



Graduate School of  
Systemic Neurosciences  
LMU Munich

---

# Cortico-hippocampal activations for high entropy visual stimulus: an fMRI perspective

---

Nisha Dalal

Dissertation an der  
Graduate School of Systemic Neurosciences  
der Ludwig-Maximilians-Universität München



**First Supervisor: Prof. Stefan Glasauer**  
**Second Reviewer: Prof. Heiner Deubel**  
**Third Reviewer: Prof. Thomas Wolbers**  
**Date of Submission: September 15<sup>th</sup>, 2017**  
**Date of Oral Examination: December 14<sup>th</sup>, 2017**

*Although our intellect always longs for clarity and certainty,  
our nature often finds uncertainty fascinating.*

– Carl von Clausewitz



*Dedicated to all that is good and pure in this world...*

*...including my beautiful family and crazy friends!*



# Abstract

We perceive the environment around us in order to act upon it. To gain the desirable outcome effectively, we not only need the incoming information to be processed efficiently but we also need to know how reliable this information is. How this uncertainty is extracted from the visual input and how is it represented in the brain are still open questions.

The hippocampus reacts to different measures of uncertainty. Because it is strongly connected to different cortical and subcortical regions, the hippocampus has the resources to communicate such information to other brain regions involved in visual processing and other cognitive processes.

In this thesis, we investigate the aspects of uncertainty to which the hippocampus reacts. Is it the uncertainty in the ongoing recognition attempt of a temporally unfolding stimulus or is it the low level spatiotemporal entropy? To answer this question, we used a dynamic visual stimulus with varying spatial and spatiotemporal entropy. We used well-structured virtual tunnel videos and corresponding phase scrambled videos with matching local luminance and contrast per frame. We also included pixel scrambled videos with high spatial and spatiotemporal entropy in our stimulus set. Brain responses (fMRI images) from the participants were recorded while they watched these videos and performed an engaging but cognitively independent task. Using the General Linear Model (GLM), we modeled the brain responses corresponding to different video types and found that the early visual cortex and the hippocampus had a stronger response to videos with higher spatiotemporal entropy.

Using independent component analysis, we further investigated which underlying networks were recruited in processing high entropy visual information. We also discovered how these networks might influence each other. We found two cortico-hippocampal networks involved in processing our stimulus videos. While one of them represented a general primary visual processing network, the other was activated strongly by the high entropy videos and deactivated by the well-structured virtual tunnel videos. We also found a hierarchy in the processing stream with information flowing from less stimulus-specific to more stimulus-specific networks.





# Overview

In this thesis, we present our fMRI study investigating the hippocampal and cortical contributions to processing high entropy visual stimulus.

In **Section 1**, we introduce the literature and the open research questions relevant to our study. Section 1.1 describes the efficiency of our visual system and how it deals with the uncertainty in the incoming visual information. Section 1.2 focusses on the hippocampal contributions to various cognitive processes, especially in regularity extraction and dealing with highly uncertain input. In section 1.3, we discuss the current frameworks describing the cortical-hippocampal networks and functional importance of these connections. Section 1.4 presents the aim of this thesis. Section 1.5 introduces the various methods we used to achieve our aim e.g. General Linear Modeling, Independent Component Analysis.

In **Section 2**, we present the two manuscripts with their respective introduction, methods, results, discussion and conclusion. Section 2.1 presents our first manuscript which investigates the hippocampal activations while watching visual stimuli with high spatiotemporal uncertainty. Section 2.2 includes our second manuscript which investigates the underlying cortico-hippocampal networks while processing visual information with varying uncertainties and how they might influence each other.

In **Section 3**, we discuss the results and implications of our studies. Section 3.1 presents the purpose this thesis fulfilled. Section 3.2 primarily consists of a summary of our experimental findings. It is divided into several subsections focussing on different aspects of our results and the functions that we propose for the hippocampus and the cortico-hippocampal networks. Section 3.3 comprises of the 'take-home' message and the possible implications of this study. It also consists of a few suggestions for future studies involving the hippocampus.

**Section 4** contains all the necessary documents needed for the successful submission of this thesis including a curriculum vitae (Section 4.1), a list of publications (Section 4.2), thesis affidavit (Section 4.3) and proof of author contribution for the manuscripts (Section 4.4). In Section 4.5, we thank all those without whom this work would not have been possible.



# Contents

<b>1</b>	<b>General Introduction</b>	<b>1</b>
1.1	Visual system and uncertainty . . . . .	1
1.2	Hippocampus and uncertainty . . . . .	4
1.3	Cortico-hippocampal networks . . . . .	6
1.4	Aim of the thesis . . . . .	11
1.5	Methodological aspects . . . . .	12
<b>2</b>	<b>Visual spatiotemporal uncertainty and cortico-hippocampal networks</b>	<b>19</b>
2.1	Hippocampal activation while watching visual stimuli with high spatiotemporal entropy . . . . .	21
2.2	Dual cortico-hippocampal network activation while watching high-entropy visual stimuli . . . . .	50
<b>3</b>	<b>General Discussion</b>	<b>84</b>
3.1	Purpose of the thesis . . . . .	84
3.2	Summary of the experimental findings . . . . .	85
3.3	Possible implications . . . . .	97
3.4	Outlook and outstanding questions . . . . .	98
<b>4</b>	<b>Appendix</b>	<b>107</b>
4.1	Curriculum Vitae . . . . .	109
4.2	List of publications . . . . .	111
4.3	Eidesstattliche Versicherung/Affidavit . . . . .	113
4.4	Proof of author contributions . . . . .	115
4.5	Acknowledgments . . . . .	117



# 1 General Introduction

## 1.1 Visual system and uncertainty

We learn from our environment on daily basis. We observe our surroundings, associate past and present events and create a model of the surrounding world. We learn to interact with the environment. We use the model learned from our environment to act upon on the incoming information to achieve the optimal outcomes of our choices.

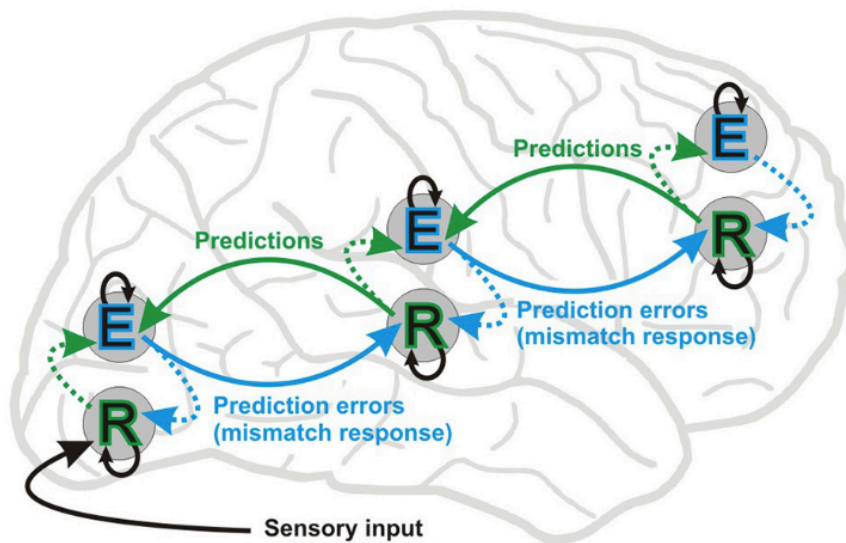
### Efficiency in our visual system

Our natural environment has regularities in both space and time. Detecting these regularities help us not only to adapt and learn from the environment but also to react to the incoming information in order to achieve a desirable outcome. Given these regularities, we can even predict the future inputs and act accordingly. For example, when we see a car on the street travelling at a specific speed, we can already predict its position in the next few seconds without actually seeing it there. We can use such information to decide whether we can cross the road before the car arrives. All these decisions are based on our environmental model of how a car works and how to cross the road.

Our visual system aims to extract behaviourally relevant information from our complex and dynamic environment and process it in the most economical way. It is structured for efficiency along all the visual processing regions, from the retina and LGN (Balasubramanian and Sterling, 2009; Ratliff et al., 2010) all the way to higher cortical regions. This efficiency is implemented both anatomically and functionally. For example, Ratliff et al. (2010) recently showed that, to accommodate for the excess of negative contrasts in natural scenes, the retina is structured so that the pathways encoding the negative contrast exceed those encoding the positive contrast twofold. This shows how our visual system has evolved anatomically over time to represent and process the visual information in the most efficient way.

### Predictive coding

Our visual system receives lots of redundant visual input. We can



**Figure 1:** Predictive coding framework. This framework shows the predictive coding mechanism at different levels of cortical hierarchy. E and R represent the error and the representation unit respectively. Here, the bottom-up stream feeds the prediction errors forward in the hierarchy and the top-down stream feeds back the predictions. At each level, E and R units interact to transform the previously obtained statistical regularities into predictions about the current sensory information. These units also receive inputs from layers immediately below and immediately above in the hierarchy. Reprinted from *Frontiers in Human Neuroscience*, Stefanics et al. (2014), Copyright 2014.

best utilize our brain resources by removing the redundant input and processing the non-redundant and behaviourally significant information only. Predictive coding explains one such hierarchical implementation in our cortex. It states that to make the best use of the brain resources, the expected information can be removed and only the unexpected information should be propagated along the cortical hierarchy (Rao and Ballard, 1999; Huang and Rao, 2011). This unpredicted part of the information is also called the residual error or prediction error. Such predictive coding mechanisms occur at every level of the cortical hierarchy and affect both feedforward and feedback connections. As shown in figure 1, while feedforward connections feed the higher cortical regions with a neural signal for the unexpected element (residual error); feedback connections, simultaneously, convey the expected information from higher cortical regions to the lower ones. If the sensory input from the lower areas matches the predicted input, the input signal is redundant and need not be processed. If it does not match, the difference between the two signals is forwarded

to the higher cortical regions.

### **Uncertainty and its impact on cognitive functions**

How information along the feedback and feedforward pathways is integrated depends on the reliability of the incoming information. If the incoming stream is highly unpredictable, the model cannot predict the next inputs and hence has to rely on the information coming in the bottom-up manner. In the latter case, more weight is given to the information from the bottom-up processing stream. But if the information is fairly predictable, the model can predict the next input correctly and top-down processing stream is given greater weight. In this case, less information is forwarded in a bottom-up manner as it is redundant. It is still unclear how these relative weights (or bias) are computed neurally for the uncertainties in the visual stream, but the hippocampus is proposed to be the region that could assign these relative weights to the processing streams (Strange et al., 2005).

Other than visual processing, the residual or prediction error could help other cognitive processes which require similar information about the statistical structure of the incoming stream. For example, if there is a prediction error, it indicates that our environmental model predicted wrong information given the previous input. To avoid repeating the same error, our environmental model has to be updated or replaced and hence learning mechanisms are employed (Schiffer et al., 2012). We can learn about relationships between objects, or relationships between objects and outcomes from our environment in either a one-shot manner or an incremental manner. Incremental learning can be used when we learn from multiple examples over time while one-shot learning is performed when we need to learn from as few examples as possible, sometimes just one. If the incoming visual stream is predictable, we employ incremental learning, otherwise one-shot learning is used. Lee et al. (2015) suggested hippocampal involvement in encoding the uncertainty signal for an environment that is highly uncertain and needs high learning rates (one-shot learning).

## Open questions

How our visual system computes and represents these visual uncertainties is still not well understood. Fraedrich et al. (2010) showed that current visual processing models, based on static images, do not work for dynamic and continuous input with higher uncertainties. The key to understanding *how these uncertainties are reflected in the brain* is identifying the brain regions and networks that react differentially to dynamic streams with different uncertainties.

The hippocampus is a strong candidate as it is sensitive to different measures of uncertainty (Strange et al., 2005; Harrison et al., 2006; Schiffer et al., 2012). It is also involved in learning the statistical structure of a continuous stream, even without the viewers’s knowledge (Turk-Browne et al., 2009). It also has strong anatomical and functional connections with different cortical regions. Given this evidence, the hippocampus seems a very likely candidate for extracting and/or representing uncertainty from the visual stream. In section 1.2, we will discuss how the hippocampus responds to the uncertainties in visual inputs.

## 1.2 Hippocampus and uncertainty

After years of debate concerning whether the hippocampus and other Medial Temporal Lobe (MTL) regions could be involved in cognitive processes outside the Medial Temporal Lobe Memory System (MTLMS) (see Suzuki (2009) and Baxter (2009) for arguments on both sides), it is now widely accepted that the hippocampus can indeed be a part of other cognitive processes like visual perception and learning (see Lee et al. (2012) for review).

### Hippocampal role in extracting regularity

The hippocampus is sensitive to the probabilistic context of event streams and reacts to different estimates of uncertainty like entropy (Strange et al., 2005), mutual information (Harrison et al., 2006), conditional entropy (Ahlheim et al., 2014) and complexity (Baumann and Mattingley, 2013). It plays an important role in learning regularities



rapidly from temporal sequences (Harrison et al., 2006; Turk-Browne et al., 2009; Schapiro et al., 2012; Bornstein and Daw, 2012; Schiffer et al., 2012; Hsieh et al., 2014; Reddy et al., 2015). Hippocampal activity patterns carry information about the temporal positions of objects in learned sequences, but not the object themselves. This role in processing improbable events is based on learning that oddballs can occur, as opposed to the encoding of their actual occurrence (Strange et al., 2005). It reacts strongly to the predictability of transitions in complex stimuli, without a priori knowledge of the stimuli structure.

Extracting these regularities rapidly from the environment, the hippocampus forms a relational representation of the environment. These newly built representations are then compared to previously stored representations in an 'associative mismatch' manner to indicate any mismatch (Schiffer et al., 2012; Olsen et al., 2012). This function captures the essence of the predictive coding in the cortex.

### **Hippocampal involvement at different levels of granularity**

Steinmetz et al. (2011) showed hippocampal responses to different levels of abstractions of the stimulus images. While some neurons reacted to the low level features of the images (contrast), other neurons reacted to higher levels like stimulus-class. These dual functional responses indicate information processing at different abstraction levels in the hippocampus. Evensmoen et al. (2015) also reported similar encoding of multiple levels of abstraction in the hippocampus with granularity gradient (coarse-grained to fine-grained) along the hippocampal and entorhinal anterior-posterior axis. While initially surprising, these results can explain the ability of the hippocampus to function at different levels of granularity. Reacting to low level features (and hence low level abstraction) can be important for creating individual episodes as declarative memory. Information with high level features (and hence a high level of abstraction) can be used to find similarities in the stimulus set and hence extract regularities from the stimulus stream. These regularities can be used in statistical learning to compute the relationship between different events or episodes.

To understand how these multiple levels of abstraction could be im-

plemented in hippocampal circuitry, Schapiro et al. (2017) created a computational model (neural network based on hippocampal anatomical connections) based on Complementary Learning Systems (CLS) theory. This theory emphasised the need for compromise between the abstraction levels in order for the hippocampus to be able to perform different functions. They found that two completely different pathways are responsible for the two functions. The monosynaptic pathway (entorhinal cortex connected to CA1 directly) enabled statistical learning while the trisynaptic pathway (entorhinal cortex connected to CA1 through dentate gyrus and CA3) helped in learning individual episodes.

### **Open questions**

Fraedrich et al. (2012) showed stronger hippocampal activity for phase scrambled videos compared to more structured and predictable virtual tunnel videos. These phase scrambled videos, with higher uncertainty in the direction of motion flow, did not reveal any clear meaning. Though hippocampal activity was associated with the unpredictability in the videos, this unpredictability could not be associated with any granularity level. *Did the hippocampus react to the low level spatial and/or temporal uncertainty; or did the uncertainty of the semantic content lead to the hippocampal activation?*

To understand how the hippocampus and the visual cortex react to visual inputs with different levels of uncertainty, we need to understand how they are connected anatomically and functionally. Therefore, we will present a brief overview of different functional accounts of these cortico-hippocampal connections in section 3.2.

## **1.3 Cortico-hippocampal networks**

The hippocampus is strongly connected to many other parts of the brain, including the temporal cortex, dorsolateral prefrontal cortex, visual cortex, lateral parietal cortex, and the midbrain and striatum via anatomical connections (see Bird and Burgess (2008) for review). These connections are presented in Figure 2. Such high interconnectivity allows multiple

models that could explain the functional importance of these pathways. We will introduce a few different models of hippocampal connectivity and its significance for various cognitive functions.

### **Extension of ventral visual stream**

Bussey and Saksida (2007) extended the ventral visual stream to a representation based ventral visual-perirhinal-hippocampal stream with the hippocampus at the top layer of the stream. This self-organizing representational stream consists of hierarchical representations with increasing complexity through various levels.

### **Balancing top-down and bottom-up processing streams**

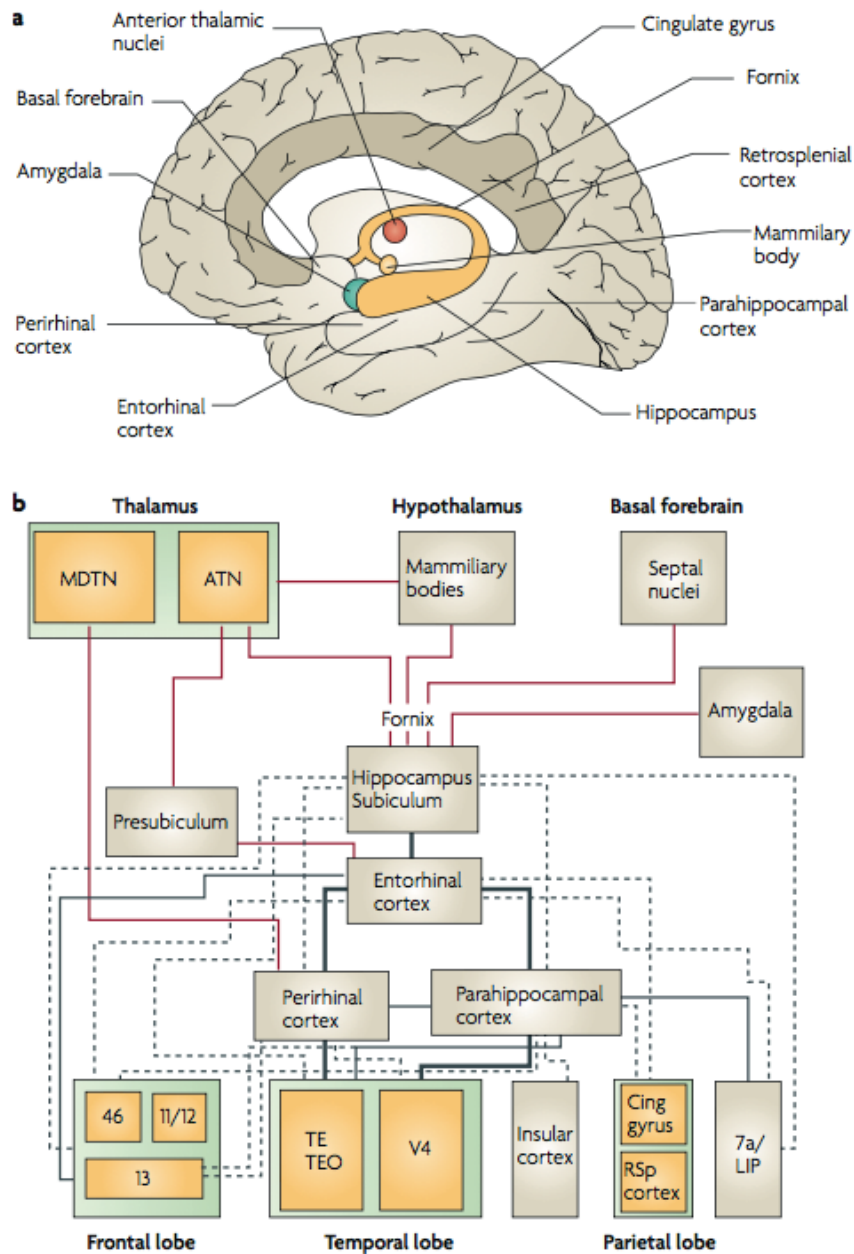
Information from top-down and bottom-up processing streams needs to be integrated to decide the best strategy for processing the incoming visual stream. For such an integration, relative weights are assigned to both streams as per the uncertainty in the input stream. Hippocampal sensitivity to statistical (ir)regularities and its connectivity to various cortical regions make the hippocampus an ideal candidate to provide such weights (Strange et al., 2005).

### **Bridge between lower and higher cortical levels**

Mozaffari (2014) suggested that the hippocampus acts as a bridge connecting higher and lower levels of visual processing. He pointed out that not all the communication between the lower and higher cortical regions needs to pass through the medial temporal lobe (MTL) but if levels are to be skipped (for rapid informational flow), a parallel connection can be established via the hippocampus and other MTL regions. This parallel architecture provides an optimal information flow by by-passing the intermediate levels and hence making predictive coding even more effective.

### **Hierarchical memory model**

The hierarchical memory model, suggested by Hasson et al. (2015), proposed processing and storing of the information in all cortical regions but at different timescales. Brain regions positioned at a lower level in



**Figure 2:** Hippocampal connections to other subcortical and cortical regions. Figure 2a labels the regions that share major anatomical and functional connections with the hippocampus. Figure 2b shows the interconnections between these major brain regions. Red lines represent the connections with subcortical regions while the black lines represent the connections to cortical regions. The thickness of the lines show the strength of the connection. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience, Bird and Burgess (2008), copyright 2008.

the hierarchy process and store information at detailed granularity level spatially and temporally (like contrast) but at shorter time scales. On

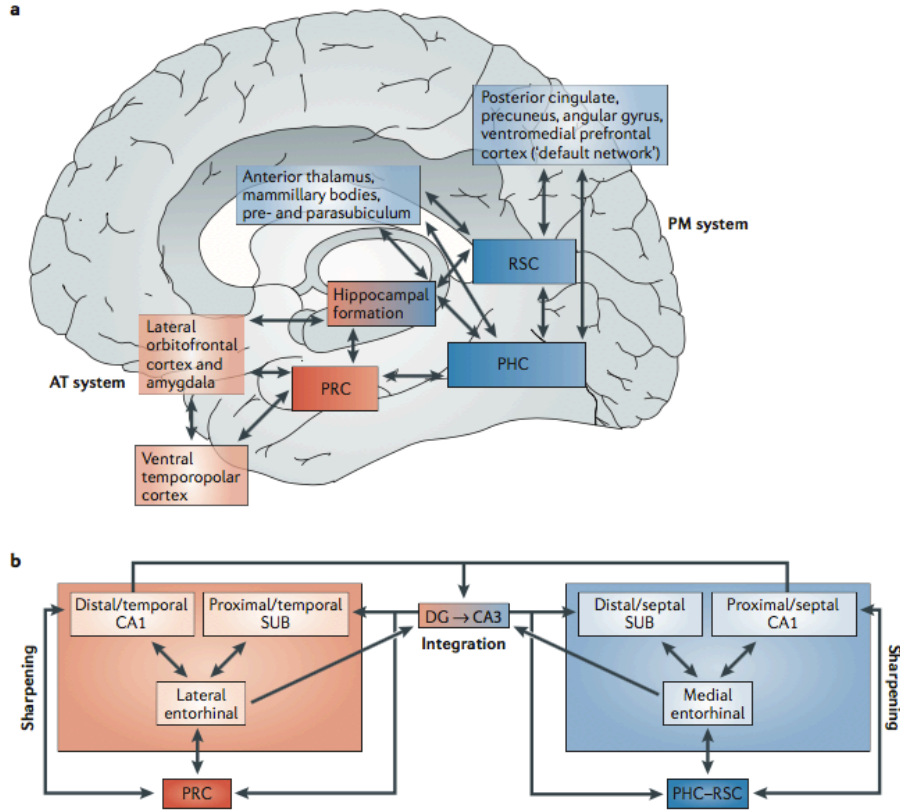
the other hand, brain regions at a higher level of the hierarchy process and store more coarse level information (like stimulus class) and function at longer timescales. Even though they included hippocampus in this model, the authors however, could not determine its time course. It does not come as a surprise given the evidence for the hippocampus working at different timescales for different cognitive functions (Squire and Zola, 1998; Hannula et al., 2006; Pertzov et al., 2013; Hannula and Duff, 2017). In other words, there might be overlapping areas of spatial and temporal activity in the hippocampus that need to be separated.

### **Posterior Medial Anterior Temporal (PMAT) framework**

Kahn et al. (2008) found two parallel cortical pathways (Posterior Medial and Anterior Temporal) extending to different subregions of the hippocampal formation: the hippocampal body and anterior hippocampus respectively (as shown in Figure 3).

The PM network consists of the parahippocampal cortex (PHC) and the retrosplenial cortex (RSC) along with the precuneus, posterior cingulate, angular gyrus, anterior thalamus, presubiculum, mammillary bodies and medial prefrontal cortex and extends to the hippocampal body. Major functions served by this network and its underlying regions include episodic memory, spatial memory, temporal context and social cognition. The AT network, on the other hand, includes the perirhinal cortex (PRC), ventral anterior temporal cortex, lateral orbito-frontal cortex and amygdala. This pathway converges at the anterior hippocampus and plays an important role in recognition and associative memory, affective processing, semantic processing, object perception and associations between objects and rewards.

The hippocampus, which is strongly connected to both these networks, could be the region where information from both streams is integrated to create a meaningful relational representation of the input (see Ritchey et al. (2015) for review). These two networks perform different cognitive functions and can be function at different abstraction levels and timescales.



**Figure 3:** Posterior Medial Anterior Temporal (PMAT) framework. Figure 2a shows the brain regions involved in the PMAT framework. Regions shown in blue belong to the PM network while the regions in red represent the AT network. Figure 2b shows how these systems extend to different subregions in the hippocampal formation, enabling it to perform different functions (sharpening and integration). Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience (Ranganath and Ritchey, 2012), copyright 2008.

### Open questions

Anatomical and functional connectivities explain how these cortico-hippocampal networks could include both parallel and hierarchical components of the networks (Bergmann et al., 2016). Such a highly interconnected position could enable the hippocampus to integrate information at different abstraction levels as per the demands of the various cognitive processes. *How these cortico-hippocampal networks are involved in processing uncertainties in the visual data* is still an open question. *At what abstraction level do they integrate information? Can they function at more than one granularity level?* We address these questions in this thesis to gain a fundamental understanding of how uncertainties are rep-

resented and processed in cortical and subcortical regions.

## 1.4 Aim of the thesis

Information about uncertainty or predictability of the visual input stream could be useful for various cognitive functions. The hippocampus has been linked not only to different measures of uncertainty but also to the information about the presence or absence of such regularities. Such information could be used to weigh the top-down and bottom-up processing stream relatively (Strange et al., 2005). Such information could also help in deciding, which learning strategy (one-shot or incremental) to employ given an input with a certain predictability (Lee et al., 2015).

This doctoral thesis aims to understand the brain’s response to a dynamic visual stimulus with high uncertainties (entropy) using fMRI measurements. Fraedrich et al. (2010, 2012) recorded brain activities while subjects watched well-structured and predictable virtual tunnel videos and their corresponding phase scrambled videos which had similar low level statistics but did not reveal any clear meaning instantly. They reported task-independent strong activations in the hippocampus and the early visual cortex for the phase scrambled videos when compared to the virtual tunnel videos.

In this thesis, we inspect the granularity of this above-mentioned hippocampal involvement. We investigate if this is due to an ongoing recognition process dedicated towards uncovering the content of the visual stimulus, or is it independent of the semantic content of the stimulus and occurs due to its high spatiotemporal entropy at low level stimulus statistics.

To answer these questions, we measured brain activities from participants using fMRI while they watched similar virtual tunnel and phase scrambled videos. We additionally used spatiotemporal stimuli (pixel scrambled videos) that do not have any meaningful content but have high entropy. Participants were instructed to perform a near threshold arrow detection task to keep the attention high all the time. This task was independent of the content of the videos and did not need any

explicit semantic processing of the videos.

We compared these stimulus types using General Linear Modeling (GLM) to find differences between brain activations for the different stimulus types. We further used multivariate Partial Least Squares Correlation (PLSC) to find the regions explaining the entropy timeline. Moreover, we examined which functional networks are involved while processing these videos using Independent Component Analysis (ICA) to discover components with maximal spatial independence. We then evaluated the functional properties of these networks, their time courses and their relationships to the experimental design. We explored which cortico-hippocampal networks were recruited and at which timescale. To understand the temporal dependence between these networks, we calculated the functional effective connectivity between components using Granger causality.

In this thesis, we reveal the reasons behind the cortical and hippocampal responses to dynamic visual input, underlying cortico-hippocampal networks and their influence on each other using neuroimaging methods. In section 1.5, we introduce the methods we used to answer the above mentioned questions.

## 1.5 Methodological aspects

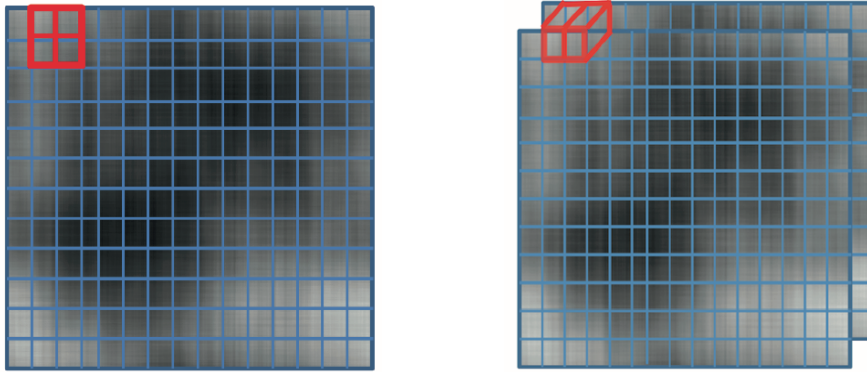
This section consists of the basic introduction to the methods we employed to analyze our fMRI data to achieve the aims described before.

### Entropy calculations

Entropy expresses the disorder or uncertainty of the system. It depicts the uncertainty of the next outcome in the system. For example, for every coin toss, the probability of receiving either a head or tail outcome is  $(1/2)$ . Every coin toss is equally uncertain and does not contain any information about the one to follow. Entropy is high for this system. Entropy is zero when one outcome is certain to occur. This statistical measure of randomness can also be used to estimate the amount of informational content in an image. It can be represented by the probabilistic



distribution of the pixel values in the image. For example, an image with a huge section of clear blue sky will have a low entropy as most of the pixels in the image will just be representing sky and will have a similar pixel value, creating a smooth texture.



**Figure 4:** Entropy calculations in our study. We calculated word-based entropy for estimating the spatial and spatiotemporal uncertainty in our stimulus videos. Figure 4 (left) shows our definition of 2X2 word length for spatial entropy. For spatial entropy, a single frame was used to calculate the frame-wise entropy, regardless of the word length. Figure 4 (right) represents our definition of 2x2 word length for computing spatiotemporal entropy. As you can see in this figure, more than one frame was used to estimate spatiotemporal entropy to include the temporal variability of the stimulus. The number of frames used to calculate the 'frame-wise' entropy depended on the word length (word length = number of frames used).

We used entropy as a measure of uncertainty in our visual stimulus. To consider both the spatial and temporal aspects of our three types of videos, we calculated word-based spatial and spatiotemporal entropies for each frame of the video. A word was defined as a matrix of size (word length)\*(word length) on consecutive pixels. We calculated the entropy of each frame for various word lengths. It had a slightly different implementation for spatial and spatiotemporal entropies. For spatial entropy with word length 2x2, we used a 2x2 sliding window (as shown on the left side of Figure 4) and calculated the frequency of all combinations of grayscale values for this window in one frame. To include the temporal variability in the videos, the word for spatiotemporal entropy was defined to span through multiple frames at once. The number of frames used for calculating the entropy was specified by word length. As in Figure 1

(right), two consecutive frames were used to compute the frequency of each 2x2 block, again by sliding the window through the frames. These frequencies were then used to compute the probabilities of all grayscale combinations. The entropy of each frame of the stimulus videos was then calculated using the following formula:

$$H = \sum_i p_i * \log(p_i),$$

where  $i$  represents a specific combination of (word length\*word length) grayscale values and  $p_i$  is the probability of this permutation of grayscale values in a frame. As a grayscale range of 256 values was large for a reasonable estimation of word-based frequencies ( $256*256*256*256$  possible combinations for word length=2), we downsampled the videos by discretising them into different grayscale levels (for details, see the method section from the first manuscript) and calculated the entropies of these new video frames.

### **General Linear Modeling (GLM)**

The GLM is a linear model, given some variable measurements, which is used to estimate the parameters of the design matrix. For example, for a matrix  $Y$  of measurements and a design matrix  $X$ , parameters in  $b$  can be estimated so as to maximize the explainable variance in  $Y$  and to minimize the error  $e$ .

$$Y = Xb + e$$

For fMRI data, GLM is used as a univariate analysis, where each voxel response is modeled separately to predict or to estimate its response in terms of a linear combination (weighted sum) of several reference functions or regressors. The time course of a single voxel is represented in  $Y$ . A set of regressors, related to the experimental paradigm, form the design matrix  $X$ .

In our analysis, we created a design matrix using our experimental design, which consisted of one column for each stimulus type and one row for each measurement in the run. This design matrix carried information

about the presence or absence (1 or 0 respectively) of each stimulus type at a certain time point. Six columns of motion regressors were also added to this model. After a high-pass temporal filter (210 seconds) has been applied, this design matrix was then convolved with the canonical hemodynamic response function (HRF). Beta weights ( $b$ ) for the experimental conditions were then estimated using the voxel measurements ( $Y$ ) and design matrix ( $X$ ) to explain the variance in  $Y$  in terms of  $X$ , weighted by  $b$ . The response of each voxel for every condition of the experimental paradigm (virtual tunnel, phase scrambled and pixel scrambled videos) was calculated for each participant. Group analysis was performed using the t-test to calculate the voxel-wise difference between conditions at group level.

### **Partial Least Squares Correlations (PLSC)**

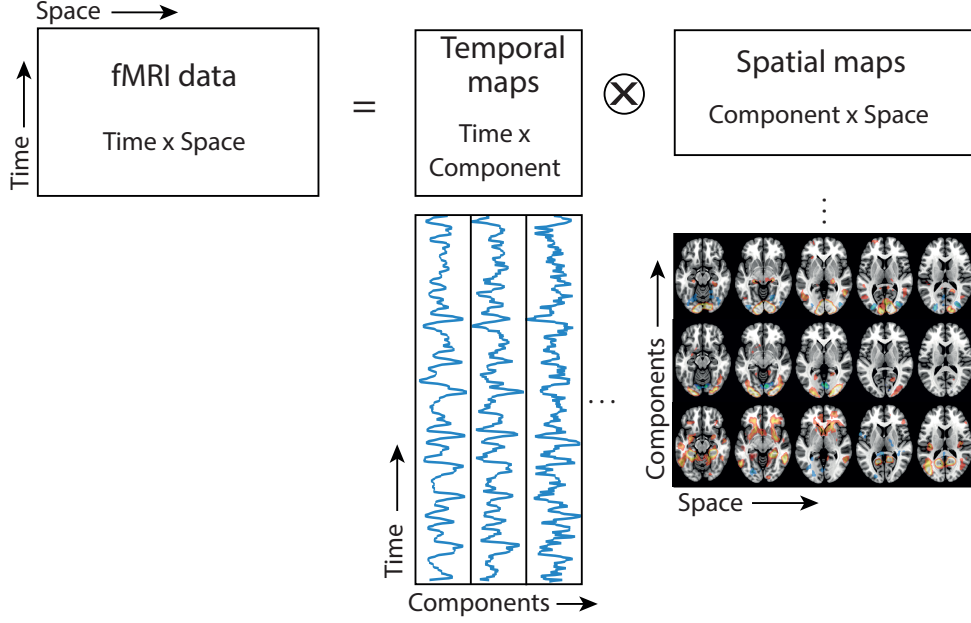
This technique was used to analyse the relationship between measures of brain activity (voxel time course) and of measures of behaviour or experimental design (entropy timeline in our case) (Krishnan et al., 2011).

The main analytical tool for PLS is the singular value decomposition (SVD) of a matrix as it provides the best reconstitution (in a least squares sense) of the original matrix by a matrix with a lower rank. The SVD of a matrix  $R$  decomposes it into three matrices:  $R=U\Delta V^T$ . The singular vectors  $U$  and  $V$  are also called saliences.  $U$ , the left singular vectors of  $R$ , represents the design or behavioral profiles that best characterizes  $R$ , on the other hand  $V$ , the right singular vectors of  $R$ , represents the voxels or brain images that best characterizes  $R$  (Krishnan et al., 2011).

In our analysis, entropies from the stimulus videos were concatenated in the order of the actual stimulus presentation for each subject to create an entropy timeline. This entropy timeline was then convolved with the HRF to match the delayed brain activity due to the slow nature of the BOLD response activity in the human brain. This entropy timeline was then correlated with the brain activity using PLSC method. We thus found that voxel weights contributions to the entropy timeline in a multivariate manner. These maps from all the subjects were then transformed into the z-scores and analyzed at the group level using the t-test.

## Independent Component Analysis (ICA)

ICA is a computational method for separating a multivariate signal into additive and independent subcomponents, assuming that its subcomponents are statistically independent non-Gaussian signals.



**Figure 5:** Separating fMRI data into components using ICA. The spatiotemporal fMRI data represents the activity of all the voxels on the brain over time, hence its dimension is Time X Space. This data is separated into independent components represented by temporal and spatial maps. Temporal maps is a (Time X Component) size matrix and represents the temporal information of the independent component. Each column of the temporal maps is the time-course of each component. Spatial maps consist of the voxel maps of the components and hence has a size of (Component X Space). Each row of the spatial map represents the voxel maps of each component.

ICA was used for the first time for separating fMRI signals into independent components by McKeown et al. (1998) as a data-driven approach. If the fMRI signal is represented by a time-space data matrix of measurements  $X_{j,t}$  where  $j = 1, \dots, J$  ( $J$  is the total number of voxels) and  $t = 1, \dots, T$ , ( $T$  is the number of time samples), it can be modeled as follows:

$$X_{t,j} = \sum_{k=1}^K S_{t,k} A_{k,j} + E_{t,j}$$

where  $A$  and  $S$  are formed by  $K$  independent components (as shown in

figure 5). Columns of  $S$  represent time courses, and rows of  $A$  represent spatial maps of the respective components, and  $E$  is spatially and temporally white noise (McKeown et al., 2003).

In our study, we used the Group ICA of fMRI Toolbox (GIFT) (Calhoun et al., 2001) to compute the independent components from our preprocessed fMRI data. This computational analysis was performed in three steps: data reduction, ICA and then back-reconstruction. Data reduction was performed by concatenating fMRI data from all the participants and removing dimensions that do not add anything to explainable data using Principal Component Analysis (PCA) (Li et al., 2007). Spatial ICA was then performed to calculate the independent components using fastICA algorithm. Finally, results from ICA were then back-reconstructed to individual subject space.

### **Functional Network Connectivity (FNC): Granger Causality**

Regions within a given component (from spatial ICA), by definition, show strong temporal coherence due to the linear mixing assumption of ICA. In spatial ICA, the resultant spatial maps are maximally independent, but the time courses are not independent and can exhibit considerable temporal dependencies. These temporal dependencies can reveal the functional connectivity between the components. Different measures, such as a lag-based method (Jafri et al., 2008) or the Granger causality (Demirci et al., 2009) can be used to quantify these dependencies.

Granger causality analysis identifies directed functional (“causal”) interactions from time-series data. G-causality says that a variable  $X$  “G-causes” another variable  $Y$ , if the past of  $X$  contains information that helps predict the future of  $Y$ , to a greater extent than the information already in the past of  $Y$  itself (and in the past of other “conditioning” variables  $Z$ ). In general terms, when this condition is satisfied one can say that there is “information flow” from  $X$  to  $Y$  (Seth et al., 2015). This is justified because G-causality is an approximation to transfer entropy (the approximation is exact for Gaussian variables), which is a directed version of Shannon’s mutual information. This, in turn, is a very general way of characterizing the statistical dependency or shared information between two variables (Barnett et al., 2009).

We used the Granger causality method from the FNC toolbox (Demirci et al., 2009) to investigate the causal relationships between spatially independent components. A few selected components (highest variability with the experimental design) were first filtered with a Butterworth filter in the range of 0.08 Hz to 0.24 Hz. Granger causality coefficients were calculated for different frequencies for all component pairs. These coefficients allowed us to infer the 'G-causal' relationships between components.

Using these methods, we identified brain regions and networks involved in processing visual stream with high entropy. We present the study and findings in the next section (section 2.1 and 2.2) of this thesis.

## **2 Visual spatiotemporal uncertainty and cortico-hippocampal networks**

This cumulative-style thesis consists of two detailed manuscripts (section 2.1 and section 2.2). This chapter presents the two manuscripts and their corresponding author contribution statements.





## 2.1 Hippocampal activation while watching visual stimuli with high spatiotemporal entropy

**Authors:** Nisha Dalal, Virginia L Flanagin, Stefan Glasauer

### Contributions

Nisha Dalal designed and programmed the experiment, collected and analyzed data, created figures, interpreted the results and wrote the manuscript.

Virginia Flanagin supervised the project, helped in interpreting results and critically reviewed the manuscript.

Stefan Glasauer supervised the project, helped in designing the experiment and in interpreting results and critically reviewed the manuscript.



# Hippocampal activation while watching visual stimuli with high spatiotemporal entropy

Nisha Dalal <sup>1,2</sup>, Virginia L Flanagin <sup>1,2,3</sup>, Stefan Glasauer <sup>1,2,3</sup>

<sup>1</sup> Department of Neurology, LMU, Munich, Germany

<sup>2</sup> Graduate School of Systemic Neurosciences, LMU, Munich, Germany

<sup>3</sup> German Center for Vertigo and Balance Disorders, LMU, Munich, Germany

**Keywords:** Hippocampus, entropy, fMRI, visual perception, Partial Least Squares



## Abstract

Uncertainty of the sensory information is an important factor in determining the reliability of the incoming information in the brain. In previous fMRI experiments, we found stronger hippocampal and early vision activity while subjects watched phase scrambled videos compared to their corresponding, but more structured, virtual tunnel videos. These phase scrambled videos did not reveal any clear meaning. Since the hippocampal involvement was not reported when subjects were presented with static phase scrambled frames, the dynamic nature of the stimulus turned out to be an important factor in hippocampal recruitment. Two apparent reasons for the hippocampal recruitment were tested in the experiment we present in this contribution. First, it could be involved in an implicit associative memory retrieval when visual stimuli have a systematic structure but still do not reveal a clear meaning. The second reason for its recruitment could be the high level of uncertainty of the visual stream itself, irrespective of stimulus content. To investigate the aforementioned reasons, we added pixel scrambled videos (very high spatiotemporal entropy but no semantic content) to the previous stimulus set. Twenty-seven subjects (mean age  $25.4 \pm 3.7$  years, 16 females) participated in an fMRI experiment with two experimental runs (and two training runs) and performed an arrow detection task, irrelevant to the content of the videos. We observed that the hippocampus and early visual areas reacted strongly to the high entropy pixel-scrambled videos compared to all the other types of videos, even though the underlying task did not involve processing the statistical structure and/or content of the videos explicitly. This highlights the possibility of the hippocampus being relevant for the entropy of the input stream itself. The results show that entropy of the incoming dynamic information, though never related directly to the hippocampal activity before, is an important determinant of the hippocampal involvement.



# 1 Introduction

Despite its complexity, our physical world is embedded with structures and regularities at every scale over space and time. These regularities help us to extract information about current and future events. Neuroimaging research, over the past few years, has identified several brain regions that are sensitive to the (ir)regularities of the sensory input stream, the hippocampus being one of them.

The hippocampal system is known to be involved in statistical learning (Turk-Browne et al., 2009, 2010; Zhao et al., 2011; Bornstein and Daw, 2012; Schapiro et al., 2012) and learning by mismatch (Schultz et al., 1997; Schultz and Dickinson, 2000). It has been consequently dubbed as an 'associative mismatch detector' (Kumaran and Maguire, 2007). It is sensitive to measures of (ir)regularity such as entropy (Strange et al., 2005; Schiffer et al., 2012) mutual information (Harrison et al., 2006) and predictability (Ahlheim et al., 2014).

Recently, we have shown higher Blood-Oxygen-Level Dependent (BOLD) activity in the hippocampus and early visual cortex when human subjects watched phase scrambled videos compared to a movement through a virtual tunnel, even though the spatiotemporal amplitude spectrum of the videos was same (Fraedrich et al., 2010, 2012). One important constraint for this hippocampal activity was the dynamic nature of the visual content because similar hippocampal recruitment has not been found in subjects watching static phase scrambled images (Wichmann et al., 2006). Our previous fMRI experiments also established that hippocampal activation in these conditions was independent of the task performed by the subjects.

There might be two possible reasons for hippocampal involvement in processing these indistinct dynamic visual stimuli. First, it is possible that cortico-hippocampal networks are recruited for associative memory retrieval and attempted recognition when visual stimuli have a systematic structure but do not reveal a clear semantic meaning immediately. Second, hippocampus activity could be directly related to the level of uncertainty of the stimulus itself (Strange et al., 2005; Harrison et al., 2006).

To distinguish between these two possibilities, we test, in the present study, the effect of spatiotemporal entropy on visual perception by adding a white noise stimulus to the previous experiment stimulus set in an fMRI experiment. The white noise videos, pixel scrambled versions of the virtual tunnel videos, have high entropy (spatial and temporal uncertainty at low level statistics) but are meaningless and are of low complexity at higher levels of perception. This meaningless nature of the videos is apparent to subjects from the beginning and hence no conscious effort of recognition is expected from the subjects while watching the videos. Therefore, if the hippocampal activity is due to the ongoing recognition process (Brown and Aggleton, 2001; Yee et al., 2014), we hypothesize that we will observe higher activations in the hippocampus during phase scrambled stimulation (due to its inherent structure) compared to the virtual tunnel and pixel scrambled videos. On the other hand, if higher entropy at low level input statistics is causing hippocampal involvement, pixel scrambled videos should activate hippocampus strongly compared to the other two types of videos. In the current study, we thus address the following question: is the hippocampal involvement during dynamic visual stimulation related to attempted recognition or is it a part of continuous visual processing and caused by the low level features of the visual stream?

## 2 Materials and Methods

### Participants

Thirty-one healthy participants (age ranging from 19 to 33; mean age  $25.4 \pm 3.7$  years, 16 females) with normal or corrected-to-normal vision, no metallic parts in their body and no documented history of neurological or psychiatric history gave their informed consent to participate in the study. The experiment was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee of the medical faculty at the Ludwig-Maximilian University of Munich. The scanning session of one participant could not be completed due to large head movements during scanning sessions. Three participants were re-



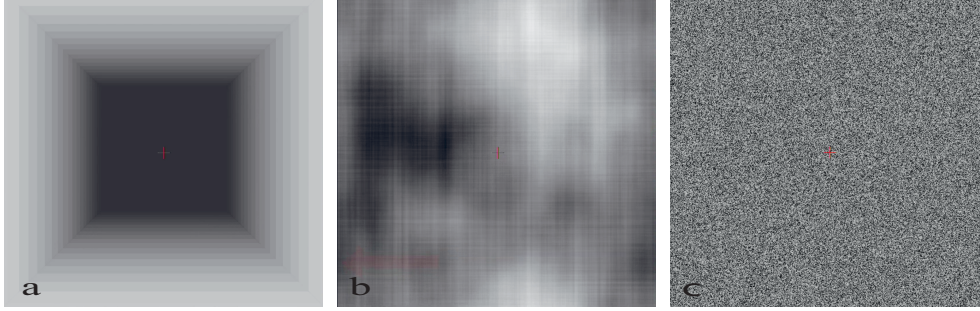
moved after preprocessing due to excessive head movements leaving 27 subjects (age ranging from 19 to 33; mean age  $25.6 \pm 3.8$  years, 14 females) for further analysis.

## **Stimuli**

The stimuli videos were created using the psychotunnel (Graham et al., 2010) software that was implemented in OpenGL. These videos were similar to the videos used in previous experiments (Fraedrich et al., 2010, 2012) except that they were 12 seconds long, instead of 6 seconds, so as to have single videos as a block. All videos (resolution of  $540 \times 539$ , 30 Hz sampling rate, 360 frames, 12 second duration) had three segments: a straight segment, a left or right turn and then another straight segment. Six different turn angles (35, 45, 75, 85, 115, 125 degrees) were used. For the second type of stimulus, virtual tunnel videos were transformed into the three dimensional Fourier space and the phase was randomly scrambled across all the dimensions while keeping the amplitude spectrum undisturbed. The resulting data were transformed back to the time domain resulting in the phase scrambled videos, which had the same local flow, average contrast and luminance as the respective tunnel videos. The local luminance difference over frames was matched for both video types. The third type of videos were created by randomly scrambling the pixels of the virtual tunnel videos resulting in the pixel scrambled videos. Some barely visible red arrows of different lengths (pointing either to the left or the right), which appeared for ten frames (0.33 seconds), were added to some of the videos. Half of the videos did not have any arrows and the other half had arrows chosen randomly between one and three. Example frames of all three types of videos are shown in Figure 1.

## **Task and training**

Participants were first trained in two training sessions (seven-eight minutes each) in a separate dark room outside the scanner. Each training session consisted of 30 videos with ten videos of each type (virtual tunnel, phase scrambled and pixel scrambled videos). Each video, followed by a dark blank screen for two seconds, was shown in a pseudorandom order on a laptop screen using Cogent 2000 software. A fixation cross was presented in the center of the screen throughout the sessions, including



**Figure 1:** Frames from the three video types a. virtual tunnel video, b. phase scrambled video and c. pixel scrambled videos. Videos were presented with a fixation cross in the center as shown in the figure. The phase scrambled frame (b) also shows one of the arrows (left bottom) to be detected by the subjects in the arrow detection task. These arrows were presented for only 0.33 seconds in the videos.

the blank screen (example in Figure 1b). Participants were asked to fixate on the fixation cross throughout the experiment and report the presence of arrows in the video by pressing a button immediately. An answer was considered correct only when the participants could detect all the arrows present in the video. Feedback was provided to participants after every video. They were allowed to perform the fMRI session only if they could perform the training session with a minimum accuracy of 60%. This task was chosen because it can be performed without analyzing or recognizing the content of the video and because it does not require short-term memory (Fraedrich et al., 2012). During pilot studies, we noted that arrow detection was more difficult in pixel scrambled videos, so the arrows in these videos were made darker than the original arrows from previous experiments (12/255 of the red component compared to 10/255 in other videos) to avoid any attentional bias due to difficult detection in a specific stimulus type during the actual experiment.

Two task-based sessions (15 min each) were performed when subjects were inside the fMRI scanner. Each session consisted of 20 videos of each type (120 videos in total), again separated by a two second blank space with fixation cross, presented in a pseudorandom order. Each video presentation was considered a block of functional scans. No feedback was provided to the participants inside the fMRI scanner.

## Entropy calculations

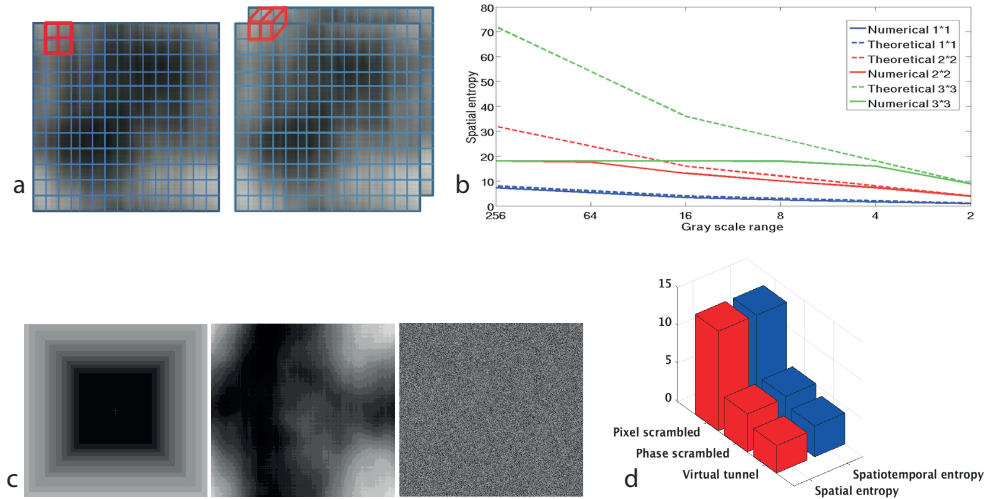
Entropy is a statistical measure of randomness that can be used to characterize the predictability of the input image. To account for the temporal unfolding of the stimulus, both spatial and spatiotemporal entropies were computed for all the videos. ‘Word-wise’ entropies were calculated per frame using sliding blocks of size  $n \times n$ , where  $n$  was the word length. Each block represented a word and hence word-wise entropies were calculated by computing distributions of these patterns on each frame.

The aim of our entropy calculation was to capture both the spatial and the temporal uncertainties of the videos. Therefore, we calculated both the spatial and the spatiotemporal entropy of each frame. This was achieved by taking  $n$  frames together at a time for word length  $n$  and then forming a  $n \times n$  block as shown in Figure 2a.

Given the size of our videos (360 frames of  $540 \times 539$  pixels each), we did not have enough sample points for all  $n \times n$  values for a range of 256, as the number of total possible combinations using a  $2 \times 2$  block in a 256 gray scale video is already  $256 \times 256 \times 256 \times 256$ . As the values were drawn from a random set, the probability of finding matching blocks was very low. For a feasible estimation of the entropy, we remapped the gray value range of 256 to different lower ranges (128, 64, 32, 16, 4, 2) by discretization (Figure 2b). This downsampling made it possible to calculate the entropy values with the limited number of frames at our disposal (360). We verified the downsampling by computing the theoretical and the actual entropy values for pixel scrambled noise videos. As shown in Figure 2b, the grayscale range 16 still yields good estimates for  $2 \times 2$  word length, so that for word length 2, we decided to use this range for further computations of entropies for all types of videos. An example of video frames corresponding to this downsampling is shown in Figure 2c.

To estimate entropy, for each new frame, we computed the probability of each  $2 \times 2$  block being present in the frame (frequency/total positions). The probabilities were then used to compute the entropy of the videos using the formula:

$$H = \sum_i p_i * \log(p_i).$$



**Figure 2:** entropy calculations. a. definition of ‘blocks’ for spatial and spatiotemporal entropy calculations. The frequency of each such block was calculated and was used to compute the probability of finding the same block in a frame. The number of such unique blocks and their probabilities were then used to calculate the entropy of each frame, b. comparison between theoretical and actual (numerical) spatial entropy values of a pixel scrambled frame with respect to the block size and the degree of downsampling (grayscale range), c. an example of downsampled (grayscale range=16) video frames, d. mean spatial and spatiotemporal entropy values of the down-sampled videos (block size=2\*2 and grayscale range=16).

To account for the temporal dynamics of the videos, we used the spatiotemporal entropy for further analysis. For a block size as small as 2, spatial and spatiotemporal entropies are not very different from each other (see Figure 2d). For further analysis, we refer to spatiotemporal entropy as entropy, unless specified otherwise.

## fMRI acquisition

fMRI data was acquired using an eight-channel head coil on a 3T whole-body MR scanner (Signa HDx, GE healthcare) with a T2\*- weighted gradient-echo, echo-planar sequence (TR= 2 sec, TE= 30ms, FOV= 220 mm, matrix= 64\*64, voxel size 3.28\*3.28\*3.49 mm). The whole brain was scanned with 38 axial slices with a slice thickness of 3.5 mm with no interslice gap. The first four functional scans were discarded at acquisition level due to T1 stabilization. Padding and adjustable head restraints were used to minimize the head motion. Apart from the functional scans, a high resolution (voxel size 0.85\*0.85\*0.7 mm) whole brain fast spoiled

gradient echo (FSPGR) anatomical image was also collected from each participant.

### **fMRI preprocessing**

Prior to preprocessing, functional and anatomical brain images of all subjects were extracted from the skull using Brain Extraction Tool (BET) v2.1 (Smith, 2002) from FSL (FMRIB Software Library) using a fractional intensity threshold of 0.5. After initial extraction, further preprocessing steps were performed using SPM12 (Wellcome Trust Centre for Neuroimaging, UCL, London, UK). Volumes from both functional sessions were aligned to the each other using the mean image as the reference scan and 0.9 quality using a second degree b-spline interpolation. These functional scans were corrected and resliced for head movements using a six parameter rigid body transformation and 4<sup>th</sup> degree b-spline interpolation. These volumes were then coregistered to the high resolution anatomical image of the participant by maximizing the normalized mutual information. Individual anatomical images were segmented using the white matter, the gray matter and the cerebrospinal fluid tissue probability maps from SPM using a mutual information affine registration to ICBM template to achieve affine regularization. Very light bias regularization (0.0001) was performed along with warping with a frequency cutoff of 25. The DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra) tool (also from SPM) was then used for nonlinear image registration to iteratively match the gray matter and white matter images of all subjects to the template mean image generated from those images. These registrations were penalized by linear elastic energy using 0.01 Levenberg-Marquardt regularization in three cycles and three iterations. The two functional scans and the gray matter images, now matched to the average template, were then normalized to the MNI template using the DARTEL flow field created above for each subject by preserving concentration. Smoothing was then applied using a gaussian blurring kernel (4 mm full width/half-maximum isotropic) to the normalized functional scans to reduce the aliasing effects.

### **Generalized linear modeling (GLM)**

Statistical analysis of fMRI data was first performed at first level (in-

dividual subject) with mass-univariate approach using GLM. The fMRI model was specified for 439 scans (TR= 2 seconds; microtime resolution= 16 seconds; microtime onset = 1) for each functional run for each subject. Three conditions (virtual tunnel, phase scrambled, pixel scrambled videos) were used as the task regressor. The baseline for each run was modeled separately. Movement parameters, computed during realignment, were also used as regressors. Task regressors were convoluted with the canonical hemodynamic response function without any derivatives. All the functional scans were temporally filtered with a high pass filter of 210 seconds to remove the scanner drift. Autoregressive modeling AR(1) was specified to account for the autocorrelation due to aliased biorhythm and other unmodeled neural activities. Model parameters (in the form of beta images) for first level fMRI analysis were estimated using the classical ReML (Restricted Maximum Likelihood).

T-contrasts were created with multiple one-dimensional contrast vectors corresponding to the difference between the three conditions (for example, pixel scrambled videos - virtual tunnel videos). Additional contrasts were created to compute the difference between activations corresponding to virtual tunnel videos and the two other videos together. Similar contrasts were created for pixel scrambled videos versus the two other types of videos. Factorial design was specified for random-effects (second level) analysis using the contrast images of all subjects from first level analysis by ordinary least squares (SPM defaults) for group inference. Implicit masking was performed to exclude the NaNs from the contrast images. These model parameters were estimated using the classical ReML at group level as well.

We used the threshold  $p < 0.05$  (FDR corrected) for all contrasts of interests from group analysis (for example, virtual tunnel – pixel scrambled videos). Clusters with less than 10 voxels were discarded. No external masking was applied. For identifying the anatomical structures in the GLM analysis, the Harvard-Oxford Cortical Probability Atlas, the Harvard-Oxford Subcortical Probability Atlas and the Juelich Histological Atlas were used (Maldjian et al., 2003).

### **Partial least squares (PLS) analysis**

In neuroimaging, PLS methods have been used to analyze the relationship between different experimental factors and the brain activity (McIntosh et al., 1996; Garrett et al., 2010; Krishnan et al., 2011). These factors could range from behavioural measures to the seed activity and the experiment designs. In this study, PLS correlations were used to examine the relationship between the brain activity and the entropy of the videos.

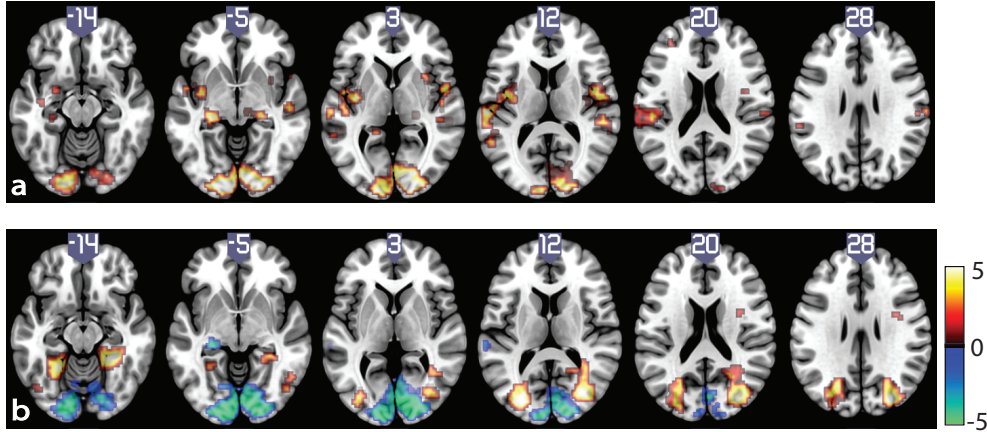
Entropies of the videos were calculated frame wise, as explained in the method section. Entropy time-series were created for each subject by concatenating the entropy values as per the onset and duration of the corresponding video during the experiment. The entropy of the blank screen with the red fixation cross was also calculated and it replaced the two second gap between the video presentations as in the experiment design. These timelines were convoluted with the canonical HRF model (from SPM) to align with the delayed hemodynamic function. The delayed timelines were then compared with the preprocessed voxel wise brain recordings over time for each subject. For PLS, the correlation matrix were calculated between the entropy timeline and the individual voxels time-series for all the subjects. Singular value decomposition (SVD) was employed to decompose this correlation matrix into correlation strength and brain salience as in (Garrett et al., 2010). Brain salience for each voxel denotes its weight in the correlation between entropy and brain activity. This multivariate approach allowed us to detect patterns in voxel wise contributions to the entropy time-course per subject. Group analysis was done by transforming these resulting maps into z-score maps and then performing the group analysis using the t-test from SPM.

## **3 Results**

### **Generalized linear model (GLM)**

First, we investigated if our results replicated the findings from the previous experiments (Fraedrich et al., 2010, 2012). Additionally, we investigated if the three types of videos are represented differently in brain and if yes, in which regions. To address this, we performed generalized

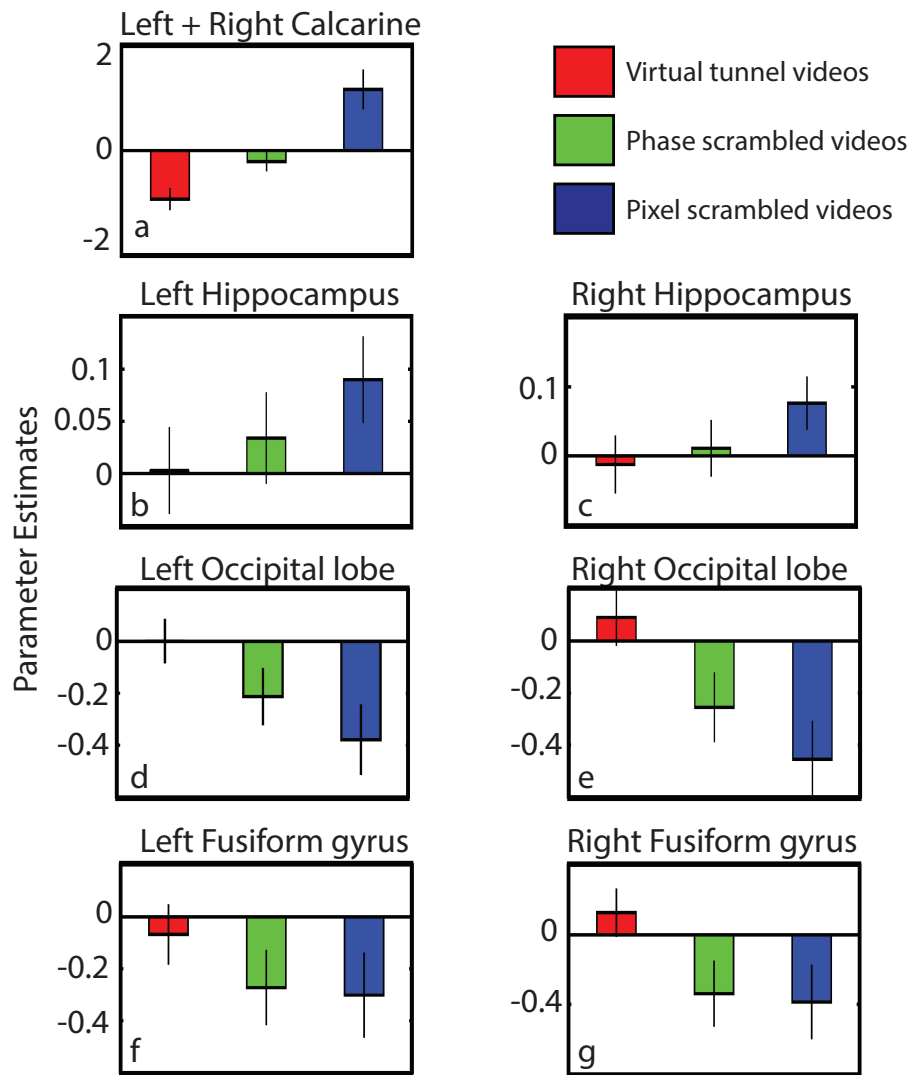
linear modeling to create two contrasts: virtual tunnel videos vs other videos and white noise videos vs other videos. We made three comparisons as follows:



**Figure 3:** Generalized Linear Modeling results of the fMRI data while the subjects watched the stimulus videos. a. group activity map for pixel scrambled videos compared to watching both other types of videos ( $T=3.19$ ,  $p<0.05$ , FDR corrected,  $df=26$ ). We found significance in positive contrast only (pixel scrambled videos>all other videos) b. group activity map for virtual tunnel videos compared to other types of videos ( $T= 3.33$ ,  $p<0.05$ , FDR corrected,  $df=26$ ).

**a. Virtual tunnel videos versus phase scrambled videos:** On the basis of our previous experiments, we formed our hypothesis that early visual areas and the hippocampal region would react more strongly for the phase scrambled videos as compared to the virtual tunnel videos. Our GLM results confirmed this stronger involvement of these brain regions for phase scrambled videos ( $p<0.05$ , FDR corrected,  $n=27$ ) and hence we were able to replicate our previous results. We found stronger activity in the left hippocampus (peak:  $x= -24$ ,  $y= -27$ ,  $z= -6$ , cluster size= 6 voxels,  $T= 4.55$ ,  $p<0.05$ , FDR corrected) for the phase scrambled videos than that for the virtual tunnel videos, but failed to find significant activation for the right hippocampus. Similar significant positive activations were found in the bilateral early visual cortex (peak:  $x= -6$ ,  $y= -96$ ,  $z= -6$ , cluster size=1210 voxels,  $T= 8.92$ ,  $p<0.05$ , FDR corrected). On the other hand, the fusiform gyrus and the parahippocampus (left peak:  $x= -24$ ,  $y= -48$ ,  $z= -18$ ,  $T=5.86$ ; right peak:  $x=27$ ,  $y=-42$ ,  $z= -12$ ,  $T=6.14$ ,  $p<0.05$ , FDR corrected), showed higher activity in response to the virtual





**Figure 4:** Parameter estimates of the clusters with significant activity in group analysis comparing the three stimulus types. a. left hippocampus, b. right hippocampus, c. left occipital lobe, d. right occipital lobe, e. left fusiform gyrus, f. right fusiform gyrus, and g. bilateral calcarine. The detailed description of individual clusters is shown in Table 1.

tunnel videos than the phase scrambled videos. Bilateral dorsal stream

**Table 1:** Brain regions from the group GLM analysis for the three contrasts: pixel scrambled>other videos, virtual tunnel>other videos and other videos>virtual tunnel videos with t-test (df= 26, p>0.05, FDR corrected). All the significant clusters with size>15 voxels are presented here with their anatomical labels.

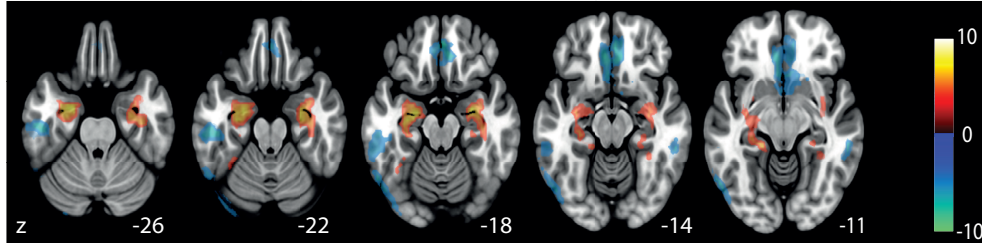
Region	Hemi-sphere	Cluster size (voxels)	Peak coordinates x,y,z (MNI)	Max t
<b>pixel scrambled videos &gt; other videos</b>				
Calcarine	L, R	777	-6, -99, -6	6.42
Hippocampus	L	38	-21, -27, -6	5.30
Postcentral + Temporal superior + Rolandic Operculum	L	272	-63, -21, 15	8.40
Hippocampus + Thalamus	R	19	24, -27, -3	4.42
Rolandic Operculum	R	83	48, 0, 9	4.99
Temporal superior	R	29	54, -30, 9	5.11
Supramarginal	R	16	66, -21, 27	3.85
Cingulum mid	R	34	9, 12, 36	4.21
Supramarginal	L	62	-51, -24, 42	4.40
Precentral + Postcentral	R	40	51, -12, 54	4.79
<b>virtual tunnel videos &gt; other videos</b>				
Fusiform + Lingual + Parahippocampus	L	97	-24, -51, -12	5.66
Fusiform + Lingual + Parahippocampus	R	138	27, -42, -12	5.70
Temporal inferior	R	20	48, -72, -6	3.71
Occipital mid/sup + Parietal sup + Precuneus	L	482	-27, -84, 12	5.81
Occipital mid/sup + Parietal sup + Precuneus	R	748	30, -81, 27	5.41
Precuneus	L	17	-9, -48, 45	3.45
Occipital Inf	L	15	-45, -69, -9	3.92
<b>other videos &gt; virtual tunnel videos</b>				
Calcarine + Lingual + Occipital mid/inf/sup	L + R	1080	-6, -99, -6	7.68
Hippocampus	L	17	-27, -30, -6	5.02

(left peak: x= -21, y= -84, z= 39, T= 5.22; right peak: x= 33, y= -78, z= 27, T =5.83, p<0.05, FDR corrected) also showed stronger activity for the tunnel videos, compared to the phase scrambled videos. All these results are in line with our previous work (Fraedrich et al., 2010, 2012).

**b. Pixel scrambled videos versus others:** We extended our hypothesis from previous contrast that the early vision areas might react more strongly for the pixel scrambled videos as compared to both other video types due to higher unpredictability in their structure. As shown in Figure 3a, the early visual areas were indeed elicited more strongly ( $p < 0.05$  FDR corrected,  $n=27$ ) for pixel scrambled videos. Figure 4a shows the parameter estimates of bilateral early visual areas for all the types of stimulus we presented. As hypothesised, higher values of parameter estimates in these areas indicate the strongest activity for pixel scrambled videos. Apart from the early visual areas we found, bilateral hippocampal regions also showed stronger activity in response to the pixel scrambled videos than that for the phase scrambled and the virtual tunnel videos together as shown in Figure 3a. Parameter estimates of the clusters comprising of the left and right hippocampus also show relatively higher activity for the pixel scrambled videos than others, as shown in Figures 4b and 4c respectively. Apart from the regions we reported before, higher activity was also found in the insula, superior temporal gyrus and superior temporal lobe for the pixel scrambled videos. Small clusters of the right cingulate gyrus and bilateral pre/post-central gyrus also showed a similar response to the pixel scrambled videos (Table 1). We did not find any significant effect for the negative contrast (other two videos versus the pixel scrambled videos).

**c. Virtual tunnel videos versus others:** As expected, the virtual tunnel videos had lower activity for the early visual regions and the left hippocampus compared to all other videos. Similarly, the bilateral dorsal stream, fusiform gyrus and parahippocampal gyrus were more strongly activated for the virtual tunnel videos than others ( $p < 0.05$ , FDR corrected,  $n=27$ ) as shown in Figure 3b. Corresponding parameter estimates of the bilateral dorsal stream (occipital lobe) and fusiform gyrus are shown in Figure 4d, 4e, 4f and 4g. Significant clusters, with cluster size  $> 15$ , are shown in Table 1. Regions like the parahippocampal gyrus and fusiform gyrus, which were strongly activated for virtual tunnel videos compared to the other non structured videos, were expected as the parahippocampus has been reported earlier to be involved in visu-

ospatial perception (Epstein and Kanwisher, 1998; Ekstrom et al., 2003; Rajimehr et al., 2011) and navigation (Aguirre et al., 1996; Maguire et al., 1997; Aminoff et al., 2013).



**Figure 5:** Group results for Partial Least Squares analysis. Activation map for positive and negative significant weights correlating to the entropy vector are shown in the figure ( $T=8.4$ ,  $p<0.001$ , FDR corrected,  $df=26$ ). Positive weights represent the regions that correlated positively with the entropy vector and vice versa.

**Table 2:** Brain regions from the PLS group results for the positive and negative contributions to the entropy vector (t-test, FDR corrected,  $df=26$ ,  $p<0.001$ ). All significant clusters with cluster size  $> 15$  voxels are shown here with their anatomical labels.

Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
<b>Positive</b>				
Hippocampus + Parahippocampus + Fusiform+ Inferior Temporal Gyrus	L	360	-33, -3, -30	8.90
Hippocampus + Parahippocampus + Fusiform + Inferior Temporal Gyrus	R	213	42, -6, -39	8.42
<b>Negative</b>				
Anterior Cingulate	L, R	405	12, 24, -6	-8.59
Inferior Temporal Gyrus	L	127	-51, -21, 24	-8.82
Inferior + Middle Tem- poral Gyrus	R	26	57, -36, -12	-10.27
Inferior Occipital Gyrus	L	127	-39,-84,21	-10.10
Superior Occipital Gyrus	L	112	-24, -81, 30	-7.23
Superior Frontal Gyrus	R	22	21, 45, 48	-6.35
Precuneus	L	32	-6, -63, 48	-7.52

### **Multivariate: Partial Least Squares Correlation analysis**

To further explain the effects of the spatiotemporal entropy on the brain, we compared the entropy timeline with the voxel time courses in a multivariate manner. We found significant voxel weights ( $p < 0.001$ , FDR corrected,  $df = 26$ ) expressing the correlation between the brain activity and the entropy timeline. The left hippocampus, parahippocampus, fusiform and inferior temporal gyrus were found to contribute positive weights towards the entropy vector (peak:  $x = -33$ ,  $y = -3$ ,  $z = -30$ , cluster size = 360 voxels, max  $T = 8.89$ ). A similar cluster was found in the right hemisphere (peak:  $x = 42$ ,  $y = -6$ ,  $z = -39$ , cluster size = 213 voxels, max  $T = 8.42$ ). These regions were prominently more active, for the pixel scrambled and the phase scrambled videos than for the virtual tunnel videos in GLM results as well. For negative weights, we found that the bilateral anterior cingulate (peak:  $x = 12$ ,  $y = 24$ ,  $z = -6$ , cluster size = 405 voxels, max  $T = -8.59$ ) and the bilateral inferior temporal gyrus contributed the most. Several smaller clusters from the occipital, parietal and frontal lobe were also found to be significant, as shown in Figure 5. Anatomical labels of all the significant clusters with cluster size greater than 15 voxels are shown in Table 2.

## **4 Discussion**

Our main result is the stronger hippocampal activity for the pixel scrambled videos when compared to the more structured virtual tunnel and phase scrambled videos. Our findings from both types of analysis thus indicate that it is the statistical structure of the visual input that determines involvement of the hippocampus rather than the indistinct meaning of the phase scrambled videos.

Interestingly, the hippocampus has not been reported to be active for static white noise images (Rainer et al., 2001; Wichmann et al., 2006). This indicates that the temporal unfolding of the stimulus adds a layer of uncertainty in the temporal domain and plays an important role in determination of hippocampal involvement in visual processing.

Unpredictability in sequences of shapes and numbers (Strange et al.,

2005; Harrison et al., 2006), images (Schapiro et al., 2012), some previously unknown symbols (Turk-Browne et al., 2009) and everyday action videos, such as making sandwiches (Schiffer et al., 2012) have been studied before. These studies established stronger hippocampal involvement for unpredictable violation of a known model than for completely novel stimuli. To the best of our knowledge, the effect of varying spatiotemporal entropy of the visual stimuli on brain activation has not been assessed before. We hypothesized that the hippocampus is recruited when the visual input stream has high entropy even if the content is meaningless or does not reveal any clear meaning. Previously, Turk-Browne et al. (2009, 2010) demonstrated hippocampal activity for statistical learning of the relationship between different symbols without subjects being aware of any ongoing learning. Other experiments used behavioral tasks that implicitly needed learning of statistical regularities for better performance (Strange et al., 2005; Harrison et al., 2006). In our experiment, subjects were asked to report detection of arrows in the videos, making the task cognitively independent of the semantic content of the videos itself and hence not involving any conscious effort of regularity/structure extraction.

### **Hippocampus: The communication channel between perception and learning?**

Stimulus-bound surprise information is reflected in the cortical activity, and the regions that react vigorously to surprise could be associated with the prediction error (Strange et al., 2005). Our study confirms that the hippocampus is sensitive to the probabilistic context of the visual stream and hence can be in a position to balance the top-down and bottom-up sensory cortical hierarchy. Adaptation, a sign of learning, in visual input has been shown with reduction in hippocampal activity over time and repetitions (Schiffer et al., 2012). In our experiment, the hippocampal system could adapt during the virtual tunnel videos but not during the pixel scrambled videos due to their high unpredictability. This leads to reduction in activity of virtual tunnel videos but not for pixel scrambled videos. The hippocampus has recently been shown to be selectively “switched on” for unpredictable input and to act as a switch in the transition between the two learning types: incremental and one-

shot learning (Lee et al., 2015). In incremental learning, the relationship between the incoming information and the outcomes is learnt in an incremental manner, over multiple instances. One-shot learning, on the other hand, implies learning at a much faster rate, sometimes even with just one instance of the relationship. For behavioral advantages, one-shot learning is employed for highly unpredictable stimuli. This ability to learn is not related to novelty but rather to the presence of irregularity in the input stream (Strange et al., 1999; Schiffer et al., 2012). We think that an interesting analogue exists between the perceptual streams and the learning schemes. For an uncertain input, one-shot learning is required and hence resources are allocated to place more emphasis on the bottom-up perceptual stream. On the other hand, when input regularities are known, the predicted input is propagated through the top-down stream which is then compared to the actual input. Any mismatch leads to incremental learning and updating of the previous model. Both of these processes occur in parallel in the brain and need similar statistical information. The hippocampus is the perfect candidate to reflect this statistical information in the brain. It can also communicate this information to other cognitive processes due to its strong anatomical and functional connectivity with different cortical and subcortical regions.

### **Buffer theory**

The brain continuously receives a lot of information. Often, the incoming inputs have to be integrated over time (in a “buffer”) to extract the meaningful information. Whether to discard the information from this buffer or to consolidate it to long-term memory depends upon the behavioral constraints. Baddeley (2000) extended his working memory model by including an episodic memory buffer. This buffer temporarily stores the information in relational codes. We propose the hippocampus as a primary candidate for this buffer because it can extract relationships from the incoming information and communicate well with cortical regions. Berlingeri et al. (2008) also suggested that the anterior hippocampus could be the brain region that functions as the episodic memory buffer. Since the hippocampus is at the top of the ventral visual-perirhinal-hippocampal processing stream (Bussey and Saksida, 2007), it is well connected to cortical regions and can consolidate the traces (in-

coming information) if needed. In our experiment, due to the complex and unpredictable nature of the noise videos, the hippocampus might be required to store (or continuously refresh) the non-structured visual information in this 'buffer', resulting in stronger BOLD activity in the hippocampus.

## **Conclusion**

We analyzed the two possibilities of hippocampal involvement while processing visual stream inputs. Hippocampal activity was found to be stronger while watching the pixel scrambled videos compared to the other more predictable videos types. Since phase scrambled videos did not exhibit stronger hippocampal activations than pixel scrambled videos, we ruled out the possibility of 'meaning-decoding' as the requirement for hippocampal activity. Instead, our analyses show that hippocampal activity increases with higher unpredictability of the visual stream, irrespective of its spatial or semantic structure.

Our study demonstrated strong hippocampal activation when processing dynamic visual input with high entropy but low complexity. The hippocampal response was observed even though no explicit instruction or need to process the statistical structure of the input stream was given. Why the hippocampus responds to this statistical information and which other brain regions are involved in the process are still open questions.

## **Acknowledgments**

Authors of this manuscript would like to thank the Department of Neuroradiology of the University Hospital, LMU Munich for their diagnostic support. They would also like to extend their thanks to DFG (GRK 1091, GSN) and BMBF (BCCN 01GQ0440, DSGZ 01EO0901) for funding this study.



## References

- Aguirre, G. K., Detre, J. A., Alsup, D. C., and D'Esposito, M. (1996). The parahippocampus subserves topographical learning in man. *Cerebral Cortex (New York, N.Y.: 1991)*, 6(6):823–829.
- Ahlheim, C., Stadler, W., and Schubotz, R. I. (2014). Dissociating dynamic probability and predictability in observed actions—an fMRI study. *Frontiers in Human Neuroscience*, 8.
- Aminoff, E. M., Kveraga, K., and Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences*, 17(8):379–390.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, 4(11):417–423.
- Berlingeri, M., Bottini, G., Basilico, S., Silani, G., Zanardi, G., Sberna, M., Colombo, N., Sterzi, R., Scialfa, G., and Paulesu, E. (2008). Anatomy of the episodic buffer: a voxel-based morphometry study in patients with dementia. *Behavioural neurology*, 19(1-2):29–34.
- Bornstein, A. M. and Daw, N. D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning: Sequential predictions in hippocampus and striatum. *European Journal of Neuroscience*, 35(7):1011–1023.
- Brown, M. W. and Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, 2(1):51–61.
- Bussey, T. and Saksida, L. (2007). Memory, perception, and the ventral visual-perirhinal-hippocampal stream: Thinking outside of the boxes. *Hippocampus*, 17(9):898–908.
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., and Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature*, 425(6954):184–188.
- Epstein, R. and Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676):598–601.

- Fraedrich, E. M., Flanagin, V. L., Duann, J.-R., Brandt, T., and Glasauer, S. (2012). Hippocampal involvement in processing of indistinct visual motion stimuli. *Journal of cognitive neuroscience*, 24(6):1344–1357.
- Fraedrich, E. M., Glasauer, S., and Flanagin, V. L. (2010). Spatiotemporal phase-scrambling increases visual cortex activity. *NeuroReport*, 21(8).
- Garrett, D. D., Kovacevic, N., McIntosh, A. R., and Grady, C. L. (2010). Blood oxygen level-dependent signal variability is more than just noise. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(14):4914–4921.
- Graham, K. S., Barense, M. D., and Lee, A. C. H. (2010). Going beyond LTM in the MTL: a synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia*, 48(4):831–853.
- Harrison, L., Duggins, A., and Friston, K. (2006). Encoding uncertainty in the hippocampus. *Neural Networks*, 19(5):535–546.
- Krishnan, A., Williams, L. J., McIntosh, A. R., and Abdi, H. (2011). Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *NeuroImage*, 56(2):455–475.
- Kumaran, D. and Maguire, E. A. (2007). Which computational mechanisms operate in the hippocampus during novelty detection? *Hippocampus*, 17(9):735–748.
- Maguire, E. A., Frackowiak, R. S., and Frith, C. D. (1997). Recalling routes around London: activation of the right hippocampus in taxi drivers. *Journal of neuroscience*, 17(18):7103–7110.
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., and Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3):1233–1239.
- McIntosh, A. R., Bookstein, F. L., Haxby, J. V., and Grady, C. L. (1996). Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage*, 3(3 Pt 1):143–157.

- Rainer, G., Augath, M., Trinath, T., and Logothetis, N. K. (2001). Non-monotonic noise tuning of BOLD fMRI signal to natural images in the visual cortex of the anesthetized monkey. *Current Biology*, 11(11):846–854.
- Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., and Tootell, R. B. H. (2011). The “Parahippocampal Place Area” Responds Preferentially to High Spatial Frequencies in Humans and Monkeys. *PLoS Biology*, 9(4):e1000608.
- Schapiro, A., Kustner, L., and Turk-Browne, N. (2012). Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Current Biology*, 22(17):1622–1627.
- Schiffer, A.-M., Ahlheim, C., Wurm, M. F., and Schubotz, R. I. (2012). Surprised at All the Entropy: Hippocampal, Caudate and Mid-brain Contributions to Learning from Prediction Errors. *PLoS ONE*, 7(5):e36445.
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, 275(5306):1593–1599.
- Schultz, W. and Dickinson, A. (2000). Neuronal Coding of Prediction Errors. *Annual Review of Neuroscience*, 23(1):473–500.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–155.
- Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., and Friston, K. J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Networks*, 18(3):225–230.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., and Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of cognitive neuroscience*, 21(10):1934–1945.
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., and Chun, M. M. (2010). Implicit Perceptual Anticipation Triggered by Statistical Learning. *Journal of Neuroscience*, 30(33):11177–11187.

- Wichmann, F. A., Braun, D. I., and Gegenfurtner, K. R. (2006). Phase noise and the classification of natural images. *Vision Research*, 46(8-9):1520–1529.
- Yee, L. T., Warren, D. E., Voss, J. L., Duff, M. C., Tranel, D., and Cohen, N. J. (2014). The hippocampus uses information just encountered to guide efficient ongoing behavior: Hippocampus Guides Efficient Behavior. *Hippocampus*, 24(2):154–164.
- Zhao, J., Ngo, N., McKendrick, R., and Turk-Browne, N. B. (2011). Mutual Interference Between Statistical Summary Perception and Statistical Learning. *Psychological Science*, 22(9):1212–1219.



## 2.2 Dual cortico-hippocampal network activation while watching high-entropy visual stimuli

**Authors:** Nisha Dalal, Virginia L Flanagin, Stefan Glasauer

### Contributions

Nisha Dalal designed and programmed the experiment, collected and analyzed data, created figures, interpreted the results and wrote the manuscript.

Virginia Flanagin supervised the project, helped in interpreting results and critically reviewed the manuscript.

Stefan Glasauer supervised the project, helped in designing the experiment and in interpreting results and critically reviewed the manuscript.



# Dual cortico-hippocampal network activation while watching high-entropy visual stimuli

Nisha Dalal <sup>1,2</sup>, Virginia L Flanagin <sup>1,2,3</sup>, Stefan Glasauer <sup>1,2,3</sup>

<sup>1</sup> Department of Neurology, LMU, Munich, Germany

<sup>2</sup> Graduate School of Systemic Neurosciences, LMU, Munich, Germany

<sup>3</sup> German Center for Vertigo and Balance Disorders, LMU, Munich, Germany

**Keywords:** Hippocampus, Cortico-hippocampal networks, fMRI, Independent component analysis, Granger causality





## Abstract

The hippocampus, in conjunction with other subcortical and cortical regions, responds to the different measures of uncertainty. In our previous studies, we recorded the brain activity (fMRI) from 27 human subjects (age ranging from 19 to 33; mean age  $25.6 \pm 3.8$  years, 14 females) while they watched dynamic visual stimulus with varying spatial and spatiotemporal entropies. Three types of videos were used: the virtual tunnel videos (mimicking the motion inside a tunnel), corresponding phase scrambled videos (created by scrambling the phase of the virtual tunnel videos while keeping the same amplitude spectrum) and the pixel scrambled videos (scrambling the pixel values of the virtual tunnel videos). Virtual tunnel videos, being the most structured, had the lowest entropy. The pixel scrambled videos, with no spatial or temporal relationship between neighbouring pixels, had the highest entropy. We found stronger activity in the hippocampus and the early visual cortex for pixel scrambled videos when compared to both other types of videos. In this manuscript, we present a follow-up study uncovering the underlying networks and information flow between those networks using data-driven methods. Forty-nine independent components were found using independent component analysis and five of them were relevant for our experimental stimulus. We found two parallel cortico-hippocampal networks functioning at different timescales. While one of the those networks was activated specifically to a particular video type, the other one presented a more general visual processing network. We further investigated the temporal relationships between the independent networks using Granger causality and found a hierarchical system where the general visual processing network 'influences' the stimulus specific networks. This manuscript presents an interesting flow of information between the independent components where a stimulus-nonspecific component drives the other stimulus specific networks.



# 1 Introduction

We perceive the world in order to act upon it. The visual system enables us to see the world, yet the visual system (as a whole) is not interested in the image on the retina, but the environment which caused that image. In other words, the purpose of the visual system is to extract meaningful spatial and temporal relationships from visual inputs (sequences of retinal image codings) in relation to experience.

Information from the past contributes actively to comprehension by continually influencing the information processing in the present. Hasson et al. (2015) proposed a memory system that accumulates information hierarchically at all cortical levels but at different timescales. These processing timescales could vary from a few milliseconds (for early sensory areas) to seconds to minutes (for higher-order areas). The processing timescale for the hippocampus was unclear in the proposed model due to its involvement at different timescales in different studies. While Riggs et al. (2009) showed an early involvement of the hippocampus (<150 ms after stimulus onset) in the processing stream, studies focussing on encoding and retrieval of episodic memories hint at processing at much longer timescales (Hannula et al., 2006; Olson, 2006).

We have previously shown that the hippocampus is involved in the information processing of the uncertain visual stimuli (Fraedrich et al. 2012). When watching movement through a relatively predictable virtual tunnel, the hippocampus is much less active than when the phase-components of this video are scrambled. One possible reason for this is the increased spatiotemporal “surprise” or entropy in the phase-scrambled films. In manuscript 1, we examined the brain activity in response to high entropy visual stimulus. We adapted the virtual tunnel and corresponding phase scrambled videos from Fraedrich et al. (2012) and extended the stimulus set further by adding corresponding pixel scrambled videos with high entropy. We found stronger activity in the higher visual regions including the parahippocampus and the dorsal visual stream for the virtual tunnel videos compared to both other types of videos. On the other hand, for the pixel scrambled high entropy videos, we found stronger early visual cortex and hippocampal activity.

These findings raised further questions about the underlying networks and processes. For example, are the hippocampus and early visual cortical areas part of the same functional network while processing such high entropy visual stimulus? At what timescales do these networks work? Could the hippocampus be a part of multiple networks and have different functions in processing the incoming visual information? What other networks are involved in processing the virtual tunnel videos and the corresponding high entropy videos? How do these networks influence each other? Information from the previous analysis was insufficient to understand the underlying process and answer these questions.

In this manuscript, we present the follow-up study performed to inspect the cortical-hippocampal networks involved. We used a data-driven approach using analysis methods such as Independent Component Analysis (ICA) for our fMRI data to uncover the underlying networks. We further examined the functional connectivity between these networks to uncover the causal dynamics between them.

## 2 Methods

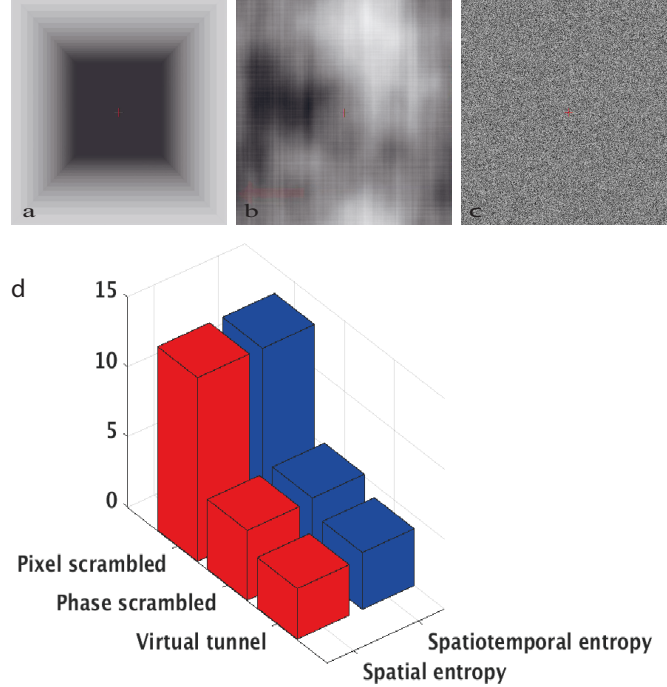
### 2.1 Stimuli and task

The stimulus set, task and fMRI data used in this manuscript have also been used and described in manuscript 1.

#### **Stimuli**

To investigate the hippocampal involvement in processing visual stimuli with different uncertainties, three types of videos were used: virtual tunnel videos, phase scrambled videos and pixel scrambled videos. Expanding on the stimuli from Fraedrich et al. (2010, 2012), 12 second long virtual tunnel videos were created using the psychotunnel toolbox implemented in OpenGL (Gramann et al., 2010). Each video was 12 seconds long and consisted of 360 frames of 540\*539 pixel size (30 Hz sampling rate). These videos, simulating a virtual motion through a tunnel, consisted of three segments: a straight segment, a left or a right turn followed

by a straight segment again. Six turn angles (34, 45, 75, 85, 115, 125 degrees) were used for both left and right turns.



**Figure 1:** Frames from the three types of stimulus videos and their entropies. a. virtual tunnel video frame, b. phase scrambled video frame, c. pixel scrambled video frame, d. spatial and spatiotemporal entropies of the three types of videos, computed in manuscript 1. Frame from the phase scrambled (b) shows an example of a red arrow on the bottom left corner from the arrow detection task that the participants performed. Each such arrow was presented for 10 frames. Pixel scrambled videos have the highest entropy while the well-structured tunnel videos have the lowest entropy. Detailed description of the spatial and spatiotemporal entropy calculations can be found in the method section of manuscript 1. Figure adapted from manuscript 1.

To generate the phase scrambled videos, virtual tunnel videos were transformed into three dimensional Fourier space and their phase was scrambled along all the three dimensions. These matrices were then transformed back to the time domain without changing the amplitude spectrum of the videos. The resultant videos were used as phase scrambled videos. They had the same amplitude as the corresponding virtual tunnel videos but did not feature anything recognizable in the video content itself. These two types of videos had the same local flow, average contrast and luminance. The local luminance difference over frames was also

matched. Pixel scrambled videos were created by scrambling pixels from the virtual tunnel videos randomly. These videos were used as the pixel scrambled videos. Spatial and spatiotemporal entropy calculations of the videos have been explained in detail in the method section of manuscript 1. Figure 1 shows an example of the video frames and their entropies.

During previous experiments (Fraedrich et al., 2010, 2012), participants could associated some 'meaning' to the phase scrambled videos due to their slow, but coherent, motion. Although inconsistent across participants, an ongoing meaning-association process was used to interpret these phase scrambled videos, even in the absence of any explicit instruction or necessity to do so. Virtual tunnel and pixel scrambled videos, on the other hand, could reveal their semantic content (or the lack of any semantic content) from the first few frames without any such ongoing decoding process. Virtual tunnel and phase scrambled videos had low entropy in spatial and spatiotemporal dimensions while pixel scrambled videos had high spatial and spatiotemporal entropies. This dichotomy between meaning and entropy allowed us to investigate the hippocampal response to uncertainty and ongoing meaning-decoding process.

### **Task and training**

We designed the task to engage the participants with the incoming visual stream while distracting them from interpreting the stimulus. Participants performed a search-task for monochrome objects within the visual scene. Similar to the second task in Fraedrich et al. (2012), red arrows were added to the videos for ten frames (0.33 seconds) each. Deviating from the previous study, each of the videos with arrows had from one to four of these arrows. Participants also had knowledge that there could be 1-4 arrows but not that only half of the films contained arrows. The goal was for the task to be challenging but to not require participants to remember the presence of arrows, as this may also elicit medial temporal lobe activity. Participants were asked to detect the arrows and respond by pressing a button as soon as they detect the arrows. During the pilot session, it was noted that arrow detection was more difficult in pixel scrambled videos. Hence, the red arrows were made darker than the original color used in the previous experiment (12/255 red component

compared to 10/255 in others) in the pixel scrambled videos to maintain the same difficulty level of the task across different video types. Subjects were trained in two training sessions (seven to eight minutes each) just before the experiment. An answer was only considered correct if they could detect all the arrows in the videos. They were informed about the task outcome after every video. Subjects with at least 60 percent accuracy were allowed to perform the task inside the fMRI scanner.

## 2.2 fMRI data acquisition and preprocessing

### Participants

Thirty-one healthy participants (age ranging from 19 to 33; mean age  $25.4 \pm 3.7$  years, 16 females) with normal or corrected-to-normal vision and no documented history of neurological or psychiatric disease gave their informed consent to participate in the study. The experiment was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee of the medical faculty at the Ludwig-Maximilians University, Munich. Participants with any metallic part in their body were not allowed to perform the experiment. The functional session of one participant was interrupted three times due to large head movements and hence the data corresponding to that participant could not be recorded and used. Functional data from three additional participants had to be discarded due to excessive head movements even after preprocessing, which left a total of 27 subjects (age ranging from 19 to 33; mean age  $25.6 \pm 3.8$  years, 14 females) for further analysis.

### Data acquisition

A 3T whole-body MR scanner (Signa HDx, GE healthcare) with an eight-channel head coil was used to record the functional data with a T2\*- weighted gradient-echo, echo-planar sequence (TR=2 sec, TE=30 ms, FOV=220 mm, matrix=64\*64, voxel size=3.28\*3.28\*3.49 mm). The whole brain was scanned with 38 axial slices with a slice thickness of 3.5 mm with no inter-slice gap. The first four functional scans were discarded at acquisition level to remove the T1 effects. Padding and adjustable head restraints were used to minimize the head motion. Apart from the



functional scans, a high resolution (voxel size  $0.85 \times 0.85 \times 0.7$  mm) whole brain fast spoiled gradient echo (FSPGR) anatomical image was also acquired from each participant.

### **Preprocessing**

To prepare the fMRI data for preprocessing, functional and anatomical brain images from all the participants were skull-stripped from the skeleton using a fractional intensity threshold of 0.5 in the Brain Extraction Toolbox (BET) (Smith, 2002) from FSL (FMRIB Software Library). Further preprocessing steps were performed using SPM12 v6909 (Wellcome Trust Centre for Neuroimaging, UCL, London, UK). For alignment between the two functional scans, volumes from both sessions were aligned with 0.9 quality using a 2<sup>nd</sup> degree b-spline interpolation with the mean image as a reference scan. Functional scans were further corrected and resliced according to the estimated head motion regressors using a six parameter rigid body transformation and 4<sup>th</sup> degree b-spline interpolation. These aligned volumes were then coregistered with the high resolution anatomical image of the participant by maximizing the normalized mutual information. Using tissue probability maps, anatomical images were further segmented into the white matter, gray matter and cerebrospinal fluid using a mutual information affine registration to ICBM template. Nonlinear image registration was performed by iteratively matching the gray matter and white matter images of all subjects to the template mean image generated from those images using the DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra) tool. These registrations were penalized by linear elastic energy using 0.01 Levenberg-Marquardt regularization in three cycles and three iterations. The two functional scans and the gray matter images, now matched to the average template, were then normalized to the MNI template using the DARTEL flow field created above for each subject by preserving concentration. The resulting functional images were then smoothed using a gaussian smoothing kernel (4 mm full width/half-maximum isotropic) to reduce the aliasing effects.

## 2.3 Independent Component Analysis (ICA)

We used ICA as a data-driven approach to uncover the networks involved in processing the incoming visual information with high uncertainties. The Group ICA of fMRI Toolbox (GIFT v3.0b)(Calhoun et al., 2001) was used to decompose the preprocessed functional data into independent components. Functional data from each session (439 functional volumes) of each participant was used as a single dataset for the toolbox. Forty-nine independent components were estimated by the toolbox using minimum description length (MDL) criteria (Li et al., 2007). Data reduction, ICA and back reconstruction were performed in a serial manner. The image mean was removed from the functional data per time point before data reduction. The first image from each subject and session was used to create the default mask. For data reduction, Principal Component Analysis (PCA) was done on each functional dataset before performing Group ICA (Beckmann and Smith, 2004). Spatial ICA was performed to estimate components by maximizing independence in space using the FastICA approach (Hyvarinen, 1999). After ICA, the aggregate spatial maps and the results from data reduction were back-constructed to individual subject components using a Group ICA (GICA) back-reconstruction method based on PCA compression and projection (Calhoun et al., 2001). The resulting components and their respective time courses were scaled to their z-scores.

Multiple linear regression and correlation measures were used for sorting the independent components temporally for their relevance with respect to our experimental stimuli. We used the General Linear Model (GLM) matrices from manuscript 1 to create an experimental model that indicated the presence and absence of a particular stimulus type at a certain time point in the experiment. The time course of this model represented the sequence of stimulus classes during the experimental sessions. The component's time courses were correlated with the model's time courses and were sorted based on the R-squared statistic. Strong correlation between the component's time course and experimental model's specific time course, hence, indicated the activation of the component in response to the specific stimulus class.

For identifying the anatomical structures in the GLM analysis, the Harvard-

Oxford Cortical Probability Atlas, the Harvard-Oxford Subcortical Probability Atlas and the Juelich Histological Atlas were used. The hippocampus was segmented into anterior and posterior as per the established convention of the anterior hippocampus corresponding to the hippocampal head, medial to the body and posterior to the tail of the hippocampus. These sub-hippocampal masks were extracted from the AAL Human Atlas using the WFU PickAtlas. Spatial templates from the FastICA fmri toolbox were used in the ICA analysis to label the respective components (Maldjian et al., 2003, 2004).

## **2.4 Functional network connectivity (FNC)**

ICA provides networks that are spatially independent and have distinct time courses; within in a network, temporal structure is shared. We investigated the relationship between these networks using Granger causality (Granger, 1969), with the methods from Functional Network Connectivity (FNC) toolbox (version 2.3a) (Jafri et al., 2008; Demirci et al., 2009). In this implementation, the Granger causality aims to investigate the causal relationships between brain networks using the variance of the residuals in the approximations as a measure of causality.

Prior to calculating dependencies, time courses of all selected components were band pass filtered within 0.008 to 0.24 Hz using a Butterworth filter; this frequency spectrum was used for all further causality calculations. These causal dependencies were calculated for both direction relationships between all of the components pairs (20 combinations). We used the Bayesian Information Criterion (BIC) (Schwarz, 1978) to select the best model to fit the data on the group level.

## **3 Results**

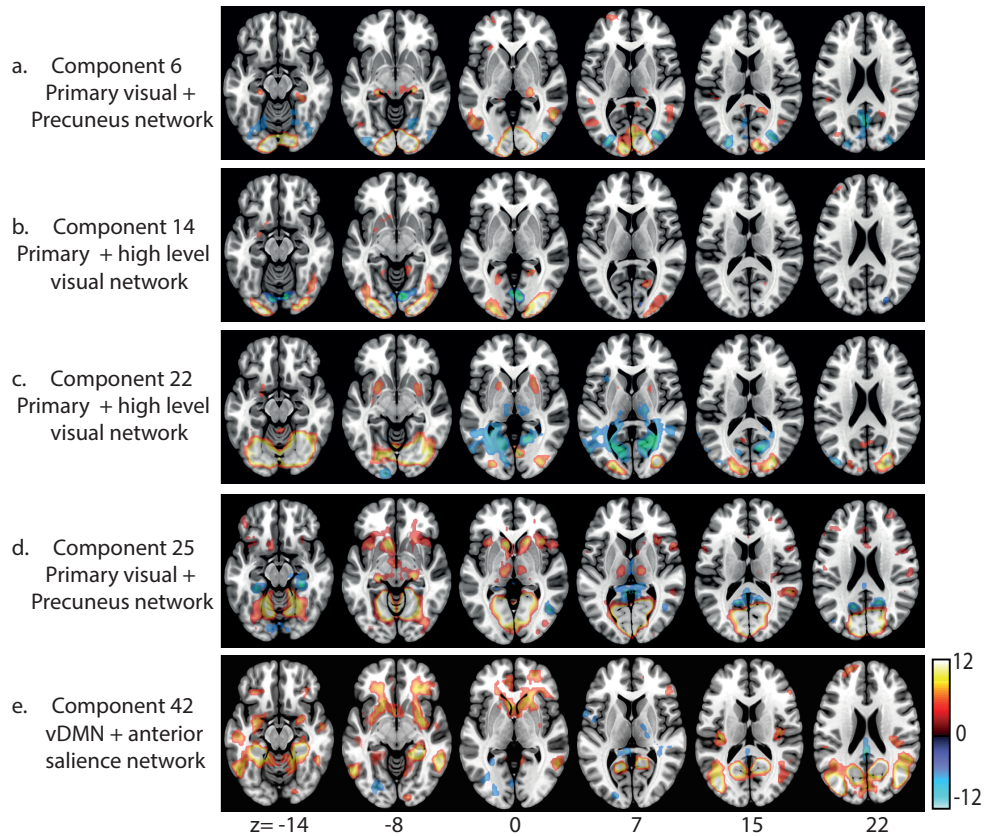
Forty-nine independent components were extracted from the fMRI data. To find the most relevant components, time courses of the components were correlated with the experimental design (onset/offset) of each of the

stimulus using multiple linear regression. Out of all these spatially independent components, only those components were selected for further analysis which could explain more than 10 percent of the data individually. Hence, the top five components that were best predicted by the experimental design (using SPM design matrix) were selected for further analysis. The spatial maps of the five selected components are shown in Figure 2.

Spatial templates of the functional ROIs provided by Shirer et al. (2012) were used to label the underlying brain regions and networks. Each component was manually checked for maximum overlap with the provided templates. Component 6 showed regions from the primary and higher visual regions and some regions from the precuneus network. Component 14 and component 22 were both labelled as high level and primary vision networks. Component 25 was labelled as primary visual and precuneus network whereas component 42 strongly resembled the ventral default mode network (vDMN), precuneus and a few regions from the anterior salience network. Component 6 and 25 were two parallel cortico-hippocampal networks from our fMRI data. The individual anatomical regions within these five components were determined and can be found in supplementary table 1.

### **Component time courses and experimental design**

When correlating the time courses of the individual components with the design matrix, components 6, 14 and 22 were positively correlated with the pixel scrambled videos and negatively correlated with the virtual tunnel videos. Component 42, on the other hand, had the highest correlation with the virtual tunnel videos and negative correlations with the other two video types. Component 25, did not correlate strongly with any individual stimulus type specifically, but explained the variance when all three stimulus were modeled together indicating its activation for all the videos rather than any specific video type. Table 1 shows the multiple regression and the subsequent correlations between the component time courses and the stimulus.



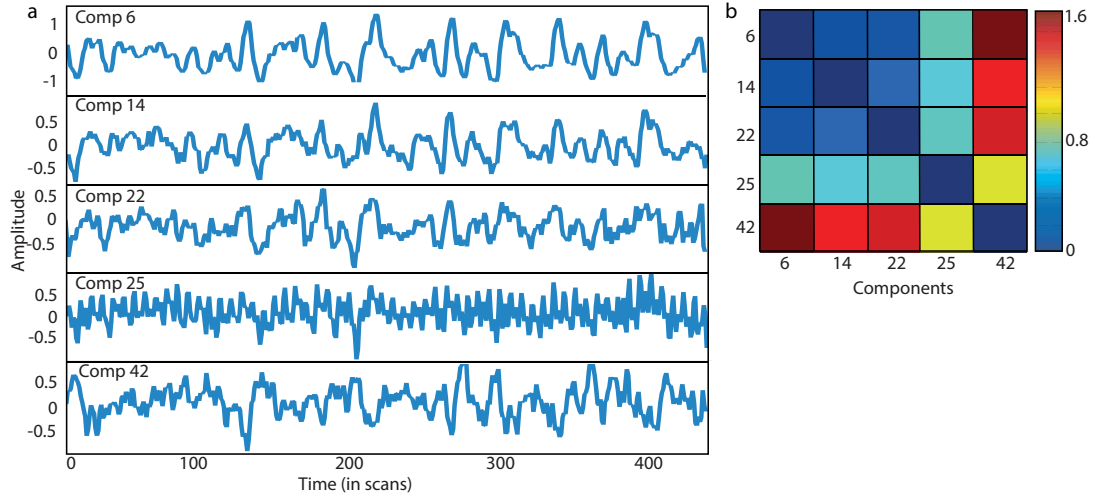
**Figure 2:** The five independent components whose time course was most related to the three experimental stimuli presented. All these spatial maps are based on one sample t-test ( $p < 0.05$ , FWE corrected).

**Table 1:** The regression value ( $R^2$ ) and correlation coefficients between the five components and the three experimental stimuli.

Component	Multiple regression $R^2$ (All stimuli)	Correlation (virtual tunnel videos)	Correlation (phase scrambled videos)	Correlation (pixel scrambled videos)
6	0.363	-0.444	-0.123	0.553
14	0.202	-0.294	-0.133	0.407
22	0.151	-0.302	0.096	0.300
25	0.124	-0.174	0.084	0.159
42	0.104	0.271	-0.129	-0.195

### Temporal correlation between components

The time courses of the components were further analyzed to find similarities between them. Figure 3a shows the mean time courses of the individual components. Figure 3b shows the dissimilarity between the time courses of each component pair calculated by correlation-based

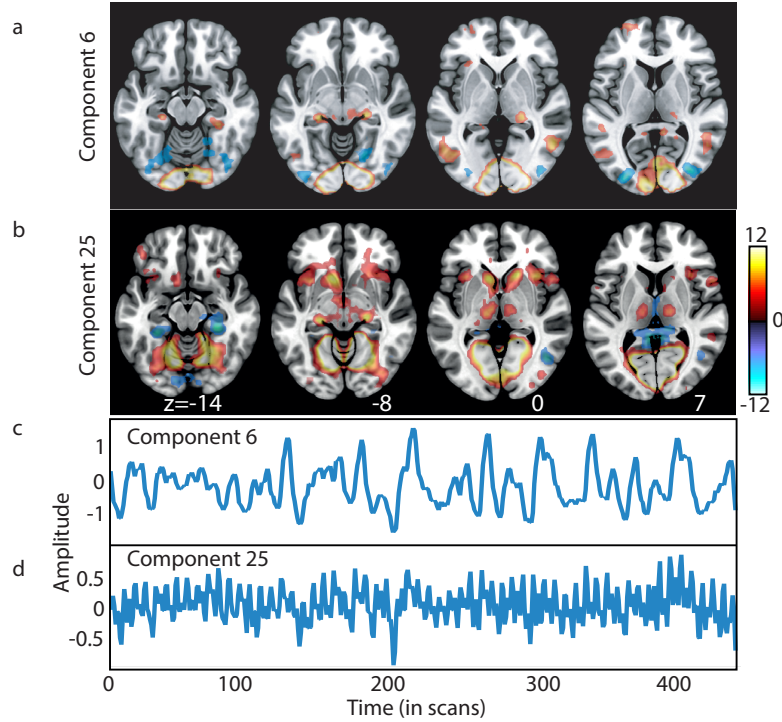


**Figure 3:** Time courses of the top five components and dissimilarities between them. a. time courses of component 6, 14, 22, 25 and 42 averaged over over participants and sessions, b. dissimilarity matrix between the component pairs. Dissimilarities between component’s time courses were calculated pair wise using correlation-based distance (one minus the sample correlation).

distance (one minus the sample correlation). It is worth noting that components 6, 14 and 22 are temporally correlated (and hence similar). This was expected as all these components correlated strongly with the pixel scrambled videos and have high anti-correlation with tunnel videos (as shown in Table 1). Component 42, on the other hand, was positively correlated with the virtual tunnel videos and negatively with pixel scrambled videos (and hence dissimilar to components 6, 14 and 22).

### Dual cortico-hippocampal involvement at different timescales

Two of the components (6 and 25) include the hippocampus in their networks. Component 6, primary visual and precuneus network, is one of the three components with a time course that is positively correlated with the pixel scrambled (high entropy) videos. Component 25, on the other hand, is a combination of the primary and high level visual networks. This component is involved in the general visual processing, independent of the actual video content (or video type). These two separate networks showed differential activities to our visual stimulus. Component 6 was activated primarily for the pixel scrambled videos (and deactivated for the virtual tunnel videos), where as component 25 was activated in

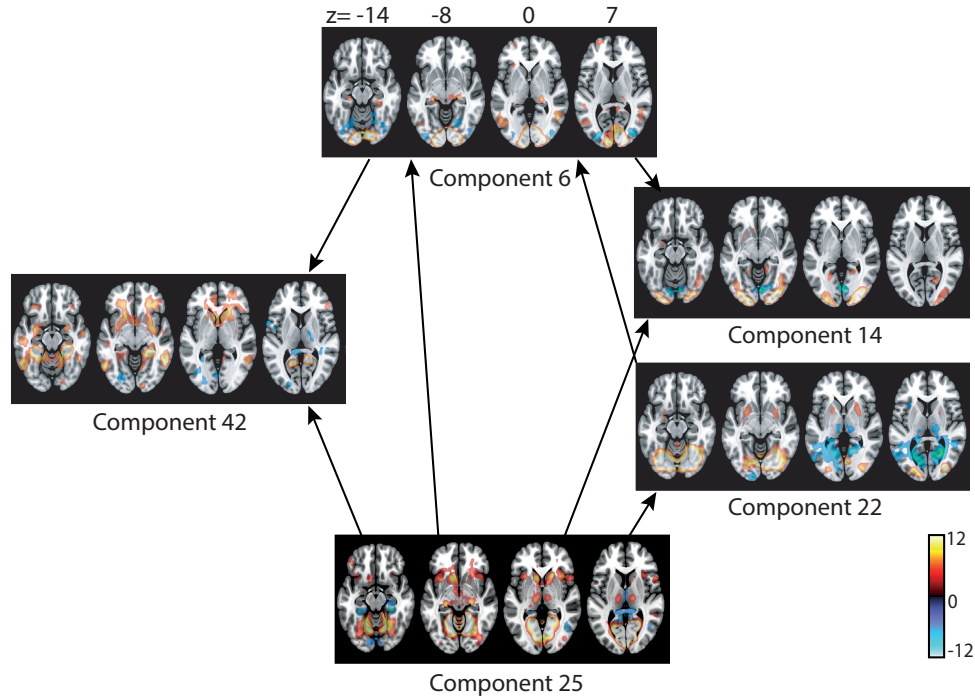


**Figure 4:** Components with a contribution from the hippocampus. a. spatial map of component 6, b. spatial map of component 25, c. time course of component 6 and d. time course of component 25.

response to all the videos, irrespective of its type. These parallel hippocampal networks indicate (at least) a dual hippocampal role while watching these videos.

### G-causal relationship between components

To understand how these independent components influence each other, we calculated the functional connectivity between them. A component X is said to 'G-cause' another component Y if the time course of component X and its past values can predict the time course of component Y better than just with the past values of Y's time course. The dependency graph based on the Granger causality is shown in Figure 5. The direction of the arrows show the direction of the G-causal relationships. For example, the arrow between component 6 and component 25 is directed from component 25 to component 6. Hence, it indicates that component 25, a general visual processing network, 'G-causes' component 6 which is a primary visual and precuneus network.



**Figure 5:** Connections between the five components and their direction. Significant G-causality between the components ( $p < 0.05$ , FDR corrected) is presented here. Direction of each arrow represents the direction of the causal relationship according to the Granger causality.

Component 25 has a G-causal influence on all other selected components. Being a combination of the primary visual system and the precuneus system and being activated for all types of visual stimulus in the experiment, this component represents a general visual processing component. As shown in Figure 5, this component, being stimulus-unspecific itself, further G-causes the stimulus-specific components (component 6, 14, 22 and 42).

Individual frequency profiles for the dependency between the components are shown in the supplementary figure, Figure 6. Component 25 drives the other components at higher frequencies (two dominant and consistent peaks at 0.067 and 0.13 Hz; see Figure 6c, 6e, 6f and 6g). The high frequency peak affects the other components strongly at high frequencies. This component's time course correlates highly with all the stimulus videos together irrespective of the stimulus type (as shown in Table 1) and has a dominant high frequency component in its time course as well (see Figure 3). This high frequency component leads to the



causality relationship between this component and all other components. Similarly, the strong g-causal drive from components 6, 14, 22 and 42 at low frequencies could be attributed to the fact that they have stronger correlations to specific stimulus sets. Hence their g-causal drive is based on the frequency of the specific stimulus type (pixel scrambled videos for components 6, 14 and 22, and virtual tunnel videos for component 42) and not the complete stimulus set.

## 4 Discussion

In the present study, we investigated the brain networks involved in processing dynamic visual stimuli with varying spatiotemporal entropy. In the five components that were related to the stimuli, we found primary and higher-level visual networks, precuneus networks as well as the ventral DMN and the anterior salience network. Multiple networks were found in each component. Three of the networks found were positively correlated with the pixel-scrambled videos, one with the videos of the virtual scene and one with the start of each of the videos. We discovered two cortico-hippocampal functional networks related to our stimulus set, functioning at different timescales. One network, operating on a slower time scale, was related to the uncertain pixel-scrambled videos, the other, operating on a faster time scale, was active at the start of all videos. This two-fold role of the hippocampus in both stimulus-specific and stimulus-unspecific visual streams presents an interesting scenario.

### **Independent components and the connectivity between them**

Component 6 shows the hippocampus and early visual areas positively active and the dorsal stream involved negatively in response to the high entropy videos. Stronger early visual activity for non-structured (and less predictive) stimuli has been suggested to be related to predictive coding before (Fraedrich et al., 2010). As there is no meaningful information in the pixel scrambled videos, negative activations in the dorsal stream might indicate suppression of information flow to the higher cortical regions for further processing. These results are consistent with our

previous findings (Fraedrich et al., 2012). Additionally, this component is deactivated (negatively activated) for virtual tunnel videos. With the GLM (General Linear Model) approach in manuscript 1, we were not able to say if the hippocampus and the early visual cortex were part of two different processes or if there was a single underlying network involving these regions and reacting positively to the high entropy videos. Component 6 here shows that these regions are part of single network. Component 14 and 22 are also activated for the pixel scrambled videos and deactivated for the virtual tunnel videos.

Component 42, on the other hand, is activated specifically for the virtual tunnel videos. It comprises of the ventral DMN and anterior salience networks. This component is deactivated for the pixel scrambled and the phase scrambled videos. As shown in figure 5, component 6 is 'influencing' component 42 and component 14. Activation of component 6 might lead to activation of component 14 whereas the deactivation might lead to activation of component 42. This relationship between these components might reflect how the information about the statistical structure of the incoming visual input is flowing between different networks.

Component 25 shows another cortico-hippocampal network in our data. This component, which is active for all the video types, represents a network involved in general visual processing irrespective of the actual content of the videos. Both these components work at different timescales. Our results are in agreement with (Timme et al., 2014) who also showed similarities and dissimilarities between hippocampal and cortical networks on different timescales. This could explain why Hasson et al. (2015) could not confirm a single timescale for hippocampal processing and found different timescales of hippocampal processing in different studies (Hannula et al., 2006; Olson, 2006; Riggs et al., 2009).

This component showed a G-causal relationship to all other components with a high frequency component. As shown in figure 5, component 25 is 'influencing' all other independent networks that are relevant to our experimental design. This general primary processing network might help to process the incoming visual information and then influence other stimulus-specific networks to take over, depending upon the statistical characteristics of the visual input. Due to the definition of the indepen-

dent components, it cannot be established distinctly if this causal influence is due to the primary visual component or the precuneus network in component 25.

### **Buffer theory**

In manuscript 1, we proposed the hippocampus as primary candidate for episodic buffer as it can extract relationships from the incoming information and can communicate well with cortical regions. This limited capacity episodic buffer holds information temporarily in terms of multimodal codes and is relevant for binding information across various underlying systems and with the long-term memory. Such binding capabilities allows this component to create a unified episodic representation of the incoming information and should be able to contribute to the online maintenance of the integrated memory traces.

For a meaningful (or well structured) stimulus, long timescale cortical regions (similar to DMN) are reported to be able to intrinsically retain some information from at least a few minutes prior (Hasson et al., 2015; Chen et al., 2016). Information embedded within a continuous predictable context (tunnel videos in our case) can persist over time outside the hippocampus, but visual information without any context or statistical regularity (for example, pixel scrambled videos) might be difficult to sustain. Greater retrieval demands (no usable statistical information) for pixel scrambled videos might cause more informational load on the hippocampus and hence higher activity. In the absence of any useful abstraction from the incoming visual stimulus, persisting or learning any informational representation is difficult and needs continuous involvement of the brain region responsible for such task or goal-based information extractions.

### **Activity localization in the hippocampus**

We found the hippocampal activity in different components to be overlapping to some extent. Component 6 and component 25 both showed an overlapping positively significant activity in the bilateral medial hippocampal body. The activity in component 6, additionally, extended to posterior subregions of the hippocampus. Component 25, on the other

hand, also showed a negatively significant bilateral activity in the anterior and posterior hippocampus.

Kahn et al. (2008) showed similar parallel cortico-hippocampal functional networks converging on distinct subregions of MTL, independent of the task using fMRI in humans. They proposed a Posterior-Medial and Anterior Temporal (PMAT) framework for cortico-hippocampal networks (Ranganath and Ritchey, 2012; Ritchey et al., 2015). The posterior medial (PM) network has the parahippocampus (PHC) and retrosplenial cortex (RSC) as the core components and extends to mammillary bodies, the anterior thalamic nuclei, the pre and parasubiculum and the default network (including the posterior cingulate, precuneus, lateral parietal cortex and medial prefrontal cortex). The anterior temporal (AT) system on the other hand has perirhinal cortex (PRC) as a core component and extends to the ventral temporopolar cortex, lateral orbito-frontal cortex and amygdala.

Component 25 from our ICA results included visual cortical areas along with the posterior hippocampus, parahippocampus, precuneus, lateral parietal cortex, anterior thalamic nuclei and some sections of the medial prefrontal cortex. These regions overlap with the major regions of the PM network mentioned above. The activation in the posterior hippocampus is consistent with our GLM results and is the region from the proposed PMAT model to which the cortical pathways from the PM network are strongly connected. This pathway has been shown to be involved with memory constructing the episodic simulations of future events (Buckner and Carroll, 2007; Hassabis et al., 2007; Schacter et al., 2007) and default modes of cognition (Buckner et al., 2008; Gusnard and Raichle, 2001). Some studies show that PHC and RSC have strong connectivity from the earlier occipital and temporal areas (Kahn et al., 2008; Ranganath and Ritchey, 2012) which also explains the cooccurrence of the visual cortical regions in the same network. This general visual processing component causally affects stimulus-specific components such as component 6, 14 and 42 and, therefore, conforms to a hierarchal but still a parallel process in the cortex, with information flow from a general to more stimulus-specific networks in the cortical regions.

We found dual hippocampal involvement while watching high entropy

visual stimulus. This duality comes from different independent networks and occurs at different timescales. One such component with more general role in visual processing drives other more stimulus-specific components at higher frequencies.

## **Conclusion**

Multiple cortico-hippocampal networks are recruited at different timescales for processing visual stimulus with high uncertainty. These networks performing different functions, in parallel, on the same visual input could be the key to understanding how these brain networks influence each other and how the information flows between such networks hierarchically.

## **Acknowledgments**

Authors of this manuscript are extremely grateful to the Department of Neuroradiology of the University Hospital, LMU Munich for their diagnostic support. They would like to thank Dr. Rainer Boegle and Kathrin Kostorz for their helpful remarks on the manuscript. They would also like to extend their thanks to DFG (GRK 1091, GSN) and BMBF (BCCN 01GQ0440, DSGZ 01EO0901) for funding the study.

## References

- Beckmann, C. F. and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE transactions on medical imaging*, 23(2):137–152.
- Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). *The Brain’s Default Network: Anatomy, Function, and Relevance to Disease*. *Annals of the New York Academy of Sciences*, 1124(1):1–38.
- Buckner, R. L. and Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, 11(2):49–57.
- Calhoun, V., Adali, T., Pearlson, G., and Pekar, J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14(3):140–151.
- Chen, J., Honey, C. J., Simony, E., Arcaro, M. J., Norman, K. A., and Hasson, U. (2016). Accessing Real-Life Episodic Information from Minutes versus Hours Earlier Modulates Hippocampal and High-Order Cortical Dynamics. *Cerebral Cortex*, 26(8):3428–3441.
- Demirci, O., Stevens, M. C., Andreasen, N. C., Michael, A., Liu, J., White, T., Pearlson, G. D., Clark, V. P., and Calhoun, V. D. (2009). Investigation of relationships between fMRI brain networks in the spectral domain using ICA and Granger causality reveals distinct differences between schizophrenia patients and healthy controls. *NeuroImage*, 46(2):419–431.
- Fraedrich, E. M., Flanagin, V. L., Duann, J.-R., Brandt, T., and Glasauer, S. (2012). Hippocampal involvement in processing of indistinct visual motion stimuli. *Journal of cognitive neuroscience*, 24(6):1344–1357.
- Fraedrich, E. M., Glasauer, S., and Flanagin, V. L. (2010). Spatiotemporal phase-scrambling increases visual cortex activity. *NeuroReport*, 21(8).
- Gramann, K., Onton, J., Riccobon, D., Mueller, H. J., Bardins, S., and Makeig, S. (2010). Human brain dynamics accompanying use of ego-

- centric and allocentric reference frames during navigation. *Journal of cognitive neuroscience*, 22(12):2836–2849.
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3):424.
- Gusnard, D. A. and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature reviews. Neuroscience*, 2(10):685.
- Hannula, D. E., Tranel, D., and Cohen, N. J. (2006). The Long and the Short of It: Relational Memory Impairments in Amnesia, Even at Short Lags. *Journal of Neuroscience*, 26(32):8352–8359.
- Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, 104(5):1726–1731.
- Hasson, U., Chen, J., and Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6):304–313.
- Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634.
- Jafri, M. J., Pearlson, G. D., Stevens, M., and Calhoun, V. D. (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage*, 39(4):1666–1681.
- Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., and Buckner, R. L. (2008). Distinct Cortical Anatomy Linked to Subregions of the Medial Temporal Lobe Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology*, 100(1):129–139.
- Li, Y.-O., Adali, T., and Calhoun, V. D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapping*, 28(11):1251–1266.

- Maldjian, J. A., Laurienti, P. J., and Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage*, 21(1):450–455.
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., and Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3):1233–1239.
- Olson, I. R. (2006). Working Memory for Conjunctions Relies on the Medial Temporal Lobe. *Journal of Neuroscience*, 26(17):4596–4601.
- Ranganath, C. and Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13(10):713–726.
- Riggs, L., Moses, S. N., Bardouille, T., Herdman, A. T., Ross, B., and Ryan, J. D. (2009). A complementary analytic approach to examining medial temporal lobe sources using magnetoencephalography. *Neuroimage*, 45(2):627–642.
- Ritchey, M., Libby, L. A., and Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition. In *Progress in Brain Research*, volume 219, pages 45–64. Elsevier.
- Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature reviews. Neuroscience*, 8(9):657.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464.
- Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012). Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cerebral Cortex*, 22(1):158–165.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–155.
- Timme, N., Ito, S., Myroshnychenko, M., Yeh, F.-C., Hiolski, E., Hotowy, P., and Beggs, J. M. (2014). Multiplex Networks of Cortical and Hippocampal Neurons Revealed at Different Timescales. *PLoS ONE*, 9(12):e115764.



## 5 Supplementary Information

Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
<b>Component 6</b>				
Calcarine + Lingual gyrus + Occipital Inf/Mid/Sup	L, R	1341	15, -90, -6	28.16
Hippocampus + Fusiform + Thalamus + Parahip- pocampus	R	117	24, -27, -3	11.23
Parietal Sup + Precuneus + Cuneus + Occipital Sup	L	86	-15, -75, 48	9.97
Cingulate gyrus + Posterior Cingulate	L, R	74	-3, -24, 30	9.79
Hippocampus + Parahip- pocampus	L	48	-24,-27,-12	9.56
Precuneus + Occipital Mid/Sup + Calcarine + Cuneus + Parietal Sup	R	248	15, -72, 42	8.98
Temporal Mid/Inf	R	87	54, -57, 0	8.88
Parietal Sup/Inf + Pre- cuneus	R	83	15, -57, 60	8.33
Temporal medial + Occipi- tal Inf/Mid	L	117	-51, -66, 3	8.16
Frontal Sup/Mid	L	54	-27, 63, 9	7.93
Insula + SupraMarginal	R	31	36, -27, 18	7.53
Post Central gyrus	R	20	66, -12, 33	7.50
Temporal Sup	L	22	-48, -36, 18	7.11
Angular Gyrus	L	32	-48, -63, 36	6.97
Fusiform gyrus + Occipital Mid/Inf	L	127	-24, -66, -18	-7.59
Lingual gyrus + Fusiform gyrus	R	78	21, -66, -9	-9.44

Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
Occipital Mid/Sup + Cuneus	L	118	-27, -87, 6	-11.17
Precuneus + Cuneus + Cin- gulate gyrus	L, R	231	0, -66, 21	-11.42
Occipital Mid/Inf/Sup + Fusiform	R	208	33, -81, 9	-13.31
<b>Component 14</b>				
Lingual gyrus + Occipital Inf/Mid/Sup + Calcarine + Temporal Inf + Fusiform gyrus	R	548	21, -96, -9	19.82
Occipital Inf/Mid + Lin- gual gyrus + Calcarine + Fusiform gyrus + Temp Inf	L	346	-27, -96, -12	16.53
Lingual gyrus + Calcarine + Posterior Cingulate	R	85	21, -48, 0	7.89
Lingual gyrus	L	24	-21, -54, 0	6.75
Amygdala	L	23	-27, -6, -12	6.75
Frontal Mid	L	27	-36, 45, 33	6.62
Putamen + Caudate	L	24	-12, 15, -9	6.61
Temporal Mid/Inf	L	30	-63, -60, -6	-6.46
Insula	L	28	-39, 15, 6	-6.54
Occipital Mid	L	23	-39, -81, 27	-6.93
Lingual gyrus + Calcarine + Occipital Sup/Mid + Fusiform gyrus	L, R	777	9, -81, -12	-16.45
<b>Component 22</b>				

Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
Lingual gyrus + Fusiform gyrus + Occipital Inf/Mid/Sup + Tem- poral Inf + Cuneus + Calcarine	L, R	2695	-18, -81, -15	26.52
Putamen	L	121	-21, 3, -12	8.99
Putamen	R	91	27, 9, -3	7.96
Precuneus + Cuneus	L	39	-6, -60, 18	7.47
Parietal Inf	L	21	-45, -54, 54	7.38
Cingulum Mid	R	23	9, -3, 36	-6.57
Precentral gyrus + Postcen- tral gyrus	R	35	36, -24, 54	-7.44
Cingulum Mid	L	40	-9, -18, 33	-7.54
Paracentral Lobule + Pre- cuneus + Supp Motor area	L, R	228	-3, -30, 57	-8.21
Parietal Sup + Occipital Mid/Sup + Angular gyrus + Precuneus	R	60	45, -78, 30	-8.44
Frontal Sup medial + Cin- gulum Mid	L, R	191	0, 39, 42	-8.45
Calcarine + Temporal Mid + Occipital Mid/inf/Sup + Lingual gyrus + Thalamus + Hippocampus + Pre- cuneus	L	903	-18, -69, 6	-12.83
Calcarine + Thalamus + Lingual gyrus + Temporal Mid/Sup + Precuneus + Cuneus + Hippocampus	R	588	15, -63, 6	-14.42
<b>Component 25</b>				

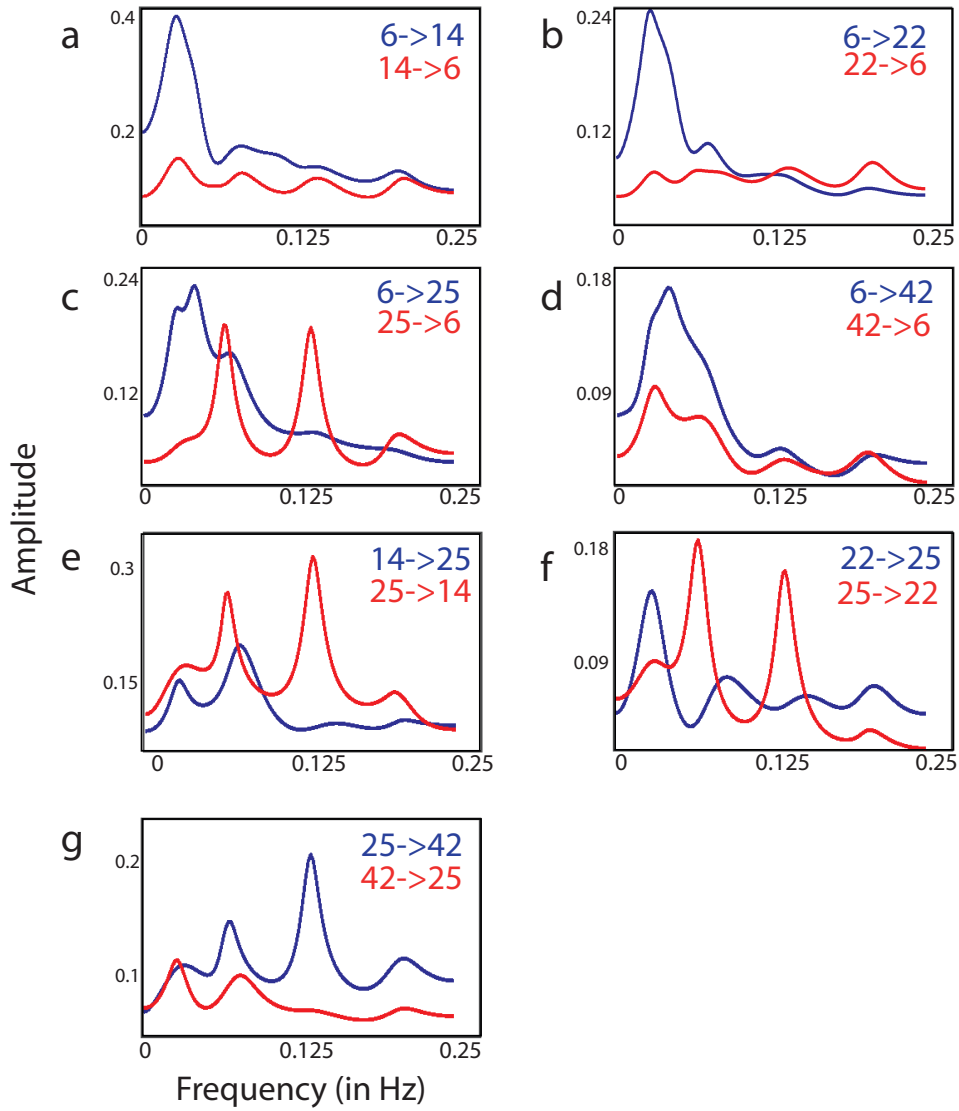
Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
Calcarine + Lingual gyrus + Cuneus + Fusiform gyrus + Occipital Sup/Inf + Pre- cuneus + Parahippocampus	L, R	3318	3, -72, 15	40.09
Supra Marginal gyrus + Postcentral gyrus + Pari- etal Inf + Precentral gyrus	R	233	51, -27, 45	11.35
Hippocampus	R	30	21, -24, -6	11.34
Caudate + Putamen + An- terior Cingulate	L	119	-12, 18, -6	11.29
Insula + Caudate + Frontal Inf Orbital + Putamen	R	163	15, 18, -3	10.97
Temporal Sup/Mid	R	37	54, -42, 12	10.96
Precuneus	L, R	61	0, -54, 42	10.69
Hippocampus + Thalamus	L	90	-24, -27, -9	10.17
Insula + Frontal Inf Orbital	L	52	-36, 15, -3	8.82
Parietal Inf/Sup	L	112	-51, -42, 51	8.42
Thalamus	R	30	15, -15, 6	6.82
Lingual gyrus	L, R	82	-6, -87, -21	-10.71
Occipital Mid	L	111	-42, -81, 33	-11.04
Temporal Mid/Inf	R	59	45, -66, 3	-11.29
Precuneus + Cingulum Mid/Post + Parietal Sup + Frontal Sup + Calcarine + Precentral gyrus + Cuneus + Supp Motor Area + Hippocampus + Thalamus	L, R	730	15, -60, 27	-12.60
Fusiform gyrus + Parahip- pocampus + Hippocampus	L, R	571	27, -33, -15	-14.01

**Component 42**

Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
Precuneus + Occipital Mid/Sup + Angular gyrus + Cuneus + Parietal Sup/Inf + Calcarine + Temporal Mid/Sup + Cin- gulum Mid + Postcentral + SupraMarginal gyrus + Lingual gyrus	L, R	4010	-33, -81, 27	25.39
Fusiform gyrus + Temporal Inf/Mid + Parahippocam- pus + Lingual gyrus + Hip- pocampus + Occipital Inf	L, R	1152	30, -36, -15	20.06
Putamen + Frontal Inf Or- bital + Caudate + An- terior Cingulate + Insula + Amygdala + Frontal Mid/Sup + Hippocampus	L, R	1165	30, 33, -12	12.76
Frontal Mid/Sup	R	89	24, 24, 45	10.80
Insula + Temporal Sup + Rolandic Oper	L	89	-36, -21, 15	10.49
Insula + Rolandic Oper	R	77	33, -24, 18	9.66
Lingual gyrus	R	117	15, -93, -9	9.27
Frontal Sup	L	51	-24, 57, 24	8.24
Temporal Mid/Sup	R	34	54, -9, -12	8.24
Frontal Mid	L	47	-27, 15, 45	7.32
Anterior Cingulate	L	48	-6, 42, 0	7.26
Temporal Sup	R	29	36, -39, 6	-6.79
Lingual gyrus	L, R	29	0, -72, 3	-7.19
Occipital Mid	L	21	-36, -69, 0	-7.76
Occipital Mid/Inf + Fusiform gyrus	L	94	-27, -81, -6	-8.20

Region	Hemi- sphere	Cluster size (voxels)	Peak coordi- nates x,y,z (MNI)	Max t
Frontal Inf Oper + Rolandic Oper	L	25	-63, 9, 6	-8.46
Supp Motor area	L, R	162	9, 3, 60	-9.09
Frontal Sup Medial + Cin- gulum Ant/Mid	L, R	64	3, 33, 30	-9.79
Cingulum Mid/Post	L, R	231	0, -42, 21	-11.53

**Table 2:** Anatomical regions from the top five independent components



**Figure 6:** Frequency profiles of the G-causal relationships between the components. Bidirectional relationships are shown over various frequency zones for the following component pairs, a. component 6 and 14, b. component 6 and 22, c. component 6 and 25, d. component 6 and 42, e. component 14 and 25, f. component 22 and 25 and g. component 25 and 42.

## 3 General Discussion

### 3.1 Purpose of the thesis

The purpose of this doctoral thesis is to identify the brain areas and networks involved in processing uncertainties in the visual stream, using neuroimaging measurements.

We found significantly higher hippocampal and early visual activity for the highly uncertain visual stream compared to more structured and predictable visual stimuli. We then revealed the underlying functional networks using Independent Component Analysis (ICA) and identified the cortico-hippocampal networks that could explain our experimental paradigm in the best possible way. Certain networks preferred specific stimulus types. A primary visual component, independent of the stimulus type, was found to represent a general visual processing network. We further evaluated the temporal dependence between the components using Granger causality and found a functional hierarchy in the components where information is flowing from the general visual processing component to relatively more stimulus-specific components.

In section 3.2, we present the interpretation and importance of the main findings from our analysis. We begin by discussing the cortical activities for meaningful and structured virtual tunnel videos. Then we discuss the hippocampal role in extracting statistical regularities from the dynamically uncertain visual input. We further suggest which other cognitive processes could use this information about the statistical structure of the visual input. We present the 'buffer theory' where we propose the hippocampus as an episodic memory buffer in line with Berlingeri et al. (2008). This framework could explain the hippocampal interest in such non-structured stimulus. We further discuss the anatomical and functional cortico-hippocampal networks in the brain and how the hippocampus could be in a position to support predictive coding. These networks could function at multiple granularities and timescales.

In the end, we explain how the current findings might change the way the hippocampus and other subcortical structures are studied in current neuroimaging studies. Rather than considering them as separate com-



partmentalized regions, we suggest treating them as an integrated part of visual perception, at least when the incoming visual information is not reliable. Another section including the outstanding questions resulting from this thesis is also included in the end.

## **3.2 Summary of the experimental findings**

Our fMRI analysis, presented in sections 2.1 and 2.2, showed surprising yet quite interesting results. While on the one hand, we found that the hippocampus was activated strongly in response to a visual stimulus with higher entropy at low-level statistics, on the other hand, we found the hippocampus to be part of a general visual processing network, independent of the task or the semantic content of the stimulus itself. To present these findings better, we discuss them individually in the following sections.

### **Cortical response to high entropy visual stimulus**

As shown in section 2.1, we found significantly stronger activity in the early visual cortex for pixel scrambled videos compared to both other types of videos. These videos have high spatial and spatiotemporal entropy but do not have any semantic meaning. On the other hand, for more structured and predictable videos, we found stronger activity in higher cortical regions, especially the dorsal stream and parahippocampal regions.

The dorsal stream, also known as the 'where pathway', has been primarily associated with processing the spatial location objects relative to the viewer (Milner and Goodale, 2006). The virtual tunnel videos were easily recognizable and mimicked the viewer's motion through a tunnel. As the tunnel videos presented a visuospatial stimulus, viewers could process the spatial context. These videos provided smooth motion flow, so the future visual input could be predicted in a top-down manner. This provides a two-fold explanation of the cortical activities we witnessed in response to the tunnel videos but not the pixel scrambled and phase scrambled videos. As the motion flow for these videos was pre-

dictable, feedback from higher cortical regions becomes more relevant, resulting in higher activation in these higher cortical regions. This could also lead to less weighting of the bottom-up processing stream and hence can also explain why activity in the early visual cortex was significantly weaker when watching the tunnel videos compared to the phase scrambled videos, which had similar low order statistical properties (matched for local luminance and flow of motion over time).

Pixel and phase scrambled videos, on the other hand, did not provide any spatial context or predictable flow of motion. This explains why the dorsal stream was activated for virtual tunnel videos but not for pixel scrambled or phase scrambled videos. While watching the pixel scrambled and the phase scrambled videos, for the lack of predictable future inputs, all the incoming information had to be processed, giving more emphasis to the bottom-up stream and hence lower cortical regions such as the early visual cortex showed significantly higher activity.

Similarly, parahippocampal activations for virtual tunnel videos are not surprising as it is known for its role in spatial information encoding. Pixel scrambled videos, on the other hand, do not contain any meaningful spatial information and hence did not engage the parahippocampus.

In section 2.2, we presented independent components involved in processing visual stimuli with varying spatial and spatiotemporal entropy using independent component analysis (ICA). Using these results, we found a component (component 42) that was activated strongly for the virtual tunnel videos. This network overlapped with the brain regions from the ventral default mode network (vDMN), anterior salience network and the precuneus network. These regions along with the hippocampus, parahippocampus and the retrosplenial cortex (RSC) are known for their role in episodic memory. As the tunnel videos were meaningful, cortical regions were able to retain this structured online information along with the information a few minutes prior, without any hippocampal involvement (Hasson et al., 2015; Chen et al., 2016).

### **Hippocampal involvement for high entropy visual stimulus**

The hippocampus has been previously reported to respond to various measure of regularities such as entropy (Strange et al., 2005), mutual in-

formation (Harrison et al., 2006), complexity (Baumann and Mattingley, 2013) and conditional entropy (Ahlheim et al., 2014), for both discrete and continuous inputs. As presented in section 2.1, we also found significantly stronger activity in the hippocampus for visual stimulus with high entropy. In section 2.2, we found that the hippocampus was a part of the functional network (component 6) that was involved in processing the pixel scrambled videos specifically.

We could, hence, answer the questions raised by Fraedrich et al. (2012). We investigated, using carefully designed stimuli, whether the hippocampal activity was due to an ongoing meaning uncovering process (uncertainty at the semantic level) or the entropy (uncertainty at low level) of the continuous visual stimuli. We found a stronger hippocampal activity for the pixel scrambled videos that had high entropy at low level statistics compared to both the virtual tunnel and the phase scrambled videos. The phase scrambled videos had lower spatial and spatiotemporal entropies than the pixel scrambled videos but while it was clear from the pixel scrambled videos that they did not contain any semantic information, we can say the same about the phase scrambled videos. These videos could entail an ongoing recognition process in an attempt to decode the semantic content. We did not find any significant activity for the phase scrambled videos in comparison to the pixel scrambled videos, indicating that the hippocampal activity was not due to an ongoing recognition process but actually due to the high uncertainty (spatial and spatiotemporal entropy) of the videos.

While it was surprising at first to observe the hippocampus, a region previously associated with higher cognitive functions such as memory, navigation and encoding, being activated in response to low-level properties in the visual stimulus, we found that similar responses have been reported before. Steinmetz et al. (2011) also found that hippocampal neurons responded to both low level (contrast) and high level (stimulus class like building, face, animal) properties of images.

Another explanation for such activations could be the non-learnability of these high entropy videos. As they did not have any relational structure, the information attained faded away quickly from the system and hence continual resampling of the stimuli was needed (Warren et al., 2015).

This resampling could be a function of the episodic memory buffer proposed by Baddeley (2000).

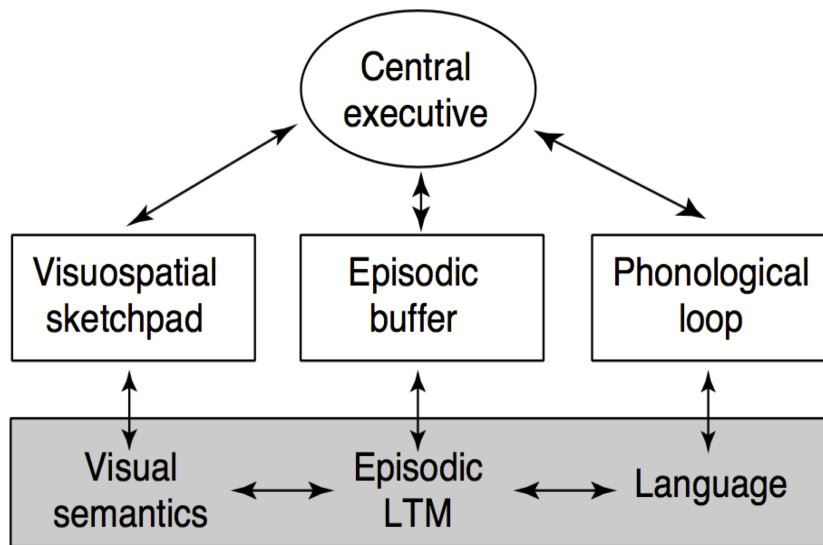
### **Is the hippocampus an episodic memory buffer?**

Baddeley and Hitch (1974) proposed a multicomponent model of working memory consisting of a supervisory central executive control component, a phonological loop and a visuospatial sketchpad. While the phonological loop deals with auditory information, the visuospatial sketchpad holds visual information. The central executive control component, as the name suggests, coordinates and controls these two components. A few years later they introduced another component called the episodic memory buffer to the working memory model (Baddeley, 2000). This new component holds information temporarily in terms of multimodal codes. It has a limited capacity and is relevant for binding information across various underlying systems and with the long-term memory. Such binding capabilities allow this component to create a unified episodic representation of the incoming information and should be able to contribute to the online maintenance of the integrated memory traces. The complete model with all the four components is shown in Figure 6.

Berlingeri et al. (2008) performed a voxel-based morphometry study using fMRI on Alzheimer's patients. They found that the anterior part of the hippocampus was significantly atrophic in patients with impaired prose recall task. They suggest that the anterior hippocampus contributes to the episodic buffer in the new working memory model.

We propose a similar function of the hippocampus as an episodic memory buffer. Such an episodic buffer is responsible for holding the incoming information until it is decided if the information is relevant and needs to be stored for the long-term or just discarded. This relevance could be influenced by various factors such as behavioural significance, attention, task requirements, irregularities etc.

In our experimental data, the hippocampus was part of the general visual processing network, which processes all the incoming visual information regardless of the content of the videos or the task at hand. For the continuously incoming information to be interpreted, it needs to be temporarily stored to extract both spatial and temporal regularities. While



**Figure 6:** Complete working memory model. This model consists of four important components. The visuospatial sketchpad holds the visual information temporarily and the phonological loop takes care of the incoming audio information. The episodic buffer stores information temporarily in terms of multimodal codes. The central executive component coordinates and controls these three components for a smooth processing of the incoming multimodal information in the working memory. Reprinted from Baddeley (2000) with permission from Elsevier.

spatial regularities could be extracted spontaneously, temporal regularities need information over time to find regularities or meaningful context. The hippocampus could be the region that holds such dynamic incoming information for further processing. Since it has strong connectivity with other cortical regions and is sensitive to the regularities in the incoming information, hippocampus for the role of an episodic memory buffer.

### **Hippocampal contributions to other cognitive processes**

Various cognitive processes can benefit from information about the statistical structure of the incoming visual stream. Such information can be attained by extracting spatial and temporal regularities from the input. The hippocampus, due to its role in representing uncertainty, could mediate the processes needing such information. It can mediate predictions based on past events. Two such processes are visual perception and visual learning.

Visual perception employs both top-down and bottom-up processing streams

to extract meaningful information in the most cost-effective manner. While the top-down processing stream transfers expectations and anticipated input to the lower level cortical regions, the bottom-up stream feeds-forward the current input from lower cortical regions to higher ones. How this information from both streams is integrated is not well understood. Based on the predictability of the current visual environment, different weights can be allocated to both streams. For a fairly predictable input stream, the top-down stream can anticipate the next inputs and hence it is not necessary to allocate more resources to process the future inputs. On the other hand, when the environment does not have regularities, it cannot be predicted based on the previous events and hence strong emphasis needs to be given to the information from the bottom-up processing stream. To assign these relative weights, we need information about the predictability of the current visual stimulus. Since the hippocampus is able to represent the predictability of the input stream, it could be the region that assigns these relative weights to the two processing streams (Strange et al., 2005).

Similarly, visual learning also needs information about the reliability and predictability of the environment to decide which learning scheme to employ: incremental or one-shot learning. While incremental learning can be used to acquire knowledge gradually through trial and error, one-shot learning needs rapid learning, sometimes even from a single instance of the relational information. When the incoming information is not predictable, rapid learning is required. Lee and Baxter (2010) investigated how the brain transitions between the two learning mechanisms. They found that the hippocampus acts as a 'switch' turning 'on' when one-shot learning is predicted to happen.

Both these cognitive processes need similar information about how predictable the current input is and hence might employ the same structure to retrieve it efficiently. Since the hippocampus can indicate the presence or absence of such regularities in the input stream, as shown from our General Linear Model (GLM) and Partial Least Squares Correlation (PLSC) results from our fMRI results in section 2.1, we propose that the hippocampus could be the region responsible for determining the predictability of the input stream and hence actively contributing to these

processes.

### **Analogue between visual perception and learning**

Other than sharing the hippocampus as an important contributor in their processing, there is an interesting analogue between visual perception and visual learning.

When the stimulus is unpredictable, more emphasis is given to the bottom-up processing stream. In this case, our current environmental model needs to be either modified or replaced to be able to correctly predict the future inputs. This model has to be learned rapidly and hence one-shot learning should be employed. On the other hand, when the incoming information is predictable, the brain puts more emphasis on the top-down processing stream to save brain resources. Due to the predictability of the input, incremental learning might suffice.

This presents an interesting analogue between visual perception and visual learning. When the information is reliable and predictable, we put more emphasis on the bottom-up processing stream and employ incremental learning but when the future inputs are not predictable, we cannot rely on the predictions from the top-down layers and hence need to learn rapidly via one-shot learning.

### **Can the hippocampus provide mediation in predictive coding?**

Predictive coding (Rao and Ballard, 1999; Huang and Rao, 2011) presents an efficient way to process the incoming information in the cortex. It helps to reduce the redundancy in the information and to allocate the brain resources in the most cost-effective manner. One way to filter out the irrelevant information is by skipping the processing of future visual inputs that can be easily predicted using our environmental model.

The hippocampus could also be useful for predictions based on the past events. Olsen et al. (2012) argued for hippocampal specialization in integration and formation of the representations, capturing the regularities in the environment. They also proposed its role in the online processing of the incoming information through an ongoing binding/comparison feedback process.

Hindy et al. (2016) related memory-based expectations in the visual system to the pattern completion mechanism in the hippocampus and proposed 'mnemonic' expectations in the brain that are different than the previously studied 'perceptual' expectations. While perceptual expectations help us to predict the changes in the current stimulus over time, mnemonic expectations are based on past events. For mnemonic expectations, the current stimulus does not need to match the physical attributes of the expected stimuli, but more the statistical regularities in space and time. These expectations need access to the past information to create the predictions for upcoming stimuli. The hippocampus was proposed to be the region responsible for such mnemonic expectations due to its role in relational binding and extraction of statistical regularities.

The hippocampus, as an episodic memory buffer, has access to information from the immediate past and hence can be used in extracting spatial and temporal regularities from it. To be part of such an ongoing general process, the hippocampus needs to be part of a general processing network, either functionally or anatomically, that functions independently of the task demands or semantic content (as it is added much later in the processing hierarchy) of the information. As shown in section 2.2, the hippocampus is, indeed, a part of such a primary visual network when processing all the videos, independent of the video type. Given this role in online processing, the facility to buffer the incoming information and its ability to analyze the statistical regularity from the incoming information, the hippocampus can make significant contributions to predictive coding.

### **Multiple cortico-hippocampal networks**

Kahn et al. (2008) examined cortico-hippocampal functional connectivity using high-resolution fMRI in 100 individuals. They found two different pathways correlating with different cortical and medial temporal lobe (MTL) regions. Regions such as the lateral parietal cortex, posterior cingulate, retrosplenial cortex (RSC) and the ventral medial prefrontal cortex correlated functionally to the posterior hippocampus and posterior parahippocampal cortex (PHC). On the other hand, regions such as the ventral anterior temporal cortex, lateral orbito-frontal cortex



and amygdala correlated with the anterior hippocampus and perirhinal cortex (PRC). Based on these networks and their spatial locations in the brain, Ranganath and Ritchey (2012) proposed the Posterior Medial Anterior Temporal (PMAT) framework. The Posterior Medial (PM) network consists of the parahippocampal cortex and the retrosplenial cortex along with the precuneus, posterior cingulate, angular gyrus, anterior thalamus, presubiculum, mammillary bodies and medial prefrontal cortex and extends to the posterior hippocampus. The Anterior Temporal (AT) network, on the other hand, includes the perirhinal cortex, ventral anterior temporal cortex, lateral orbito-frontal cortex and amygdala. This pathway converges at the anterior hippocampus.

Component 25 from our analysis overlapped with brain regions from the PM network mentioned above. The PM network is suggested to have a central role in the construction and application of situation models, that could be helpful in processing incoming information by signifying which particular model will be applicable. This network also plays a significant role in episodic memory and representing statistical information in the sensory system (Ranganath and Ritchey, 2012).

As shown in section 2.2, component 25, overlapping with the PM network, represented a general primary visual processing network as it responded strongly to all the stimulus videos. This activity was found to be independent of the actual content of the video. Similar rapid (<150 ms) and obligatory involvement have been reported previously using MEG (Riggs et al., 2009). They found this activity to be independent of the task demands and suggested an influence of prior experiences on the ongoing contextual processing.

Using Functional Network Connectivity (FNC) in section 2.2, we reported that this specific component, which is not stimulus-specific itself, 'g-causes' other components (component 6, 14, 22 and 42) that are more stimulus-specific. This directional functional connectivity presents an interesting hierarchy in the information flow where a general processing network, responding similarly to all the incoming visual information itself, influences other stimulus-specific networks to take over depending on some specific features or properties of the visual input. This hierarchy of specificity brings together an interesting scenario where networks that

are less specific to details proceed (and even influence) other networks that react to more detailed features of the information.

### **Hippocampal functions at different timescales and abstraction levels**

Brain regions operate at diverse timescales. In the resting human brain, cortical regions organize into a hierarchy of functionally-coupled networks characterized by distinct timescales. Recently, Sundaresan et al. (2017) showed the presence of such slow interactions in functionally networked human brain regions for both the simulated and the real fMRI data. Coordinated activity among the brain regions that mediate the cognitive processes would manifest in the form of functional connections among these regions at the corresponding timescales. Characterizing patterns of the functional connectivity that occur at these different timescales can enhance the understanding of how the brain produces behavior.

We performed spatial Independent Component Analysis to reveal the different networks involved in the stimulus processing. While the components were spatially independent, there was no explicit constraint of independence on their temporal structure. Due to the ICA assumption of linear mixing, the regions were temporally coherent within a component. Little had been explored about the temporal dependence between the different components. These temporal dependencies among the components, though smaller than the within-component dependencies, could have been significant (Calhoun et al., 2001). To evaluate these dependencies, resulting component time-courses were further analyzed.

Different timescales were hence analyzed in a post-hoc manner when cortico-hippocampal networks with different the frequencies emerged in the Independent Component Analysis. While analyzing their role in stimulus processing, it was discovered that these components respond differentially to our stimulus set. While some components/networks were active strongly for a particular stimulus, some of these networks were more indifferent to the presented stimuli. Temporally, higher areas can activate selectively for stimuli that are coherent over longer periods of time. Component 42, comprising of the brain regions in the anterior

saliency network, represented a similar behavior for the virtual tunnel videos (only meaningful and salient stimulus type in our study). Similarly, component 25, the primary visual network could represent the low-level visual processing. Such processing would be mandatory for all the incoming visual information, irrespective of its meaning and saliency. As reflected by the time course of this component, it is indeed activated for all the stimulus videos.

Moreover, we found two parallel cortico-hippocampal networks (component 6 and component 25) functioning at the different timescales and stimulus specificity. While one component (component 6) reacted strongly to the pixel scrambled videos, the other one (component 25) reacted to all the videos, independent of the video type (the virtual tunnel, phase scrambled or pixel scrambled videos). Component 6 reacted only to a specific stimulus type and functioned on a much longer timescale. Component 25, on the other hand, functioned on a shorter timescale and hence had a high frequency. We further investigated the temporal relationships between the independent networks using the Granger causality and found a hierarchical system where the general visual processing network 'influences' the stimulus-specific networks. Analyzing the different timescales of the components thus provided us a better insight into the relative contribution of the networks.

Hasson et al. (2015) proposed a hierarchical memory system where each cortical region functions at different timescales and stores the information it needs for processing. They proposed that lower cortical regions process information at a much faster rate and hence store small chunks of information. On the other hand, higher cortical regions process information at larger granularity and work at much longer timescales.

While they included the hippocampus as a major component in this hierarchy, they could not establish the time course for hippocampal processing given that previous studies have shown that it responds at different timescales (Squire and Zola, 1998; Hannula et al., 2006; Pertzov et al., 2013; Hannula and Duff, 2017). The reason why they could not find a specific timescale could be because the hippocampus contributes to multiple cognitive functions at different timescales and hence there is no specific timescale associated with hippocampal processing. We presented

evidence for dual hippocampal involvement at different timescales and granularity in section 2.2. The two components involving the hippocampus and other cortical regions were activated at different timescales. They also processed information at different abstraction levels. While component 6 was more specific to pixel scrambled videos, component 25 was the general primary visual network processing all the incoming visual information.

Steinmetz et al. (2011) reported hippocampal neurons responding to different features of images. Some of these neurons responded to low-level features such as contrast and illumination. Some other neurons reacted to higher level properties like stimulus class etc. Such an ability to process information at different granularity levels explains hippocampal contributions to different cognitive processes needing information at different abstraction levels. For example, the hippocampus is involved in both episodic memory and regularity extraction. While episodic memory needs detailed information to successfully encode an event, regularity extraction needs information about the similarity of current and past experiences. Evensmoen et al. (2015) also reported a granularity gradient (coarse-grained to fine-grained) along the hippocampal and entorhinal anterior-posterior axis.

The Complementary Learning System (CLS) explained how these different levels of abstractions could be implemented in hippocampal circuitry (Schapiro et al., 2017). They created a neural network based on the hippocampal anatomical connections. They proposed a theory about hippocampal functioning emphasising the need for compromise between the abstraction levels for the hippocampus. This compromise could enable the hippocampus to perform different functions in parallel. Through the computational model, these different functions are embedded in two different pathways in the hippocampal system. The monosynaptic pathway (connecting the entorhinal cortex to CA1 directly) enabled the hippocampus to process information for statistical learning. For episodic memory, the trisynaptic pathway (connecting the entorhinal cortex to CA1 through the dentate gyrus and CA3) helped the hippocampus in learning individual episodes.

## Summary

To summarize, we would like to argue for hippocampal contributions to cognitive functions outside memory. Our experimental data showed that the hippocampus is involved in two different cortico-hippocampal networks to process the visual input. The first cortico-hippocampal network processed all the incoming information, independent of the stimulus type. Another cortico-hippocampal network showed higher activity for a specific stimulus type, stimulus with high uncertainty. Both these networks process information on different abstraction levels and timescales; and the hippocampal formation is composed of anatomical pathways to support these different granularity levels.

## 3.3 Possible implications

The hippocampus is an important brain region residing inside the depths of the Medial Temporal Lobe (MTL). Its involvement in various cognitive functions such as short-term memory, navigation, statistical learning, visual perception, visual learning and many other important functions has left researchers uncertain about its fundamental functioning. Experimental studies designed to study the hippocampus and other MTL regions have usually focussed on a specific function and hence cannot be used to form an integrated model of the hippocampus. Therefore it is important to focus on the computations performed by the hippocampus rather than focussing on singleton functional elements.

We found various interesting aspects of hippocampal processing that have never been explored before. We found that the hippocampus and the primary visual network worked together in processing the incoming visual information, irrespective of the actual semantic content. This network supports the role of the hippocampus as an episodic memory buffer, as suggested by (Berlingeri et al., 2008). This network further 'influences' other more stimulus-specific networks.

In this thesis, we suggest differential hippocampal involvement in processing high entropy stimulus. While the hippocampus, functioning as a buffer, is involved in processing all the visual inputs, we found another

process, involving the hippocampus, dedicated to videos with high uncertainty. Such sensory information can be of behavioral advantage and hence is treated separately. Computations involved in extracting such regularities could be the shared connection between various functions of the hippocampus. Understanding these computations could be the key to understanding the unifying model of the hippocampal functions.

To understand what each brain region is doing, we need to investigate the underlying networks. Brain regions, such as the hippocampus, can be part of multiple networks at the same time, working at different timescales. Univariate and whole brain approaches, due to their methodological shortcomings, cannot be used to separate the activities resulting from different processes. Therefore, we emphasize the need for network-based methods in understanding the underlying mechanisms.

### 3.4 Outlook and outstanding questions

This thesis sheds light on the cortico-hippocampal involvement in processing videos with varying entropies. It further raises some orthogonal questions that, in spite of their relevance to the underlying study, could not be answered from our experiment. Taking inspiration from this thesis, the following questions can be answered through carefully designed experiments:

**Is the hippocampal involvement thresholded for a certain minimal uncertainty or is it gradual and hence increases with increasing uncertainties?**

In this thesis, we were able to demonstrate the hippocampal involvement for high entropy visual stimuli. Since we had three levels of uncertainty in the visual stimuli, we were unable to answer whether the cortico-hippocampal networks emerge after a certain threshold of uncertainty or whether it was a gradual development with increasing visual uncertainty. To study this systematically, stimuli with more controlled entropy could be created. Rather than a dynamic visual stimulus with a categorical entropy (high, low and intermediate), future stimuli could be created with gradual changes in the underlying uncertainties.

### **Is this hippocampal activity consistent in response to uncertainty at other levels of visual information?**

Different studies have shown strong hippocampal activity in response to uncertainties in relational, sequential and other high-level visual inputs. This thesis focused majorly on the low-level entropy (pixel-wise, frame-wise and between-frames) hippocampal involvement for uncertain but meaningful actions. An intersection of low-level and high-level uncertainties would be an interesting direction to take this study further. This hierarchy in entropy can be exploited using stimuli with varying entropy magnitudes and levels, from pixel-wise entropy to all the way to the daily life uncertain actions.

### **Is hippocampus involved in processing high entropy stimulus across other modalities as well?**

We investigated the hippocampal role in processing high entropy visual stimuli, but it was outside the scope of our study to answer if there are similar networks for uncertainty across the different modalities, for example auditory stimuli. Is the cortico-hippocampal networking limited to the visual information only? It would be interesting to see if there are some modality-specific and/or modality-general networks involved. To answer this, similar stimuli with varying entropies should be tested for the different modalities.

### **Homogeneous hippocampal clusters or heterogeneous clusters with both the hippocampus and thalamus/LGN?**

A few voxels from the right hippocampal cluster were categorized as LGN/Thalamus by the Human Atlas from WFU PickAtlas (Paper 1: Table 1). Since the majority of the significant voxels from our bilateral clusters, including the peak voxel, were marked as hippocampus by the Atlas in all both our analyses (GLM and PLS, see Paper 1: Table 1 and 2), we did not perform any functional differentiation in the unified cluster. To separate any such LGN activity from the underlying hippocampal cluster systematically, an LGN localizer could be employed. Using the localizer upfront, it would be easier to further break down the hippocampal cluster and separate the underlying LGN activity, if any.

## References

- Ahlheim, C., Stadler, W., and Schubotz, R. I. (2014). Dissociating dynamic probability and predictability in observed actions—an fMRI study. *Frontiers in Human Neuroscience*, 8.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, 4(11):417–423.
- Baddeley, A. D. and Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation*, volume 8, pages 47–89. Elsevier.
- Balasubramanian, V. and Sterling, P. (2009). Receptive fields and functional architecture in the retina. *The Journal of Physiology*, 587(12):2753–2767.
- Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Physical Review Letters*, 103(23).
- Baumann, O. and Mattingley, J. B. (2013). Dissociable Representations of Environmental Size and Complexity in the Human Hippocampus. *Journal of Neuroscience*, 33(25):10526–10533.
- Baxter, M. G. (2009). Involvement of Medial Temporal Lobe Structures in Memory and Perception. *Neuron*, 61(5):667–677.
- Bergmann, E., Zur, G., Bershady, G., and Kahn, I. (2016). The Organization of Mouse and Human Cortico-Hippocampal Networks Estimated by Intrinsic Functional Connectivity. *Cerebral Cortex*, 26(12):4497–4512.
- Berlingeri, M., Bottini, G., Basilico, S., Silani, G., Zanardi, G., Sberna, M., Colombo, N., Sterzi, R., Scialfa, G., and Paulesu, E. (2008). Anatomy of the episodic buffer: a voxel-based morphometry study in patients with dementia. *Behavioural neurology*, 19(1-2):29–34.
- Bird, C. M. and Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience*, 9(3):182–194.



- Bornstein, A. M. and Daw, N. D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning: Sequential predictions in hippocampus and striatum. *European Journal of Neuroscience*, 35(7):1011–1023.
- Bussey, T. and Saksida, L. (2007). Memory, perception, and the ventral visual-perirhinal-hippocampal stream: Thinking outside of the boxes. *Hippocampus*, 17(9):898–908.
- Calhoun, V., Adali, T., Pearlson, G., and Pekar, J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14(3):140–151.
- Chen, J., Honey, C. J., Simony, E., Arcaro, M. J., Norman, K. A., and Hasson, U. (2016). Accessing Real-Life Episodic Information from Minutes versus Hours Earlier Modulates Hippocampal and High-Order Cortical Dynamics. *Cerebral Cortex*, 26(8):3428–3441.
- Demirci, O., Stevens, M. C., Andreasen, N. C., Michael, A., Liu, J., White, T., Pearlson, G. D., Clark, V. P., and Calhoun, V. D. (2009). Investigation of relationships between fMRI brain networks in the spectral domain using ICA and Granger causality reveals distinct differences between schizophrenia patients and healthy controls. *NeuroImage*, 46(2):419–431.
- Evensmoen, H. R., Ladstein, J., Hansen, T. I., Møller, J. A., Witter, M. P., Nadel, L., and Håberg, A. K. (2015). From details to large scale: The representation of environmental positions follows a granularity gradient along the human hippocampal and entorhinal anterior-posterior axis: A Granularity Gradient Along Hippocampal Long Axis. *Hippocampus*, 25(1):119–135.
- Fraedrich, E. M., Flanagan, V. L., Duann, J.-R., Brandt, T., and Glasauer, S. (2012). Hippocampal involvement in processing of indistinct visual motion stimuli. *Journal of cognitive neuroscience*, 24(6):1344–1357.
- Fraedrich, E. M., Glasauer, S., and Flanagan, V. L. (2010). Spatiotemporal phase-scrambling increases visual cortex activity. *NeuroReport*, 21(8).

- Hannula, D. E. and Duff, M. C., editors (2017). *The Hippocampus from Cells to Systems*. Springer International Publishing, Cham.
- Hannula, D. E., Tranel, D., and Cohen, N. J. (2006). The Long and the Short of It: Relational Memory Impairments in Amnesia, Even at Short Lags. *Journal of Neuroscience*, 26(32):8352–8359.
- Harrison, L., Duggins, A., and Friston, K. (2006). Encoding uncertainty in the hippocampus. *Neural Networks*, 19(5):535–546.
- Hasson, U., Chen, J., and Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6):304–313.
- Hindy, N. C., Ng, F. Y., and Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience*, 19(5):665–667.
- Hsieh, L.-T., Gruber, M., Jenkins, L., and Ranganath, C. (2014). Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. *Neuron*, 81(5):1165–1178.
- Huang, Y. and Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(5):580–593.
- Jafri, M. J., Pearlson, G. D., Stevens, M., and Calhoun, V. D. (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage*, 39(4):1666–1681.
- Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., and Buckner, R. L. (2008). Distinct Cortical Anatomy Linked to Subregions of the Medial Temporal Lobe Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology*, 100(1):129–139.
- Krishnan, A., Williams, L. J., McIntosh, A. R., and Abdi, H. (2011). Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *NeuroImage*, 56(2):455–475.

- Lee, A. C. and Baxter, M. G. (2010). Hippocampus and configural-relational information: A relationship confined to memory? *Proceedings of the National Academy of Sciences*, 107(6):E21–E21.
- Lee, A. C. H., Yeung, L. K., and Barense, M. D. (2012). The hippocampus and visual perception. *Frontiers in Human Neuroscience*, 6.
- Lee, S. W., O’Doherty, J. P., and Shimojo, S. (2015). Neural Computations Mediating One-Shot Learning in the Human Brain. *PLOS Biology*, 13(4):e1002137.
- Li, Y.-O., Adali, T., and Calhoun, V. D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapping*, 28(11):1251–1266.
- McKeown, M. J., Hansen, L. K., and Sejnowsk, T. J. (2003). Independent component analysis of functional MRI: what is signal and what is noise? *Current opinion in neurobiology*, 13(5):620–629.
- McKeown, M. J., Jung, T. P., Makeig, S., Brown, G., Kindermann, S. S., Lee, T. W., and Sejnowski, T. J. (1998). Spatially independent activity patterns in functional MRI data during the stroop color-naming task. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3):803–810.
- Milner, A. D. and Goodale, M. A. (2006). *The visual brain in action*. Number 43 in Oxford psychology series. Oxford University Press, Oxford ; New York, 2nd ed edition. OCLC: ocm70328132.
- Mozaffari, B. (2014). The medial temporal lobe—conduit of parallel connectivity: a model for attention, memory, and perception. *Frontiers in Integrative Neuroscience*, 8.
- Olsen, R. K., Moses, S. N., Riggs, L., and Ryan, J. D. (2012). The hippocampus supports multiple cognitive processes through relational binding and comparison. *Frontiers in Human Neuroscience*, 6.
- Pertsov, Y., Miller, T. D., Gorgoraptis, N., Caine, D., Schott, J. M., Butler, C., and Husain, M. (2013). Binding deficits in memory following

- medial temporal lobe damage in patients with voltage-gated potassium channel complex antibody-associated limbic encephalitis. *Brain*, 136(8):2474–2485.
- Ranganath, C. and Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13(10):713–726.
- Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1).
- Ratliff, C. P., Borghuis, B. G., Kao, Y.-H., Sterling, P., and Balasubramanian, V. (2010). Retina is structured to process an excess of darkness in natural scenes. *Proceedings of the National Academy of Sciences*, 107(40):17368–17373.
- Reddy, L., Poncet, M., Self, M. W., Peters, J. C., Douw, L., van Dellen, E., Claus, S., Reijneveld, J. C., Baayen, J. C., and Roelfsema, P. R. (2015). Learning of anticipatory responses in single neurons of the human medial temporal lobe. *Nature Communications*, 6:8556.
- Riggs, L., Moses, S. N., Bardouille, T., Herdman, A. T., Ross, B., and Ryan, J. D. (2009). A complementary analytic approach to examining medial temporal lobe sources using magnetoencephalography. *Neuroimage*, 45(2):627–642.
- Ritchey, M., Libby, L. A., and Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition. In *Progress in Brain Research*, volume 219, pages 45–64. Elsevier.
- Schapiro, A., Kustner, L., and Turk-Browne, N. (2012). Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Current Biology*, 22(17):1622–1627.
- Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., and Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711):20160049.

- Schiffer, A.-M., Ahlheim, C., Wurm, M. F., and Schubotz, R. I. (2012). Surprised at All the Entropy: Hippocampal, Caudate and Mid-brain Contributions to Learning from Prediction Errors. *PLoS ONE*, 7(5):e36445.
- Seth, A. K., Barrett, A. B., and Barnett, L. (2015). Granger Causality Analysis in Neuroscience and Neuroimaging. *Journal of Neuroscience*, 35(8):3293–3297.
- Squire, L. R. and Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, 8(3):205–211.
- Stefanics, G., KremlÁk, J., and Czigler, I. (2014). Visual mismatch negativity: a predictive coding view. *Frontiers in Human Neuroscience*, 8.
- Steinmetz, P. N., Cabrales, E., Wilson, M. S., Baker, C. P., Thorp, C. K., Smith, K. A., and Treiman, D. M. (2011). Neurons in the human hippocampus and amygdala respond to both low-and high-level image properties. *Journal of Neurophysiology*, 105(6):2874–2884.
- Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., and Friston, K. J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Networks*, 18(3):225–230.
- Sundaresan, M., Nabeel, A., and Sridharan, D. (2017). Mapping distinct timescales of functional interactions among brain networks. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30*, pages 4109–4118. Curran Associates, Inc.
- Suzuki, W. A. (2009). Perception and the medial temporal lobe: evaluating the current evidence. *Neuron*, 61(5):657–666.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., and Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of cognitive neuroscience*, 21(10):1934–1945.

Warren, D. E., Duff, M. C., Cohen, N. J., and Tranel, D. (2015). Hippocampus contributes to the maintenance but not the quality of visual information over time. *Learning & Memory*, 22(1):6–10.

## 4 Appendix





## 4.1 Curriculum Vitae

**Nisha Dalal**

**March 1<sup>st</sup>, 1989**

**Nationality: Indian**

---

### Education

**Oct 2012 - Sept 2017:** PhD student (Neuroscience)

Graduate School of Systemic Neurosciences, LMU Munich, Germany

**July 2010 - Aug 2012:** Masters of Engineering (Info. Technology)

Indian Institute of Information Technology Allahabad, UP, India

**Aug 2006 - July 2010:** Bachelors of Technology (Computer Science)

University Institute of Engineering and Technology, MDU Rohtak, India

---

### Research experience

**Oct 2012 - Sept 2017:** Research assistant

Department of Neurology, Klinikum Grosshadern, Munich, Germany

**Aug 2015 - Sept 2015:** Summer school researcher

Center for Brains, Minds and Machines, Woods Hole (MA), USA

**July 2015 :** Summer school researcher

Bernstein Center for Computational Neuroscience, Tuebingen, Germany

**April 2010 - July 2010:** Intern/Programmer

Defence Research and Development Organization, New Delhi, India

---

### Technical skills

**Programming skills:** Matlab, SQL, C, C++, Java, OpenCV, HTML

**Software skills:** Tableau, MS Office, SPM, Adobe Illustrator, LaTeX

**Analytical skills:** Machine learning, Data visualization, Big data

**Languages:**Hindi (native), English (fluent), German (basic)



## 4.2 List of publications

### Manuscripts

1. **Dalal N**, Flanagin VL, Glasauer S. Hippocampal activation while watching visual stimuli with high spatiotemporal entropy, *Not submitted yet*
2. **Dalal N**, Flanagin VL, Glasauer S. Dual cortico-hippocampal network activation while watching high-entropy visual stimuli, *Not submitted yet*

### Published abstracts

1. **Dalal N**, Fraedrich E, Flanagin VL and Glasauer S. (2013) Investigating hippocampal activation elicited by watching indistinct motion stimuli, *Perception ECVF Abstract Supplement*, August 2013, 58
2. **Dalal N**, Flanagin VL and Glasauer S. (2015) Spatio-temporal uncertainty and cortical-hippocampal interactions: fMRI study, *Journal of Vision* September 2015, 991

### Selected conference presentations

1. **Dalal N**, Flanagin VL and Glasauer S. Cortical and hippocampal activations in response to visual spatial-temporal uncertainties, *presented at BCCN conference 2014*, Berlin
2. **Dalal N**, Flanagin VL and Glasauer S. Investigating hippocampal and cortical activations elicited by indistinct motion and noise stimuli, *presented at Open human brain mapping meeting 2014*, Hamburg



### 4.3 Eidesstattliche Versicherung/Affidavit

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation 'Cortico-hippocampal activations for high entropy visual stimulus: an fMRI perspective' selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation 'Cortico-hippocampal activations for high entropy visual stimulus: an fMRI perspective' is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

Munich, October 25, 2018

Nisha Dalal



## 4.4 Proof of author contributions

**Manuscript 1:** Hippocampal activation while watching visual stimuli with high spatiotemporal entropy, *Prior to submission*

**Authors:** Nisha Dalal, Virginia L. Flanagin, Stefan Glasauer

**Contributions:** Nisha Dalal designed and programmed the experiment, collected and analyzed data, created figures, interpreted the results and wrote the manuscript.

Virginia L Flanagin supervised the project, helped in interpreting results and reviewed the manuscript.

Stefan Glasauer supervised the project, helped in designing the experiment and in interpreting the results and reviewed the manuscript.

**Manuscript 2:** Dual cortico-hippocampal network activation while watching high-entropy visual stimuli, *Prior to submission*

**Authors:** Nisha Dalal, Virginia L. Flanagin, Stefan Glasauer

**Contributions:** Nisha Dalal analyzed the data, created figures, interpreted the results and wrote the manuscript.

Virginia L Flanagin supervised the project, helped in interpreting the results and reviewed the manuscript.

Stefan Glasauer supervised the project, helped in interpreting the results and reviewed the manuscript.

Munich, October 25, 2018

1<sup>st</sup>author: Nisha Dalal

1<sup>st</sup> supervisor: Stefan Glasauer





## 4.5 Acknowledgments

I have been lucky to receive support and motivation from so many incredible people during my thesis.

First and foremost, I would like to thank my Thesis Advisory Committee (TAC) members, Stefan, Virginia and Heiner for all their support, helpful remarks and scientific discussions that helped a computer science engineer submit a doctoral thesis in neuroscience.

Next, I would like to thank Rainer, Kathrin and Katie for being my last moment saviour with their helpful comments and revisions on my manuscripts and thesis.

I also want to convey my deepest gratitude towards various funding agencies (GRK 1091, BCCN, Department of Psychology and GSN), that made this thesis possible. The superlative GSN team especially Benedict, Catherine, Stefanie, Lena, Renate and Birgit deserve special thanks for their continuous support and uplifting spirit.

I would also like to thank my friends and colleagues from the Forschungshaus for constant motivation, intellectual discussions and creating a wonderful working environment. Be it sushi lunch or cocktails at Negroni's, I will cherish each and every moment spent with you guys. Especially the big and crazy India trip! Thanks for making my wedding an eternal experience for me, my family and all the hundreds of relatives present at the ceremony. My friends in distant lands deserve special applause for standing by me through thick and thin and making me realize that true friendship always stands the test of time and distance.

My parents, siblings and parents-in-laws are the best things in my small world. Thanks for making me believe that I can do this. Thanks for putting up with my crazy working hours last few months. To my Uncle and grandfather, I lost two months ago, I know that you will always be with me guiding me in life with your beautiful heartfelt smiles.

Last but not the least, all the love in the world to Kshitij, my husband, for being the only stable thing in my life when things went wild and chaotic. Your contribution in completion of this thesis is beyond words.

Last five years have been an amazing ride but wouldn't have been successful without all you guys! Cheers to you all!



