47	0.50
laughing	2.18	0.78	2.21	0.77
calmly	1.55	0.50	1.79	0.61
loudly	3.23	1.00	2.95	1.05
startled	2.09	0.79	2.05	0.83
Number				
each	2.18	0.94	2.63	1.22
ten	1.05	0.21	1.16	0.36
six	1.05	0.21	1.05	0.22
quarter	1.55	0.58	1.68	0.65
per	1.95	0.88	2.26	0.96
set	1.95	0.98	2.78	1.15
plus	1.64	0.64	1.89	1.02
collect	2.59	1.07	2.74	1.02
Social				
daughter	1.59	0.65	1.74	0.71
family	1.23	0.42	1.32	0.46
relatives	1.36	0.48	1.58	0.59
police	2.41	1.07	2.68	0.98
wife	1.50	0.58	1.79	0.69
husband	1.45	0.50	1.84	0.74
refused	2.72	1.01	2.84	1.09
died	2.68	0.87	2.95	0.94
parent	1.45	0.50	1.47	0.60

Tactile				
brittle	2.45	0.89	2.32	1.03
melting	3.00	0.90	3.37	1.09
fluid	2.68	0.92	2.53	1.04
flame	2.82	0.89	3.11	1.21
swallow	3.14	1.10	3.21	1.24
grip	1.77	0.73	1.68	1.03
touches	1.32	0.70	1.42	0.82
squeeze	1.82	0.83	2.47	1.23
Outdoor				
clouds	1.45	0.58	1.16	0.36
waves	1.55	0.72	1.79	0.77
drifting	2.23	0.95	2.63	0.93
flooded	2.14	0.69	2.16	0.74
breathe	2.55	1.12	2.58	1.09
atmosphere	1.64	0.64	2.05	1.05
heavens	2.00	0.74	1.11	0.31
explosions	2.23	0.79	2.74	1.07
Place				
driveway	1.86	0.69	1.95	0.94
unlocked	2.36	1.07	3.32	1.22
bus	2.59	0.94	2.53	1.04
parking	1.68	0.70	2.05	0.69
neighborhood	1.55	0.58	1.26	0.44
attendance	2.45	1.20	2.11	0.55
university	1.41	0.49	1.42	0.49
students	2.91	1.38	3.21	1.06