Kausale Analysen auf fMRT-Daten

Bachelorarbeit

an der Fakultät für Mathematik, Informatik und Statistik

Ludwig-Maximilians-Universität München

Institut für Statistik

Arbeitsgruppe: Data Science

Prof. Dr. Moritz Grosse-Wentrup

Eingereicht von:

Katharina Ring

München, 29. August 2018

Causal Inference from fMRI Data

Bachelor Thesis

at the Faculty for Mathematics, Informatics and Statistics

Ludwig-Maximilians-Universität München

Department of Statistics

Work Group: Data Science

Prof. Dr. Moritz Grosse-Wentrup

Handed in by:

Katharina Ring

Munich, August 29, 2018

Contents

List of Figures

List of Tables

List of Abbreviations

fMRI functional magnetic resonance imaging

EEG electroencephalography

BOLD blood-oxygenation-level-dependent

FDR false discovery rate

1 Introduction

Understanding neural activity in the human brain and how it leads to what we experience as cognition can be seen as one of the central goals in neuroimaging [42, 45]. This task is increasingly approached by performing causal inference on brain imaging data [41]. Recently, Weichwald et al. [45] proposed a compelling approach in this area, where encoding and decoding models of neuroimaging data are used together to reveal more on the causal structure of the data than could previously be extracted by analyzing these models separately. However, their method has only been tested on EEG data [1, 45] and there are specific challenges associated with fMRI data [41] which render it unclear whether the approach is applicable to fMRI data. But as fMRI is the leading neuroimaging technique used in this field [22, 36], being able to utilize the approach on these data is very valuable. Therefore, the applicability of the approach of Weichwald et al. is examined in this thesis.

The field of cognitive neuroscience has seen two developments in the recent past: Firstly, brain imaging has become more powerful with the spatiotemporal resolution of images increasing rapidly and with it the size and complexity of the data [7]. Secondly, the statistical methods available for analyzing these data have shifted towards more powerful multivariate approaches in predictive modeling [14, 16].

Brain Imaging – fMRI

Brain imaging or neuroimaging refers to graphically scanning the brain, measuring – in the case of functional imaging – the neural activity in a (human) brain [36]. The two major methods utilized in functional neuroimaging research are electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) [7]. Both have their advantages and shortcomings: While fMRI has a high spatial resolution, it does not perform well regarding the temporal resolution [36]. Additionally, fMRI only measures the neural activity indirectly through the haemodynamic response [23]. EEG on the other hand performs well regarding temporal precision but exhibits an inferior spatial mapping [36]. Since its development in the early 1990s [7], fMRI technology has improved steadily [35] and is now the leading neuroimaging tool for mapping neural activity [22, 36].

As already mentioned, fMRI does not measure the neural activity directly but uses the haemodynamic response, specifically the blood-oxygenation-level-dependent (BOLD) signal, which measures the concentration of oxygen within blood vessels in the brain [15]. It thereby exploits the functionality of the brain energy metabolism which transports oxygen and glucose to the areas in the brain with an increased need for it, i. e., the areas with high neural activity [22].

Predictive Modeling – Encoding/Decoding

A newer approach to predictive modeling in neuroimaging research is implementing computationally expensive machine learning techniques, which are receiving increasingly greater attention in the context of fMRI applications [39]. Additionally, there has been a development towards

voxel-based modeling [8, 35]. Here, the information of the neuroimage is used in the form of voxels, the 3D equivalent of pixels, of a certain size without using any additional brain region information in the statistical modeling [15].

An example of machine learning models in neuroimaging are encoding and decoding models [16]. Both models require a brain state vector X representing the neural activity per voxel/brain region and a label y, which is either a stimulus or response [45]. Encoding models predict brain states from the experimental condition or response, in probabilistic terms: $p(X|y)$ [35]. In the last decade, more researchers have used this type of modeling, for example: [10, 11, 19, 27, 30]. Decoding models work the other way round. They predict the stimuli or responses from the brain state, i. e.: $p(y|X)$ [17]. When y represents a mental state decoding models are often referred to as 'brain reading' [15]. Decoding models have also accrued interest over the past decade, see [9, 15, 24, 26, 39].

Encoding and decoding models shed light on different questions: While determining the origin of cognitive functions (e. g., relevant to a brain surgeon) can be done using an encoding model, extracting neural processes belonging to a mental state (e. g., important for brain-computer interfaces for communicating with comatose patients) is a typical decoding case [14]. Therefore, the two types of models are complementary [18]. However, decoding models lack the possibility of causal inference which is why a transformation of a decoding in an encoding model has been proposed in the linear case [14]. Additionally, it has been proposed to use a decoding model as proof of concept to an encoding model [16].

Weichwald et al. [45] also developed a method for using encoding and decoding models together in order to gain a deeper insight into causal relationships using causal Bayesian networks, one of the typical methods to study the causal structure in neuroimaging [13, 33, 40, 41], following Pearl [37] and Sprites et al. [43]. For this approach, Weichwald et al. not only differentiate between decoding and encoding but also look at whether the label is a stimulus or a response and incorporate the resulting causal direction of the model. An encoding model $p(X|s)$ if the label is a stimulus s is causal as well as a decoding model $p(r|X)$ with a response r. The other models, a decoding model with a stimulus $p(s|X)$ and an encoding model with a response $p(X|r)$, are anti-causal as they predict from the future to the past.

In their research, Weichwald et al. use causal Bayesian networks as mentioned above. While the concepts utilized for their conclusions are explained in detail in their publication a short summary of important aspects is given here. When the causal relationship is represented in the form of directed acyclic graphs, the concept of *d-separation* corresponds to conditional independence, assuming *faithfulness* and the *causal Markov condition*. Two nodes are d-separated by a set of other nodes A, if and only if all paths between the two are *blocked* by the set. A path is blocked if and only if a path over one element x is either (i) tail-to-tail ($\leftarrow x \rightarrow$) and $x \in A$, (ii) head-to-tail ($\rightarrow x \rightarrow$) and $x \in A$ or (iii) head-to-head ($\rightarrow x \leftarrow$) and neither x nor any descendant $\in A$. Faithfulness states that all independences are implied by the graphs causal

structure. The causal Markov condition states that each node given its direct causes becomes independent of its non-descendants. [45]

Weichwald et al. [45] use univariate statistical tests on independence in order to find relevant and irrelevant features. For the encoding model a variety of tests are available, e. g., the general linear model or a non-linear independence test. Features are called relevant if their corresponding null-hypothesis of independence was rejected and irrelevant otherwise. For decoding models, tests on independence are not so diverse. Typically, the feature of interest is permuted in the dataset and then it is observed whether the accuracy of the decoding model decreases significantly. Again, the features for which the null hypothesis of independence is rejected are called relevant with regard to decoding and the other features irrelevant.

Weichwald et al. then derive the following causal interpretation for relevant and irrelevant features seen in Table 1.

	encoding	decoding	causal interpretation	rule
stimulus-based	✓	n/a	X_i effect of s	S1
	×	n/a	X_i no effect of s	S2
	n/a	✓	*inconclusive*	S3
	n/a	×	*inconclusive*	S4
	✓	✓	X_i effect of s	S5
	✓	×	X_i indirect effect of s	S6
	×	✓	provides brain state context	S7
	×	×	neither effect nor provides brain state context	S8
response-based	✓	n/a	*inconclusive*	R1
	×	n/a	X_i no cause of r	R2
	n/a	✓	*inconclusive*	R3
	n/a	×	*inconclusive*	R4
	✓	✓	*inconclusive*	R5
	✓	×	X_i no direct cause of r	R6
	×	✓	provides brain state context	R7
	×	×	neither effect nor provides brain state context	R8

Table 1: The causal interpretation rules for a feature X_i which is either relevant (✓) or irrelevant (×) in encoding and decoding adapted from [45]

The table shows that the decoding model alone does not allow any causal interpretation, but in combination with an encoding model it can provide additional information on the causal relationship of the features.

While it has been stated, that encoding and decoding models should be used together when analyzing fMRI data [35] and many researchers have done so [4, 19, 30, 34, 44], the particular

approach of Weichwald et al. [45] has only been used on EEG data [1, 45]. Testing their findings on fMRI data is interesting because it is not clear whether this approach also works in this context. One major challenge for this approach stems from the indirect nature of measurement that was already mentioned [15]. Other problems include that the strength of the BOLD signal may vary among participants and that the time passing from the neural activity to the effect in the haemodynamic response can differ among different brain regions even for one subject [41]. These and other issues will be addressed in more detail in section 2.2.

The contribution of this work is to get first empirical results on whether the approach of Weichwald et al. [45] works in the context of fMRI. This is achieved by looking only at the rules S5 and S6 from Table 1. Due to the indirect measurement by fMRI no features should show relevant in encoding but irrelevant in decoding (more on that in section 2), rendering the proposed combined look at encoding and decoding models unfeasible for fMRI data. However, since the results show the combination of a feature being relevant in encoding but irrelevant in decoding, this work provides a first indication that at least in some cases the approach of Weichwald et al. [45] can be used with fMRI data.

In the next section the methods of this work are explained including a close look at the dataset used, the specific problems faced in causal inference on fMRI data, and the statistical procedure with encoding and decoding models. The section is followed by the results. The thesis closes by discussing the results and highlighting the limitations of its findings.

2 Methods

2.1 Dataset

The data used in this thesis were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The part of the Human Connectome Project data used encompasses the preprocessed task fMRI datasets on motor functions from the 1200 subjects data release. [28]

The protocol for this task was adapted from Buckner et al. [5]. Participants of the study were provided visually with written cues which asked them to move a body part, specifically right hand, left hand, right foot, left foot, and tongue. Each movement was conducted twice at twelve seconds each after the cue was shown for three seconds. The protocol also included three fixation blocks of fifteen seconds. All of these tasks were conducted in the same order for every individual, i. e., were not randomized, and were preceded by an eight seconds countdown. [2]

In the motor functionality task, 1086 individuals participated. The study was comprised of

healthy young adults typically between the age of 22 and 35. Many of the participants were related to other participants and twin siblings were especially selected for the study. [29]

All participants of the study were scanned on a customized Siemens 3T "Connectome Skyra" using a standard 32-channel. For more information see [28]. During the task protocols one image with 91x109x91 voxels was taken per 720 ms [2]. The data preprocessing pipeline included minimization of spatial distortions as well as alignment across modalities and subjects [28].

2.2 Challenges of causal inference from fMRI

As mentioned before, fMRI data bear certain challenges in combination with causal inference, so in order to apply the method of Weichwald et al. [45] to fMRI data, these challenges need to be addressed. Ramsey et al. [41] have discussed six of the most pressing problems that occur in this context. These will be described in the following paragraphs along with their consequences for the analyses in this work.

Number of possible causal structures The number of alternative causal models to investigate is astronomical as there are a multitude of brain regions and all possible relationships between them have to be considered. While this problem is not specific to fMRI data, it represents a critical hurdle to causal inference research on the brain. However, the procedure proposed by Weichwald et al. [45] provides only the first step of reaching a causal model, namely, identifying what the data in question can reveal about the causal structure between pairs of two brain regions using encoding and decoding models. Therefore, the next step, deriving a complete causal model, does not lie within its scope.

Indirect measurement As discussed previously, fMRI does not measure the neural activity directly, but can be thought of as a non-linear function of it. This causes problems for the typical causal inference in graphical models as seen in the example of a causal chain in Figure 1. The disturbance terms ε_i represent other influences that are not of interest in this situation and are independent of each other as well as of positive variance. Since the connection between the neural activity X_i and the BOLD measurements M_i is non-linear, the measurements show no causal relationship between each other, i. e., $M_1 \not\perp\!\!\!\perp M_3 | M_2$, as M_2 does not block the path from M_1 to M_3, in this example even though the underlying neural activity is causally linked, i. e., $X_1 \perp\!\!\!\perp X_3 | X_2$, since X_2 does block the path between X_1 and X_3. Deconvolution methods represent a possible solution for this issue as they attempt to estimate the latent neural activity from the BOLD signal [12]. But the results of deconvolution are very

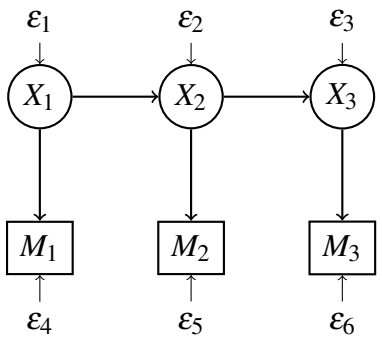

Figure 1: Indirect measurement of a causal chain adapted from [41]

noisy, so this approach is to be treated with caution and was not implemented for this thesis.

Time series The BOLD signal does not provide only one measurement per neural activity, but shows a development over time, as the haemodynamic response slowly increases and then decreases again. Therefore, a decision has to be made as to how to incorporate the BOLD signal into the model. For this work, several fMRI images will be averaged over time, leading to one value per task, as suggested by Grosse-Wentrup et al. [13].

Varying BOLD signal delay The delay with which the haemodynamic response sets in can vary among brain tissues. Some parts of the brain may respond later than others to a stimulus resulting in time differences among the variables. A solution would be to neutralize this effect by backshifting delayed responses. However, Ramsey et al. did test for this problem in their work, but found out that it had no relevant impact which suggests that the negative consequences of this problem may be negligible. Additionally, backshifting is not necessary for this thesis since BOLD measurements taken over a period of time are averaged in this work as mentioned above. Therefore, this problem, if it occurs here, should disappear or at least be reduced.

Strength of effect across participants Individuals in a study are bound to differ with regard to the strength of the BOLD signal measured. Even though the general processing of a task may be similar among participants of a study, the strength of the response to a stimulus or another neural activity will most likely differ among individuals. For this reason, random effects models are often utilized in order to take the individual differences into account [20, 32]. Alternatively, a modification for multiple datasets as conducted by Grosse-Wentrup et al. [13] could be used. However, since the Human Connectome data only provide two trials per task and person, this is not feasible.

Different variable sets Different participants of the study may cope with a given task differently resulting in diverse "variable of interest" sets. However, this can also happen if the participants approach the task in the same way due to anatomical differences of the haemodynamic response as discussed in the paragraph above. The different variable sets do then have to be integrated for statistical processes, for example, by taking only variables which are relevant for every individual, even though this can lead to misrepresented causal relationships. This problem, again, can not be addressed here because the data only provide one repetition of a task per person. For this work, the data will be treated as if they had been taken from a single participant, which constitutes an important limitation of this study.

Other problems may influence the applicability of Weichwald et al. on fMRI data, for example, measurement errors or issues related to the acquisition and preprocessing of the data [41].

However, these issues cannot be addressed in the analysis of the data, but have to be addressed beforehand during the data acquisition stage. Therefore, they fall beyond the scope of this thesis.

2.3 Statistical procedure

As mentioned before, in this section, rules S5 and S6 from Table 1 are utilized, meaning relevant features for encoding are sought and later on tested for their relevance in decoding. In order to determine the relevance and irrelevance of feature in encoding and decoding, correlations and partial correlations are used respectively, due to their simplicity. However, the linearity of these models represent a strong limitation for this study. The procedure for finding relevant and irrelevant features draws heavily from Grosse-Wentrup et al. [13], although they do not use partial correlations for decoding. This work focuses on causal chains in a stimulus-based setting as depicted in Figure 2, where X_i represent the neural activity, M_i are the BOLD measurements of the neural activity, and ε_i represent disturbance terms as in Figure 1 and s is the stimulus.

In this example, X_1 and X_2 are both effects of s and are therefore dependent on s, i. e., $X_1 \not\perp\!\!\!\perp s$ and $X_2 \not\perp\!\!\!\perp s$ (corresponding to rule S5). Since X_1 is a direct effect, there is no conditional independence ($X_1 \not\perp\!\!\!\perp s|X_2$), but for the indirect effect X_2 there is ($X_2 \perp\!\!\!\perp s|X_1$), as X_1 blocks the path between s and X_2 (corresponding to rule S5). The first part (S5) is also true when considering the measurements M_i instead of the neural activity X_i: both effects are dependent of the stimulus, $M_1 \not\perp\!\!\!\perp s$ and $M_2 \not\perp\!\!\!\perp s$. As for the second part (S6), the direct effect still shows no conditional independence ($M_1 \not\perp\!\!\!\perp s|M_2$), but neither does the indirect effect ($M_2 \not\perp\!\!\!\perp s|M_1$) because M_1 does not block the path between s and M_2. [45]

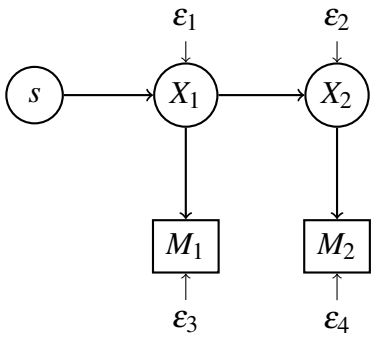

Figure 2: Indirect measurement of a causal chain with a stimulus

This is caused by the indirect nature of the measurement addressed before. However, it is not clear whether this also represents an empirical problem, as – depending on the nature of the measurement – this effect could be negligible, meaning that while there is no conditional independence, the data are close to independence and therefore specific encoding and decoding models will still be able to detect indirect effects. Testing the applicability of Weichwald et al. [45] empirically is the subject of this thesis. The precise procedure used for this work is explained in this section.

First, in order to be able to extract features relevant in encoding and decoding, the number of stimuli of the dataset was reduced from five – right hand, left hand, right foot, left foot, and tongue – to two – only left hand and right hand – in order to receive a binary stimulus that can be numerically represented as a single dummy encoded vector. As the primary brain region responsible for a motor task regarding the left hand lies far on the right side and vice versa, these

regions would be far apart from each other and therefore easily differentiable [38]. Additionally, these brain regions should be very similar apart from the fact that they lie on opposite sides of the brain [38] and this means results for one side can be verified using results from the other. This will be utilized when deciding on brain regions in the next section.

In the following, the exact steps for deriving brain regions relevant in encoding and later on testing these brain regions for relevance in decoding are presented.

2.3.1 Encoding

In order to identify brain regions relevant for encoding, significant correlations of voxels with the stimulus are extracted from the data. Clusters of significantly correlated voxels are then merged into a brain region that is deemed relevant in encoding.

However, in a first step the data need to be adjusted. The raw data contain several fMRI pictures per task for the participants. These are averaged in order to have one measurement per task. As mentioned before, BOLD measurements lag behind the neural activity and this lag can differ among brain regions [41]. Thus, the time window used for averaging the neural activity per voxel is determined empirically using the data. For this, a subset of over 1000 randomly chosen voxels was averaged over 100 randomly chosen subjects and plotted. Among these plots, voxels that are shown responsive to the respective stimulus are used in order to determine a time window. Two voxels showing a clear response were identified for each stimulus condition. The identified voxels all lay in the primary motor area for right and left hand [38] which additionally confirms this procedure. One voxel for each side is depicted in Figures 3 and 4. The figures for the other two voxels, which were used as a confirmation for the choice of the time window, can be found in Appendix A.

The windows chosen were of equal size and delay after the movement starts for both sides. While brain regions differ with regard to the delay of the BOLD signal, regions as similar as

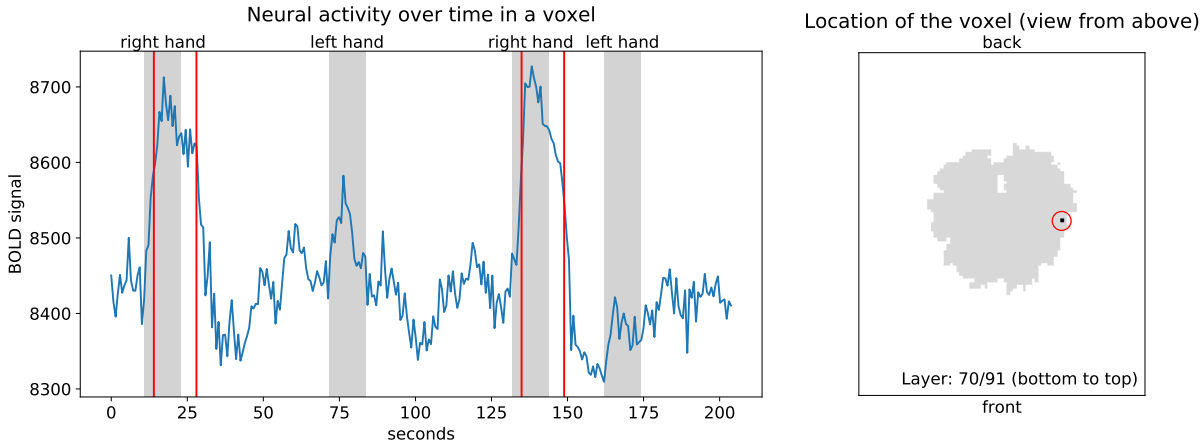

Figure 3: The time window (red) for the movement of the right hand as shown by a voxel responsive to that task

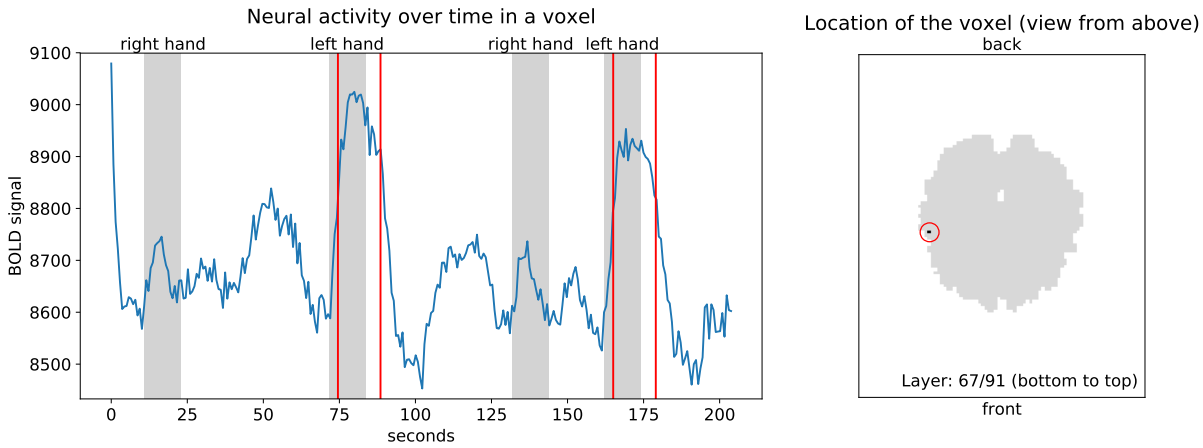

Figure 4: The time window (red) for the movement of the left hand as shown by a voxel responsive to that task

"right hand movement" and "left hand movement" should have a similar delay [41]. This was also supported by the data as seen below. The windows begin three seconds after the movement starts (so six seconds after the cue) and are fourteen seconds long, i. e., two seconds longer than the pure movement, but one second shorter than the task including the cue.

Within these windows, the measurements of the neural activity were averaged leading to one value per voxel and trial of a task conducted. Using trial-averaged BOLD activity was suggested by Grosse-Wentrup et al. [13].

In the next step, the correlations of the dummy encoded stimulus and these averaged measurements were calculated and a p-value derived through a permutation test as done by Grosse-Wentrup et al. [13]: For each feature, the correlation ρ_{s,X_j} between the feature X_j and the stimulus is calculated. Then, a corresponding p-value can be estimated by permuting the stimulus s 10^4 times and calculating the resulting correlation ρ^{H0}_{s,X_j} with X_j, giving an estimated distribution under the null-hypothesis of independence. The frequency with which ρ^{H0}_{s,X_j} exceeds ρ_{s,X_j} divided by the number of permutations gives the estimated p-value. A permutation test was chosen instead of, for example, a t-test, as it requires no assumptions regarding the distribution of the two random variables involved [46].

However, the permutation test as conducted by Grosse-Wentrup et al. [13] had to be adjusted for a specific issue in the data caused by treating the many subjects as if they were just one. Figure 5 shows the neural activity averaged over 1000 voxels for the first 25 subjects (each gray line indicates the first measurement for a new subject). As the figure

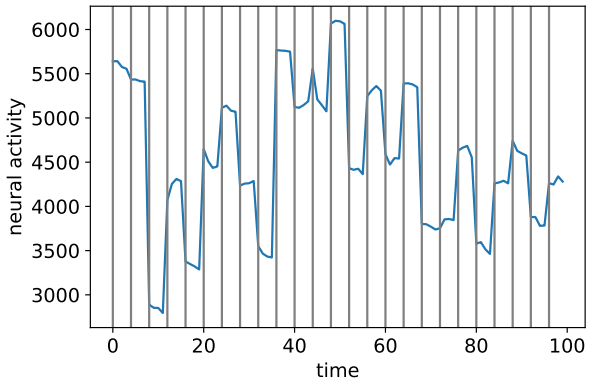

Figure 5: Neural activity averaged over 1000 voxels for the first 25 subjects

shows, the inter-subject differences in neural activity are much bigger than the intra-subject differences. For the stimulus, the tasks are balanced across the subjects, meaning every subject conducts two right hand and two left hand movement tasks. If this balancing is dropped in the permutations, the possible correlations are more extreme. Consequently, the p-values resulting from an unbalanced permutation test would be pushed towards 0.5, which is not desirable. Hence, in order to receive useful p-values, the permutations have to be balanced across the subjects as well, which was done for this thesis.

After the p-values were derived for every subject, their significance was decided on using the false discovery rate (FDR) [3]. The FDR is useful in multiple testing problems where several tests support the same or a similar hypothesis. It determines the percentage of falsely rejected null hypotheses. In this case, forming brain regions out of significant voxels is of interest. For this, it is not necessary that every significant voxel is truly correlated as the search for a brain region is based on clusters of significant voxels. The decision for the threshold of the FDR was made by visually inspecting the results for different values based on a trade-off between keeping the noise of voxels not belonging to brain regions low and still finding as many regions as possible. Here, the FDR was set to 0.3 %, corresponding to a p-value of 0.0001. The resulting significant voxels are depicted in Figure 6.

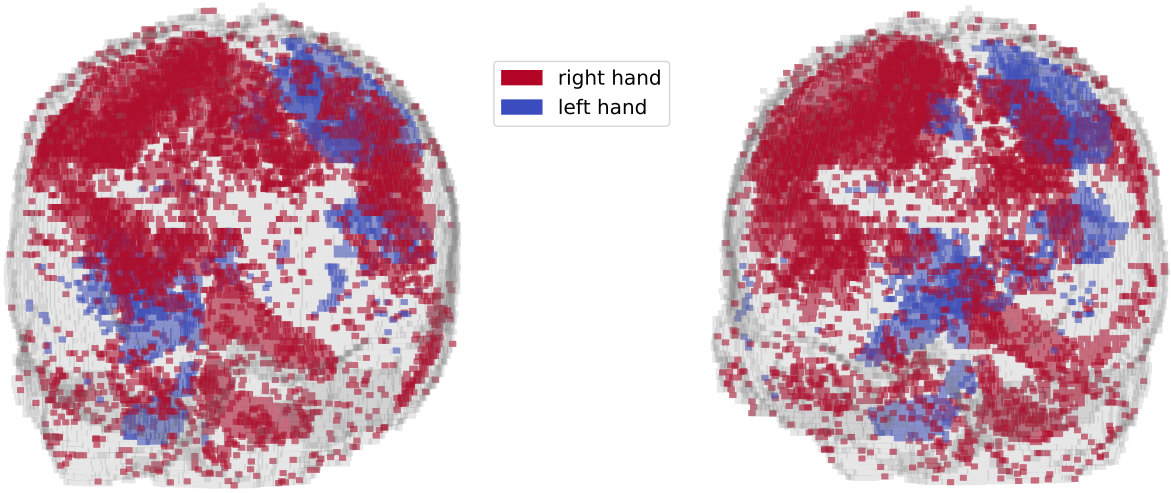

Figure 6: 3D frontal view on the significant voxels (left side: view shifted 15 degrees to the left, right side: view shifted 15 degrees to the right)

Afterwards the brain regions were determined by visually identifying clusters of voxels. First, it was decided to only include voxels correlated with the left hand movement. Both hand movements should show a similar pattern, therefore only one movement needs to be investigated. Additionally, the p-values for the right hand movement show a lot of noise which might stem from a correlation with time as the right hand task always came before the left hand task. The left hand movement shows less noise and should therefore allow for better results.

In order to exclude other time effects and artifacts as well as possible, only clusters which show a similar pattern for the right hand movement on the opposite side of the brain were

included. When two clusters were only slightly connected, the clusters were included as different regions in order to increase the number of brain regions found. Voxels were then sorted into either one of the two regions; no voxel belongs to two regions. The determined regions only consist of the significant voxels and were not smoothed over, allowing, e. g., voxel-sized holes in the regions.

Using these rules, eight brain regions which are relevant for encoding were discovered as depicted in Figure 7.

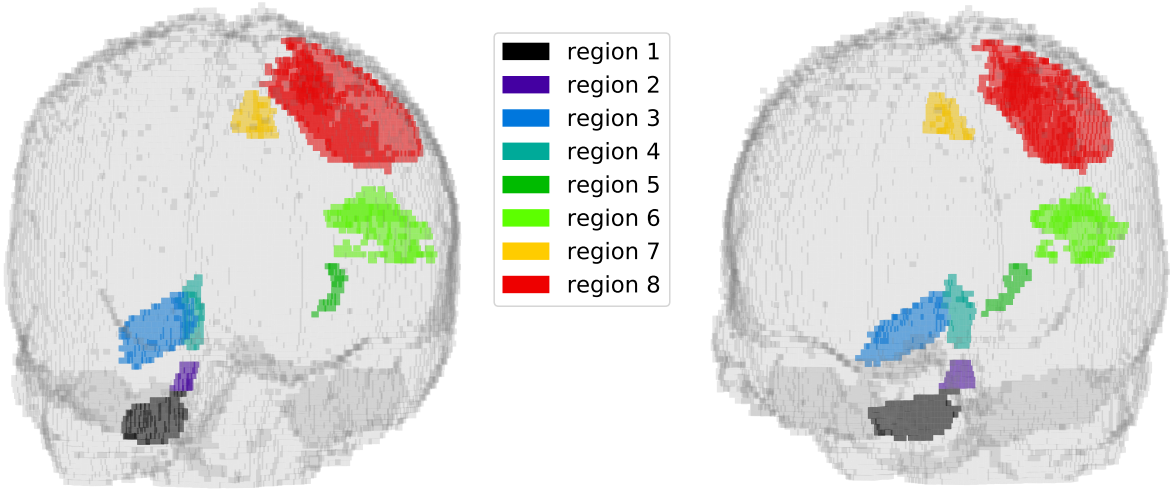

Figure 7: 3D frontal view on the brain regions relevant in encoding (left side: view shifted 15 degrees to the left, right side: view shifted 15 degrees to the right)

2.3.2 Decoding

In order to examine which of the brain regions that are relevant in encoding are relevant and irrelevant in decoding, the neural activity within the brain regions was averaged. While correlations were used for encoding, partial correlations were used for identifying features relevant and irrelevant in decoding. Partial correlations describe the correlation of two vectors given a set of controlling variables and can be computed as the correlations of residuals from linear regressions on the vectors in question using the controlling variables as features [6]. As this computation is based on normal correlations again, a permutation test on the correlation between the residuals can be conducted as done for the encoding model in order to get p-values [21].

Expanding on Weichwald et al. [45], for the set of controlling variables, i. e., the features in the linear regressions, all subsets of brain regions (excluding the respective brain region examined) were tested in order to find causal chains in situations as depicted in Figure 8. In this example, X_2 is an indirect effect, but $X_2 \not\perp\!\!\!\perp s | X_1, X_3$, because while X_1

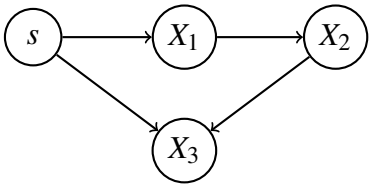

Figure 8: Example of a causal chain with a stimulus

blocks the path, conditioning on X_3 "opens" another path.
However, $X_2 \perp\!\!\!\perp s | X_1$, since only X_1 is conditioned on, which blocks the path. By testing all subsets of X_1, X_2, and X_3, X_2 can be identified as an indirect effect via X_1. [45]

In a last step, the p-values of the partial correlations calculated from a permutation test give some information on where the residuals are independent. This step is to be treated with caution as independence is the null hypothesis and the p-values are used to decide whether the null hypothesis is true. Due to the interest in the null hypothesis, the FDR threshold was set to be rather high with 50 %. This threshold is rather arbitrary, but a threshold of 10 % identified the same regions as relevant and irrelevant in decoding, as did a threshold of 90 %.

3 Results

Using a false discovery rate of 50 %, three regions were identified as irrelevant in decoding: region 2, region 4, and region 7. All other regions showed no partial correlation with a p-value above the threshold. The fact that regions irrelevant in decoding were found provides a first indication that despite of the caveats stated in section 2.2, the approach of Weichwald et al. [45] could still be useful for fMRI research. According to rule S6, the features irrelevant in decoding are indirect effects of the stimulus. Within this section, the partial correlations with a p-value of higher than 0.4471, which corresponds to the FDR threshold of 50 %, are examined. The results discussed below are highly consistent, which serves as an additional support for the applicability of the approach in fMRI.

3.1 Region 2

Many partial correlations with regard to region 2 exceeded the FDR threshold, most of them containing region 3 in the subset of variables controlled for. A few of these partial correlations are depicted in Table 2; the full table can be found in Appendix B. Since conditioning on region 3 alone appears to be enough to lead to (linear) independence between the stimulus and region 2, the results show the brain activity in region 2 to be an indirect effect of the stimulus via region 3. As including other regions in the subset of variables that are controlled for in the test sometimes increases the p-value – for example including region 1 increases the p-value from 0.77 to 0.93 – other regions could also have an (almost negligible) effect on region 2 originating from the stimulus. However, these increases could also be attributed to chance.

Null hypothesis (linear test)	P-value
$s \perp\!\!\!\perp reg.2 \mid reg.3$	0.7695
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3$	0.9327
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4$	0.7716
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.5$	0.8155
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.8$	0.8608

Table 2: Significant p-values for region 2 (partial)

3.2 Region 4

For region 4, far fewer partial correlation tests delivered a p-value above the FDR threshold. All of them including their corresponding p-value can be seen in Table 3. The table shows that all of the subsets controlled for include region 3 and controlling for region 3 alone also leads to a p-value exceeding the FDR threshold. Therefore, it can be concluded that region 4 is an indirect effect of the stimulus via region 3 as well. Including other regions does increase the p-value slightly, meaning that other regions could have (negligible) effects on region 4 as well which originate from the stimulus. However, these increases are rather small and could also be attributed to chance.

Null hypothesis (linear test)	P-value
$s \perp\!\!\!\perp reg.4 \mid reg.3$	0.4831
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3$	0.5544
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.5$	0.4936
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.8$	0.7009
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.5, reg.8$	0.6406
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.6, reg.8$	0.5204
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.7, reg.8$	0.6612
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.5, reg.6, reg.8$	0.5413
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.5, reg.7, reg.8$	0.6290
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.6, reg.7, reg.8$	0.4586
$s \perp\!\!\!\perp reg.4 \mid reg.1, reg.3, reg.5, reg.6, reg.7, reg.8$	0.5135

Table 3: Significant p-values for region 4

3.3 Region 7

Region 7 is the last region which was determined to be an indirect effect. The partial correlation test including the subset of region 5 and region 8 achieved a very high p-value of 0.99. Including more regions into the subset only decreases that value as shown in Table 4. Therefore it can be concluded from the data that region 7 is an indirect effect with a path via regions 5 and 8. Since including other brain regions in the set of controlling variables only decreases the p-value, no other brain regions appear to be involved.

Null hypothesis (linear test)	P-value
$s \perp\!\!\!\perp reg.7 \mid reg.5, reg.8$	0.9901
$s \perp\!\!\!\perp reg.7 \mid reg.1, reg.5, reg.8$	0.7562
$s \perp\!\!\!\perp reg.7 \mid reg.2, reg.5, reg.8$	0.7643
$s \perp\!\!\!\perp reg.7 \mid reg.4, reg.5, reg.8$	0.4711
$s \perp\!\!\!\perp reg.7 \mid reg.5, reg.6, reg.8$	0.8036
$s \perp\!\!\!\perp reg.7 \mid reg.1, reg.2, reg.5, reg.8$	0.6288

Table 4: Significant p-values for region 7

4 Discussion

Within this work, the approach of Weichwald et al. [45] – using encoding and decoding models together in neuroscience research in order to gain more insight into the causal structure of the data compared to evaluating both models separately – which had previously only been used on EEG data, was empirically tested on fMRI data. Due to several characteristics of fMRI data, notably the fact that instead of the neural activity the haemodynamic response is measured, it was unclear whether this approach could also be applied in the context of fMRI. It was especially doubtful whether brain regions that are relevant features in encoding models could show irrelevant in decoding models due to the indirect nature of measurement. However, this thesis shows that brain regions relevant in encoding and irrelevant in decoding can exist for fMRI data as well. Therefore, the results of this work are a first indication that the approach of Weichwald et al. [45] is also applicable in combination with fMRI data. Since neuroscientists work frequently with fMRI data, the findings of this work could help advance the fields of fMRI and neuroscience.

However, this work exhibits several limitations. Firstly, it is not clear whether all indirect effects could be identified as more indirect effects could exist. Further research should therefore focus on the overarching applicability of the approach Weichwald et al. [45] on fMRI data.

Within this work, it is also not clear how the limitations with regard to the dataset are connected to the results. Due to the limited scope of this thesis, only secondary data sources were available and the data used did not meet all desired criteria. First, the trials conducted by the Human Connectome Project were not randomized, so the results are interpreted under the assumption that time effects do not influence them. Secondly, the data originate from many different participants of the study, but are treated as if they originate from one person. It is unclear whether and how the results would improve if the data really originated from one person.

Additionally, Ramsey et al. [41] suggest a few adjustments, e. g., using random effects models, when conducting causal inference on fMRI data. Due to the scope of this work, they were not implemented. Therefore, their impact on the approach of Weichwald et al. [45] remains to be tested.

Two more important limitations stem from the testing procedures implemented. Firstly, only linear tests were conducted even though the characteristic of interest was (in-)dependence which is not necessarily linear. This means that non-linear dependences may not be discovered by the tests [13]. Empirically, however, most dependences between features in neuroscience seem to be linear [31, 35]. Secondly, features were marked as irrelevant in encoding based on a failure to reject the null hypothesis of no correlation. This is a flaw in the statistical procedure that could not be avoided as a test with the null hypothesis correlation is not possible [47]

Furthermore, all limitations of the approach proposed by Weichwald et al. [45] itself still apply in combination with fMRI data. For this work it is especially relevant that the approach cannot differentiate between a causal chain ($s \to X_1 \to X_2$) and a tail-to-tail relationship ($s \to X_1 \leftarrow X_2$).

Other limitations of this approach are the untestable assumptions of faithfulness and the causal Markov condition. However, the fact that the set of all unfaithful distributions has a measure of zero compared to the set of all distributions that can be generated by a given directed acyclic graph [25]. Therefore, the assumption of faithfulness might not be a strong limitation of the approach of Weichwald et al. [45].

These limitations should be addressed in further research in order to gain certainty of the applicability of the approach of interpreting encoding and decoding models together proposed by Weichwald et al. [45] on fMRI data. However, this work is a promising first step in confirming the applicability on fMRI data as indirect effects could be identified using this type of neuroimaging.

5 References

[1] Bach, Dominik R. et al. “Whole-brain neural dynamics of probabilistic reward prediction”. In: *Journal of neuroscience* 37.14 (2017), pp. 3789–3798.

[2] Barch, Deanna M. et al. “Function in the human connectome: task-fMRI and individual differences in behavior”. In: *Neuroimage* 80 (2013), pp. 169–189.

[3] Benjamini, Yoav and Hochberg, Yosef. “Controlling the false discovery rate: a practical and powerful approach to multiple testing”. In: *Journal of the royal statistical society. Series B (Methodological)* (1995), pp. 289–300.

[4] Brouwer, Gijs Joost and Heeger, David J. “Decoding and reconstructing color from responses in human visual cortex”. In: *Journal of neuroscience* 29.44 (2009), pp. 13992–14003.

[5] Buckner, Randy L. et al. “The organization of the human cerebellum estimated by intrinsic functional connectivity”. In: *Journal of neurophysiology* 106.5 (2011), pp. 2322–2345.

[6] Cohen, Patricia, West, Stephen G., and Aiken, Leona S. *Applied multiple regression/correlation analysis for the behavioral sciences*. Psychology press, 2014.

[7] Cohen, Ronald A. and Sweet, Lawrence H. *Brain imaging in behavioral medicine and clinical neuroscience*. Springer science & business media, 2010.

[8] Davis, Tyler et al. “What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis”. In: *Neuroimage* 97 (2014), pp. 271–283.

[9] De Martino, Federico et al. “Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns”. In: *Neuroimage* 43.1 (2008), pp. 44–58.

[10] Di Liberto, Giovanni M., O’Sullivan, James A., and Lalor, Edmund C. “Low-frequency cortical entrainment to speech reflects phoneme-level processing”. In: *Current biology* 25.19 (2015), pp. 2457–2465.

[11] Dumoulin, Serge O. and Wandell, Brian A. “Population receptive field estimates in human visual cortex”. In: *Neuroimage* 39.2 (2008), pp. 647–660.

[12] Gitelman, Darren R. et al. “Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution”. In: *Neuroimage* 19.1 (2003), pp. 200–207.

[13] Grosse-Wentrup, Moritz et al. “Identification of causal relations in neuroimaging data with latent confounders: An instrumental variable approach”. In: *NeuroImage* 125 (2016), pp. 825–833.

[14] Haufe, Stefan et al. “On the interpretation of weight vectors of linear models in multivariate neuroimaging”. In: *Neuroimage* 87 (2014), pp. 96–110.

[15] Haynes, John-Dylan and Rees, Geraint. “Neuroimaging: decoding mental states from brain activity in humans”. In: *Nature reviews neuroscience* 7.7 (2006), pp. 523–534.

[16] Holdgraf, Christopher R. et al. “Encoding and decoding models in cognitive electrophysiology”. In: *Frontiers in systems neuroscience* 11.61 (2017), pp. 1–24.

[17] Huth, Alexander G. et al. “Decoding the semantic content of natural movies from human brain activity”. In: *Frontiers in systems neuroscience* 10.81 (2016), pp. 1–16.

[18] Kay, Kendrick N. and Gallant, Jack L. “I can see what you see”. In: *Nature neuroscience* 12.3 (2009), pp. 245–246.

[19] Kay, Kendrick N. et al. “Identifying natural images from human brain activity”. In: *Nature* 452.7185 (2008), pp. 352–355.

[20] Lazar, Nicole A. et al. “Combining brains: a survey of methods for statistical pooling of information”. In: *Neuroimage* 16.2 (2002), pp. 538–550.

[21] Legendre, Pierre. “Comparison of permutation methods for the partial correlation and partial Mantel tests”. In: *Journal of statistical computation and simulation* 67.1 (2000), pp. 37–73.

[22] Logothetis, Nikos K. and Pfeuffer, Josef. “On the nature of the BOLD fMRI contrast mechanism”. In: *Magnetic resonance imaging* 22.10 (2004), pp. 1517–1531.

[23] Logothetis, Nikos K. et al. “Neurophysiological investigation of the basis of the fMRI signal”. In: *Nature* 412.6843 (2001), pp. 150–157.

[24] Martin, Stéphanie et al. “Decoding spectrotemporal features of overt and covert speech from the human cortex”. In: *Frontiers in neuroengineering* 7.14 (2014), pp. 1–15.

[25] Meek, Christopher. “Strong completeness and faithfulness in Bayesian networks”. In: *Proceedings of the eleventh conference on uncertainty in artificial intelligence*. Morgan kaufmann publishers Inc. 1995, pp. 411–418.

[26] Mesgarani, Nima and Chang, Edward F. “Selective cortical representation of attended speaker in multi-talker speech perception”. In: *Nature* 485.7397 (2012), pp. 233–236.

[27] Mesgarani, Nima et al. “Phonetic feature encoding in human superior temporal gyrus”. In: *Science* 343.6174 (2014), pp. 1006–1010.

[28] WU-Minn, Human Connectome Project. *1200 Subjects Data Release Reference Manual*. 2017.

[29] WU-Minn, Human Connectome Project. *ConnectomeDB*. 2018.

[30] Mitchell, Tom M. et al. “Predicting human brain activity associated with the meanings of nouns”. In: *Science* 320.5880 (2008), pp. 1191–1195.

[31] Müller, K R., Anderson, Charles W., and Birch, Gary E. “Linear and nonlinear methods for brain-computer interfaces”. In: *IEEE transactions on neural systems and rehabilitation engineering* 11.2 (2003), pp. 165–169.

[32] Mumford, Jeanette A. and Poldrack, Russell A. “Modeling group fMRI data”. In: *Social cognitive and affective neuroscience* 2.3 (2007), pp. 251–257.

[33] Mumford, Jeanette A. and Ramsey, Joseph D. “Bayesian networks for fMRI: a primer”. In: *Neuroimage* 86 (2014), pp. 573–582.

[34] Naselaris, Thomas et al. “Bayesian reconstruction of natural images from human brain activity”. In: *Neuron* 63.6 (2009), pp. 902–915.

[35] Naselaris, Thomas et al. “Encoding and decoding in fMRI”. In: *Neuroimage* 56.2 (2011), pp. 400–410.

[36] Passingham, Richard E. and Rowe, James B. *A short guide to brain imaging: The neuroscience of human cognition*. Oxford university press, USA, 2016.

[37] Pearl, Judea. “Causality: models, reasoning, and inference”. In: *Econometric theory* 19.675-685 (2003), p. 46.

[38] Penfield, Wilder and Rasmussen, Theodore. *The cerebral cortex of man; a clinical study of localization of function.* Macmillan, 1950.

[39] Pereira, Francisco, Mitchell, Tom, and Botvinick, Matthew. “Machine learning classifiers and fMRI: a tutorial overview”. In: *Neuroimage* 45.1 (2009), S199–S209.

[40] Ramsey, Joseph D., Hanson, Stephen José, and Glymour, Clark. “Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the Smith et al. simulation study”. In: *Neuroimage* 58.3 (2011), pp. 838–848.

[41] Ramsey, Joseph D. et al. “Six problems for causal inference from fMRI”. In: *Neuroimage* 49.2 (2010), pp. 1545–1558.

[42] Smith, Stephen M. et al. “Network modelling methods for FMRI”. In: *Neuroimage* 54.2 (2011), pp. 875–891.

[43] Sprites, Peter, Glymour, Clark N., and Scheines, Richard. *Causation, prediction, and search*. MIT press, 2000.

[44] Thirion, Bertrand et al. “Inverse retinotopy: inferring the visual content of images from brain activation patterns”. In: *Neuroimage* 33.4 (2006), pp. 1104–1116.

[45] Weichwald, Sebastian et al. “Causal interpretation rules for encoding and decoding models in neuroimaging”. In: *Neuroimage* 110 (2015), pp. 48–59.

[46] Wilcox, Rand R. *Introduction to robust estimation and hypothesis testing*. Academic press, 2011.

[47] William, L. THOMPSON. “Null hypothesis testing: problems, prevalence, and an alternative”. In: *Journal wildlife management* 64.4 (2000), pp. 912–923.

Appendix

Appendix A: Voxels confirming the time window for encoding

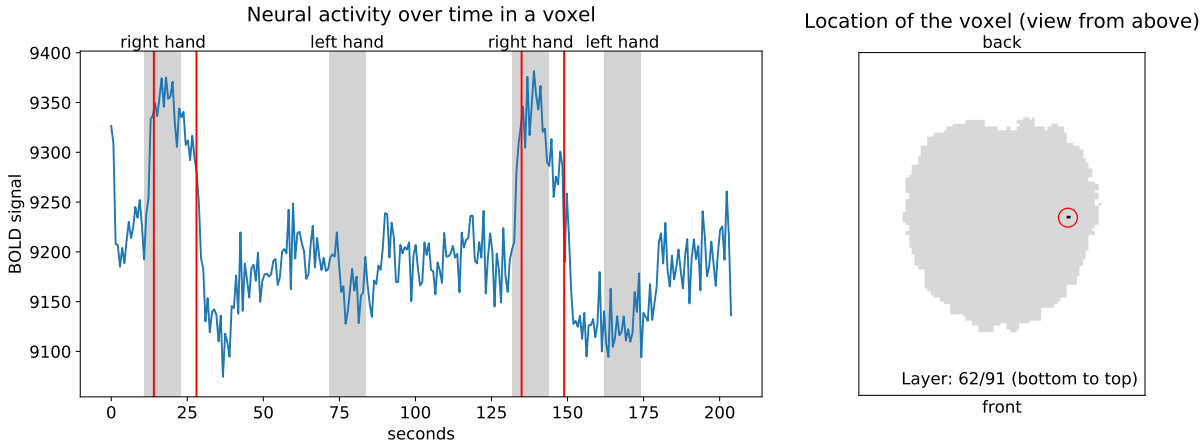

Figure 9: Appendix A: The time window (red) for the movement of the right hand as shown by a second voxel responsive to that task for confirmation

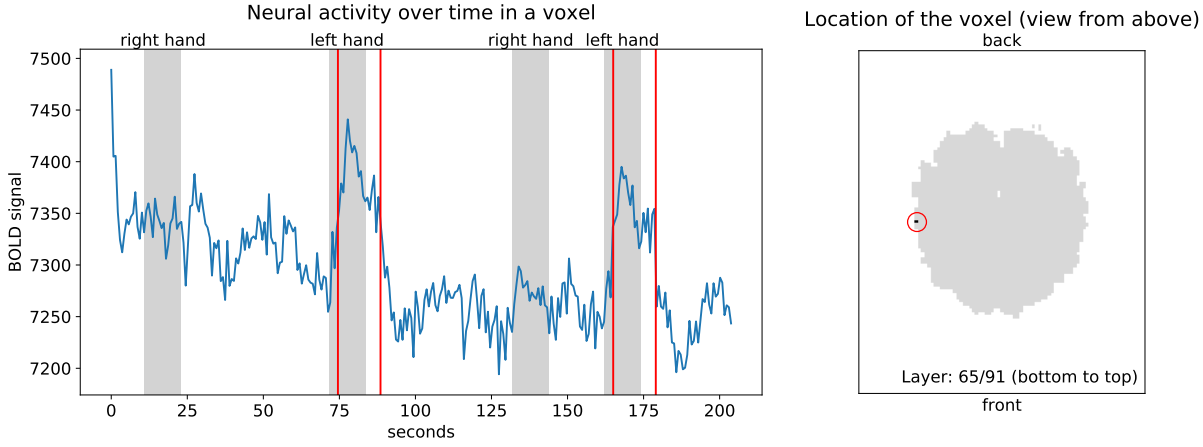

Figure 10: Appendix A: The time window (red) for the movement of the left hand as shown by a second voxel responsive to that task for confirmation

Appendix B: Significant p-values for region 2 (full table)

Null hypothesis (linear test)	P-value
$s \perp\!\!\!\perp reg.2 \mid reg.3$	0.7695
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3$	0.9327
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4$	0.7716
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.5$	0.8155
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.8$	0.8608
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.6$	0.9526
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.6$	0.5885
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.6, reg.8$	0.6553
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.7$	0.9636
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.7, reg.8$	0.8075
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4$	0.9264
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5$	0.9669
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.8$	0.6318
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5$	0.5423
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.8$	0.8815
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5$	0.8346
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.5, reg.8$	0.9025
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.6, reg.7$	0.9719
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.6, reg.7$	0.5344
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.6, reg.7, reg.8$	0.5548
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.6$	0.9685
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.6, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.6$	0.9683
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.6, reg.8$	0.8067
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.6$	0.7067

Table 5: Appendix B: Significant p-values for region 2 (part 1)

Null hypothesis (linear test)	P-value
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.6, reg.8$	0.7573
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5, reg.6$	0.5528
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.5, reg.6, reg.8$	0.813
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.7$	0.9699
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.7$	0.9796
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.7, reg.8$	0.5961
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.7$	0.4684
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.7, reg.8$	0.8409
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5, reg.7$	0.5279
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.5, reg.7, reg.8$	0.8906
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5$	0.9672
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.8$	0.8852
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5, reg.8$	0.9256
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.6, reg.7$	0.9863
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.6, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.6, reg.7$	0.9804
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.6, reg.7, reg.8$	0.7736
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.6, reg.7$	0.6447
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.6, reg.7, reg.8$	0.6849
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.5, reg.6, reg.7, reg.8$	0.7801
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.6, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.6$	0.9796
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.6, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.6, reg.8$	0.9198
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5, reg.6, reg.8$	0.8802
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.7$	0.9842
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.7, reg.8$	0.8817
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5, reg.7, reg.8$	0.9194
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.6, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.6, reg.7$	0.9893
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.5, reg.6, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.4, reg.5, reg.6, reg.7, reg.8$	0.9087
$s \perp\!\!\!\perp reg.2 \mid reg.3, reg.4, reg.5, reg.6, reg.7, reg.8$	0.8572
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.6, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.7, reg.8$	1
$s \perp\!\!\!\perp reg.2 \mid reg.1, reg.3, reg.4, reg.5, reg.6, reg.7, reg.8$	1

Table 6: Appendix B: Significant p-values for region 2 (part 2)

Eigenständigkeitserklärung

Ich versichere hiermit, dass ich die vorgelegte Bachelorarbeit eigenständig und ohne fremde Hilfe verfasst, keine anderen als die angegebenen Quellen verwendet und die den benutzten Quellen entnommenen Passagen als solche kenntlich gemacht habe. Diese Bachelorarbeit ist in dieser oder einer ähnlichen Form in keinem anderen Kurs und/oder Studiengang als Studien- oder Prüfungsleistung vorgelegt worden. Hiermit stimme ich zu, dass die vorliegende Arbeit von der Prüferin/ dem Prüfer in elektronischer Form mit entsprechender Software überprüft wird.

_____	_____
Ort, Datum	Katharina Ring