



**DECISIONS, PREDICTIONS AND LEARNING
IN THE VISUAL SENSE**

B. V. EHINGER

**DECISIONS, PREDICTIONS AND LEARNING
IN THE VISUAL SENSE**

B. V. EHINGER

Decisions, Predictions and Learning in the visual sense

Dissertation
zur Erlangung des Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
im Fachbereich Humanwissenschaften der
Universität Osnabrück

vorgelegt von
Benedikt V. Ehinger

May 2018

©MAY 2018 – BENEDIKT V. EHINGER

COVER: MODIFIED FROM **NIKOLAI LUTOHIN**, GALAKSIJA P.37 ISSUE #41,
SEPTEMBER 1975, BIGZ, BELGRAD

Abstract

We experience the world through our senses. But we can only make sense of the incoming information because it is weighted and interpreted against our perceptual experience which we gather throughout our lives. In this thesis I present several approaches we used to investigate the learning of prior-experience and its utilization for prediction-based computations in decision making.

Teaching participants new categories is a good example to demonstrate how new information is used to learn about, and to understand the world. In the first study I present, we taught participants new visual categories using a reinforcement learning paradigm. We recorded their brain activity before, during, and after prolonged learning over 24 sessions. This allowed us to show that initial learning of categories occurs relatively late during processing, in prefrontal areas. After extended learning, categorization occurs early during processing and is likely to occur in temporal structures.

One possible computational mechanism to express prior information is the prediction of future input. In this thesis, I make use of a prominent theory of brain function, predictive coding. We performed two studies. In the first, we showed that expectations of the brain can surpass the reliability of incoming information: In a perceptual decision making task, a percept based on fill-in from the physiological blind spot is judged as more reliable to an identical percept from veridical input. In the second study, we showed that expectations occur between eye movements. There, we measured brain activity while peripheral predictions were violated over eye movements. We found two sets of prediction errors early and late during processing. By changing the reliability of the stimulus using

the blind spots, we in addition confirm an important theoretical idea: The strength of prediction-violation is modified based on the reliability of the prediction.

So far, we used eye-movements as they are useful to understand the interaction between the current information state of the brain and expectations of future information. In a series of experiments we modulated the amount of information the visual system is allowed to extract before a new eye movement is made. We developed a new paradigm that allows for experimental control of eye-movement trajectories as well as fixation durations. We show that interrupting the extraction of information influences the planning of new eye movements. In addition, we show that eye movement planning time follow Hick's law, a logarithmic increase of saccadic reaction time with increasing number of possible targets.

Most of the studies presented here tried to identify causal effects in human behavior or brain-computations. Often direct interventions in the system, like brain stimulation or lesions, are needed for such causal statements. Unfortunately, not many methods are available to directly control the neurons of the brain and even less the encoded expectations. Recent developments of the new optogenetic agent Melanopsin allow for direct activation and silencing of neuronal cells. In cooperation with researchers from the field of optogenetics, we developed a generative Bayesian model of Melanopsin, that allows to integrate physiological data over multiple experiments, include prior knowledge on bio-physical constraints and identify differences between proteins.

After discussing these projects, I will take a meta-perspective on my field and end this dissertation with a discussion and outlook of open science and statistical developments in the field of cognitive science.

Contents

Publications	1
1 General Introduction: Vision, shaped by experience	5
1.1 How is the visual system organized?	7
1.2 Formalization of experience as prior information .	15
1.3 Predictive coding	18
1.4 Randomization and the quasi-experimental nature of eye movements	24
1.5 Ways to study causal interactions: Optogenetics . .	26
1.6 Overview over this thesis	28
2 Experience shapes computation: Learning new cate- gories	29
2.1 Layman’s summary	30
2.2 Abstract: Extensive training leads to temporal and spatial shifts of cortical activity underlying visual category selectivity	31
2.3 Introduction: Category Learning	32
2.4 Methods: The MEG adaptation paradigm	36
2.5 Results: Category selectivity at 116ms after exten- sive category learning	52
2.6 Discussion	61
3 Predictions and Decision making: Expectations trump reality	69
3.1 Layman summary	70

3.2	Abstract: Humans treat unreliable filled-in percepts as more real than veridical ones	72
3.3	Introduction: Decision making and reliabilities of inferred stimuli	73
3.4	Results: A bias towards filled-in stimuli	76
3.5	Discussion	90
3.6	Methods: An ambiguous 2-AFC task and a Bayesian mixed model	93
4	Action and Expectations: Predictions over eye movements and filled-in percepts	105
4.1	Layman summary	106
4.2	Abstract: Predictions of visual content across eye movements and their modulation by inferred information	107
4.3	Introduction: Eye movements test visual hypotheses	108
4.4	Methods: Gaze-dependent stimulus changes and inferred stimuli	110
4.5	Results: Early and late prediction errors and an interaction with inferred information	122
4.6	Discussion	132
5	Action and Sampling: Eye movements as an exploration-exploitation dilemma	139
5.1	Layman Summary	140
5.2	Abstract: Probing the temporal dynamics of the exploration-exploitation dilemma of eye movements	141
5.3	Introduction: Where and when to look at things? .	142
5.4	Methods: A new gaze-dependent paradigm to control the where and when of eye movements	149
5.5	Results: Longer exploitation time and less available options result in shorter decision time	160
5.6	Discussion	172

6 Bayesian Generative Models and Optogenetics: Modeling the Kinetics of Melanopsin	187
6.1 Layman Summary	188
6.2 Abstract: Understanding melanopsin using Bayesian generative models - an introduction	189
6.3 Introduction: Why should we develop Bayesian generative models of Melanopsin?	190
6.4 Method and Results: From Model Building, Parameter Estimation and Diagnostics	194
6.5 Model Extensions: Multiple cells, priors, different melanopsin variants and latent models	206
6.6 Chapter conclusion	219
7 General Discussion	221
7.1 Summary	221
7.2 Predictions for eye movements, and decision making in the blind spot	225
7.3 Eye movements and Predictions	228
7.4 Eye movements and Uncertainty	230
7.5 A Statistical Evolution in the Cognitive Sciences . .	231
7.6 An Open Science Evolution of the Cognitive Sciences	235
7.7 Concluding remarks	237
References	239
Appendix: Thesis Art	277

Publications

Peer-reviewed Publications

- Ehinger**, BV, P König, and JP Ossandón (2015). “Predictions of visual content across eye movements and their modulation by inferred information”. In: *Journal of Neuroscience* 35.19, pp. 7403–7413. DOI: 10.1523/jneurosci.5114-14.2015.
- Kietzmann, TC, BV **Ehinger**, D Porada, AK Engel, and P König (2016). “Extensive training leads to temporal and spatial shifts of cortical activity underlying visual category selectivity”. In: *NeuroImage* 134, pp. 22–34. DOI: 10.1016/j.neuroimage.2016.03.066.
- König, P, N Wilming, TC Kietzmann, JP Ossandón, S Onat, BV **Ehinger**, RR Gameiro, and K Kaspar (2016). “Eye movements as a window to cognitive processes”. In: *Journal of Eye Movement Research* 9, pp. 1–16. DOI: 10.16910/jemr.9.5.3.
- Spoida, K, D Eickelbeck, R Karapinar, T Eckardt, D Jancke, BV **Ehinger**, P König, D Dalkara, S Herlitze, and OA Masseck (2016). “Melanopsin variants as intrinsic optogenetic on and off switches for transient versus sustained activation of G protein pathways”. In: *Current Biology*. DOI: 10.1016/j.cub.2016.03.007.
- Ehinger**, BV, K Häusser, JP Ossandón, and P König (2017). “Humans treat unreliable filled-in percepts as more real than veridical ones”. In: *eLife* 6. DOI: 10.7554/eLife.21761.
- Ehinger**, BV, L Kaufhold, and P König (2018). “Probing the temporal dynamics of the exploration–exploitation dilemma of eye movements”. In: *Journal of Vision* 18.3, pp. 6–6. DOI: 10.1167/18.3.6.

This list contains only publications which are part of this dissertation

Non peer-reviewed Publications

Ehinger, BV, D Eickelbeck, K Spoida, S Herlitze, and P König (2016). “Understanding melanopsin using bayesian generative models- an Introduction”. In: *bioRxiv*. DOI: 10.1101/043273.

Talks

Ehinger, BV, P König, and J Ossandón (2012). “Early effects in visual category learning”. OCCAM, Osnabrück, Germany.

Ehinger, BV, P König, and J Ossandón (2015). “Prediction of visual content across eye movements and their modulation by inferred information in the blind spot”. VSS, Florida, USA.

Ehinger, BV, K Häusser, J Ossandón, and P König (2016). “Mind over matter: Decision Bias and Prediction Errors in the Blind Spot”. Silver Lab, Berkeley, USA.

Ehinger, BV, K Häusser, J Ossandón, and P König (2017a). “Humans treat filled-in information as more reliable than veridical ones”. Donders Discussions, Nijmegen, Netherlands.

Ehinger, BV, K Häusser, J Ossandón, and P König (2017b). “Humans treat filled-in information as more reliable than veridical ones”. ECVF, Berlin, Germany.

Posters (first author only)

Ehinger, BV, D Porada, A Engel, P König, and TC Kietzmann (2012). “100ms to Category-Selective Response Patterns: Ultra-rapid Categorization in Occipitotemporal Cortex as a Result of Extensive Category Learning”. Neural Coding, Decision-Making & Integration in Time, Rauschholzhausen, Germany.

Ehinger, BV, P König, and J Ossandón (2014a). “Prediction of visual content across eye movements and their modulation by inferred information in the blind spot”. 7th MMN Conference Leipzig, Germany.

- Ehinger, BV, P König, and J Ossandón (2014b).** “Prediction of visual content across eye movements and their modulation by inferred information in the blind spot”. Brain Conference, Copenhagen.
- Ehinger, BV, P König, and J Ossandón (2014c).** “Prediction of visual content across eye movements and their modulation by inferred information in the blind spot”. ECEM, Vienna, Austria.
- Ehinger, BV, P König, and J Ossandón (2014d).** “Robust and Explorative Analysis of EEG”. Statistical Challenges in Neuroscience, Warwick, United Kingdom.
- Ehinger, BV, P König, and J Ossandón (2014e).** “Robust and Explorative Analysis of EEG”. Donders Discussion, Nijmegen, Netherlands.
- Ehinger, BV, L Kaufhold, and P König (2015).** “Disentangling fixation duration and saccadic planning using gaze-dependent guided viewing”. Bernstein Sparks Workshop on Active Perceptual Memory, Berlin, Germany.
- Ehinger, BV, J Ossandón, and P König (2015).** “A logistic bayesian mixed model”. Bayesian Summerschool, Amsterdam, Netherlands.
- Ehinger, BV, D Eickelbeck, K Spoida, S Herlitze, and P König (2016).** “A Bayesian hierarchical model for the biophysical properties of melanopsin”. SFN, San Diego, USA.
- Ehinger, BV, K Häusser, J Ossandón, and P König (2016).** “In case of doubt: choose the made-up-one. Filled-in percepts are preferred in perceptual decision making”. FENS, Copenhagen, Denmark.
- Ehinger, BV, L Kaufhold, and P König (2016).** “Disentangling fixation duration and saccadic planning using gaze-dependent guided viewing”. ECVP, Berlin, Germany.
- Ehinger, BV, D Porada, TC Kietzmann, and P König (2016).** “A spatio-temporal analysis of MEG Adaptation Paradigms applied to extensive Visual Category Learning”. OHBM, Geneva, Switzerland.
- Ehinger, BV, L Kaufhold, and P König (2017a).** “Humans treat unreliable filled-in percepts as more real than veridical ones”. OCCAM, Osnabrück, Germany.
- Ehinger, BV, L Kaufhold, and P König (2017b).** “Humans treat unreliable filled-in percepts as more real than veridical ones”. ICON, Amsterdam, Netherlands.
- Ehinger, BV, B Wahn, and P König (2017).** “Predicting when humans lose a moving target in a multiple object tracking task: An EEG time-resolved Decoding Study”. Machine Learning Summer School, Tübingen, Germany.
- Ehinger, BV and P König (2018).** “Update: Humans treat filled-in information as more reliable than veridical ones”. Mind and Brain Symposium, Berlin, Germany.

General Introduction: Vision, shaped by experience

Our sensory systems are remarkably complex. Even in a simple tasks, like responding to a visual or auditory cue, the total rate of incoming signals through our senses to the brain is staggeringly high. The eye alone transfers light from the retina to the lateral geniculate nucleus (LGN), the former with 65 million photoreceptors getting input from the environment (Jonas et al., 1992) which is condensed to an output of 1 million ganglion cells (Curcio et al., 1990). Hearing has 11.000-16.000 outer cochlea hair cells (Bredberg, 1968; Nadol, 1988) with around 60.000 efferent ganglion cells in the auditory nerve (Nadol, 1988). These cells fire between 0 and 100 times per second (Levine et al., 1992; Warland et al., 1997). Of course, these numbers and rates are only rough estimates of information transfer because the information transfer code is unknown. It follows that the brain is an extremely complex system that processes a vast amount of incoming information, in our example ultimately compressing it to a single action: choosing to press a button. How does the brain do this? What kind of structural or prior knowledge does it use? In this thesis I investigate how prior experience of the world and predictive computations can help process this information.

The focus of this thesis is on the visual sense. The benefits to study the visual sense are many and here I will give only three

general reasons. One, it is commonly regarded as the most important human sense (Rock and Victor, 1964) dominantly used in Education, Art and Media. Two, it has a historical advantage over several other senses due to its beneficial anatomical location to study it in primates (Felleman and Van, 1991; Tootell et al., 1988). And three, its the sense with the largest amount of cortex dedicated to processing its information (Felleman and Van, 1991; Palmer, 1999; Kaas, 2008) ¹.

This chapter is structured as follows: I will first introduce some basics on visual processing. Then I will concentrate on the role of learning and experience in perceptual processing and the ubiquitous feedback connections in the visual cortex. These feedback connections lead us to predictive coding, a computational algorithm proposed to explain basic brain functioning. I will then discuss eye movements and finally, ways to study causal influences in neuroscience.

¹Of course, most cortex is inherently multimodal. A slightly less provocative statement would be, that among the primary sensory cortices, the visual cortex is the largest.

1.1 How is the visual system organized?

In the following a short overview of the visual system is given. For an in-depth introduction see Palmer (1999). I will start with the function of the eye, then the LGN, followed by the cortex with V1 and then the dorsal and ventral stream. Finally, I will discuss the ubiquitous feedback connections.

Light is transduced to chemical signals in the photoreceptors of the eye. There are two main types of photoreceptors, rods and cones. We have many more rods than cones in the retina, but cones are sensitive to specific, different wavelengths of the light, giving us the ability to perceive colors. Both rods and cones use proteins of the family of opsins to transduce light waves to electrochemical responses. Rhodopsin is used in rods and S/M/L-cone opsins in cones. Interestingly, a third opsin Melanopsin is expressed in the retina, not in photoreceptors but directly in special ganglion cells. I will present further details on Melanopsin in the fifth chapter of this thesis (Spoida et al., 2016; Ehinger, Eickelbeck, et al., 2016).

The photoreceptors are arranged on the retina in a concentric pattern with increasing cone density towards the center. The area with the highest photoreceptors density is called the fovea. It is the point with highest acuity, that is, resolution. (Purves et al., 2001).

The spatial arrangement and connectivity of the photoreceptors allows one to see color and details using the information from the cones. The connectivity of ganglion cells in the periphery is one-to-many, that is, many photoreceptors are integrated by a single ganglion cell. This is the reason for a blurred perception in the periphery. The one-to-one connectivity in the fovea allows us

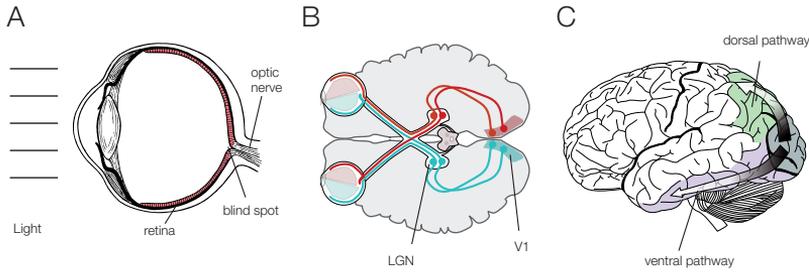


Fig. 1.1 A) The eye. Light is transduced to electro-chemical signals at the retina. The nerve fibers (ganglion cells) are bundled to the optic nerve and pierce through the retina at the blind spot, displacing all photoreceptors at that location. B) The optic nerve is split in left and right visual field and passes through the thalamus at the LGN. Finally it reaches V1. C) Visual processing is commonly split in two main pathways: The ventral "what" and dorsal "where" pathways. Image-Sources: Public Domain, M.P. Nieto (Wikimedia Commons), "Selket" (Wikimedia Commons)

to have a region with highly-detailed perception. Having only a single fovea makes it necessary to rotate the retina in a controlled way. We will discuss these eye movements in more detail in a later section, but make ample use of them in Chapter 4 & 5.

The ganglion cells are collected to a cell bundle, the optic nerve. The area where the optic nerve passes through the retina is called the physiological blind spot. At this location, no photoreceptors exist. Nevertheless, we do not perceive nothing: The brain automatically fills in information from the surrounding visual input (Figure 1.1). This is discussed in a later section, and we make use of the blind spot in chapter 3 & 4. The optic nerve mostly terminates in the thalamus, specifically in the LGN. A small portion targets the superior colliculus directly, bypassing further processing at the cortex. From the LGN, information is transferred to the superior colliculus and mainly the striate cortex (V1) in the neocortex. The properties and calculation of V1 cells are usually understood in terms of simple and complex cells. They both act as edge or grating detectors. The simple cell classically has a fixed receptive field, the complex cell has a stronger spatial invariance

(Carandini et al., 2005). In total, V1 is often referred as a filter bank (Olshausen and Field, 1996) of Gabor patches, or as local edge detectors. The reality is a lot more complex. There are strong context effects, that is, the activity of one cell depends on activity of neighboring cells. This is easily explained by inhibitory input from neighboring cells in V1 (Haslinger et al., 2012). The edge-detectors should not only be understood in the spatial domain, that is in x- or y-coordinates, but edges can also be observed in a spatio-temporal domain (DeAngelis et al., 1995), that is changing position over time. Some of these cells respond to motion (Hubel and Wiesel, 1968; Adelson and Bergen, 1985), often attributed only to higher level areas. V1 is one of the better understood cortical areas. Nevertheless a multitude of theories on its function in the larger context of the brain exists. They range from feature detection (Olshausen and Field, 1996; Bell and Sejnowski, 1997), over a necessary part of conscious perception (Keliris et al., 2010), to a cognitive blackboard (Roelfsema and Lange, 2016). It is quite likely, that this large cortical area is used for many different purposes, depending on the prevailing requirement of the system (TS Lee et al., 1998).

Continuing on from V1, there is evidence for a ventral and a dorsal pathway, loosely separated (Mishkin and Ungerleider, 1982). The ventral stream dominantly resolves the questions "what" and "how". This visual stream analyses features with more and more complex structure (Riesenhuber and Poggio, 2002). It culminates in areas that are highly selective for specific categories and objects. Areas of high interests for researchers are the face fusiformus area (FFA) and the parahippocampal place area (PPA). They are both located in the inferiotemporal cortex (IT). In the anterior path of IT very specific neurons were found, including the so-called grandmother cell (Quiroga et al., 2005) which is so specific that it responds only to ones own grandmother. We will discuss in Chapter 2 how such high specificity can arise during category learning.

The second, dorsal, stream, roughly relates to the location, the *where* of objects. Areas of interest here are usually V3, middle temporal (MT) and higher level areas as for example lateral intraparietal cortex (LIP). MT is famous for directed motion detection (Maunsell and Essen, 1983). The LIP is interesting, because it also shows strong connections to the ventral path. It is an area involved in the generation of saccades (Goldberg et al., 2006), the integration of rewards (Platt and Glimcher, 1999), in working memory (Gnadt and Andersen, 1988) and possibly episodic memory (Wagner et al., 2005). This concludes our summary of the visual hierarchy and the two perceptual pathways.

1.1.1 Experience shapes perception

It is unlikely, that the whole hierarchy presented above is predetermined by the genetic code and anatomical structure just develops without external influence. Instead, it is highly likely that the massive exposure to perceptual information we receive while growing up, strongly influences the structure of the brain. Here I present three examples that show how our organism adapts to the visual environment by learning from exposure.

The first example shows how developmental experience shapes what we can experience. The second one shows how the functional definition of cortical areas can be formed through exposure to certain categories. The third one shows how we learn new categories and how such categories can influence how we perceive the world.

In a seminal study, Blakemore and Cooper (1970) raised a cat in an environment, where its only visual input was daily exposure for five hours to vertical bars using the apparatus shown in Figure 1.2A. After five month of maturation, the cats were tested

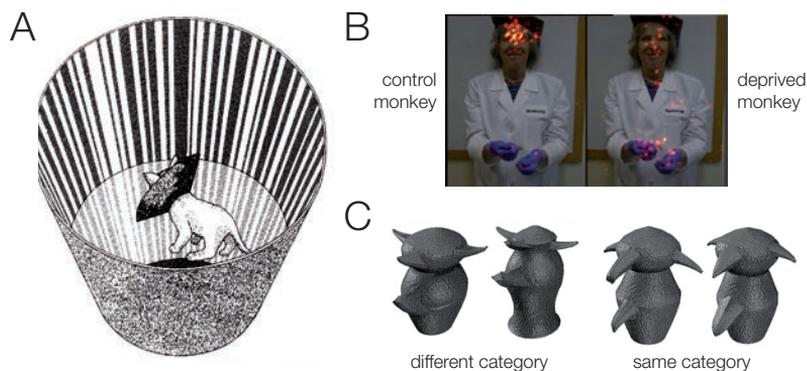


Fig. 1.2 A) Experiment of Blakemore and Cooper (1970). A cat is raised in an environment with only vertical bars but no horizontal ones. Subsequent tests show, that the cat is unable to see horizontal features. B) Arcaro et al. (2017) raised a monkey without seeing another monkey or human face. They showed for the first time, in opposition what was thought before, that also high level areas like FFA are shaped through experience. C) Gauthier and Nelson (2001) invented these so-called "greebles" to teach completely new categories to their subjects. Others show that this also influences the features that are extracted from the environment.

for direction selectivity and startling response of horizontal bars using behavioral and electrophysiological methods. Neither electrophysiological responses to horizontal bars, nor a startling response could be found. The cats acted as if they were blind towards horizontal bars. Because orientation selectivity to horizontal bars is such a well described feature of V1, one can conclude that the environment is critical for the visual development. This is in line with a study on ferrets, that show that silencing neurons in V1 precludes directional cells from forming (Chapman et al., 1996), but goes against a recent study which genetically silenced visual feedforward signals in mice (Hagihara et al., 2015). It is clear that behavioral and perceptual exposure to the environment is necessary to experience basic features of the environment.

Is this also true for higher areas and thus more complex features? In a recent study, Arcaro et al. (2017) show that monkeys do not form a monkey face-selective area, if they were not exposed to monkey face-like stimuli (Figure 1.2B). This shows that high

level cognitive areas are also shaped by exposure. Alternatively, it is possible, that distinct functional areas exist based on structure in the first place, but then are overwritten, or lost, by subsequent lack of exposure to the right stimuli.

In order to differentiate these possibilities, a pre- and post-test experimental design is needed. Such example can be found in the vast literature of category learning. In a typical category learning task, subjects are confronted with a new category. This new category could be based on novel objects (e.g. two types of greebles (Gauthier and Nelson, 2001)). Subjects learn to extract the features necessary to separate the two sets of one category (Figure 1.2C). It has been shown, that such category-experts activate areas (in this case FFA) similar to face-experts but for novel categories (Gauthier, Skudlarski, Gore, and AW Anderson, 2000; Gauthier and Nelson, 2001). Sigala and Logothetis (2002a) showed that extensive category learning in macaque monkeys leads to neurons that specifically spike for features that help differentiate categories, but not for others. They used Brunswick face stimuli (identical to ours in Chapter 2) with two diagnostic dimensions, eye height and eye width, and two non-diagnostic dimensions, nose height and mouth height. After category learning, most cells in the inferior temporal cortex that were specific to the stimuli, responded to the diagnostic dimensions, but only rarely did a neuron respond to the non-diagnostic dimension. Thus in the inferior temporal cortex, features that are specific to categories are extracted. Based on these three findings it follows that experience shapes perception.

1.1.2 Feedback in the visual cortex

So far we only discussed the feedforward stream. But feedback is ubiquitous in the cortex. Figure 1.3 shows that most areas are connected both ways to each other. We will see later, that the many

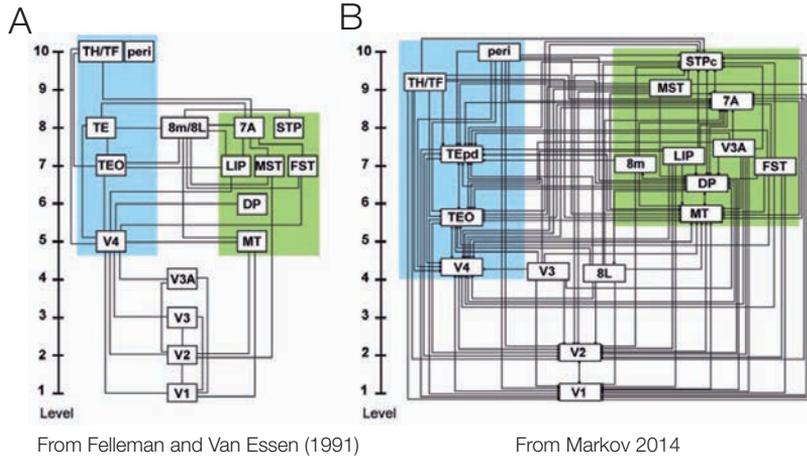


Fig. 1.3 A) Felleman and Van (1991) B) The updated map by Markov, Vezoli, et al. (2014) highlights the plentiful feedback connections. The hierarchy is build based on the ratio of incoming feedforward (lower level) and outgoing feedback connections (higher level)

feedback connections may allow the cortex to implement generative models in the framework of predictive coding. In LGN, around 90% of incoming fibers are not from the retina (Updyke, 1975; Cudeiro and Sillito, 2006), with around 60% of those incoming fibers representing local feedback and 30% cortical feedback. In V1 at least 85% of incoming fibers are local feedback from inhibitory cells (Markov, Ercsey-Ravasz, et al., 2013). Only around 2-3% of total connections are input-connections from the LGN. In the LGN it seems that the feedforward connections are the active drivers of the area, that is feedforward connections can elicit spikes. Feedback connections only modulate the response (Crick and Koch, 1998) and do not elicit spiking on their own. This classical view of modulation-only feedback has long been thought to generalize to the whole cortex (Sherman and Guillery, 1998). There are certain exceptions from this rule. For instance Mignard and Malpeli (1991) showed, that V1 neurons in supragranular layers can be also be activated by feedback connections from V2. More recent examples are the work of Shermann (Covic and Sherman, 2011; De Pasquale

and Sherman, 2011) that shows similar feedback driving activity in V1/V2 (and also other senses like auditory A1/A2). The general observation of feedback modulation and feedforward driving seems nevertheless quite accurate.

In humans, V1 feedback effects are reported in many different paradigms: E.g. visual imagery (Albers et al., 2013; Harrison and Tong, 2009), scene segmentation (Mumford et al., 1987; Lamme, 1995; Roelfsema, Lamme, et al., 2002), or perceptual decision-making (Roelfsema and Spekreijse, 2001; Nienborg and Cumming, 2009; Kok et al., 2012). Feedback signals are even more numerous the higher we follow the visual stream. In the visual cortex, it is estimated that more than two thirds of all connections are feedback connections (Markov, Vezoli, et al., 2014).

We can try to summarize the visual cortex as a hierarchical system which analyses the incoming data with increasing complexity of features. Noteworthy are the abundant modulatory feedback connections that in some cases even generate top-down activity.

1.2 Formalization of experience as prior information

Perception through our senses depends on the feedforward information passed through the cortex. But how is this input interpreted and combined with knowledge about our world? What are all these feedback connections used for? One popular explanation of the algorithmic level² of the brain (Marr, 1982) is termed predictive coding. It is usually referred to in the context of the Bayesian Brain and will be introduced in the following paragraphs. First, I need to introduce a mathematical formalism that allows one to describe the integration of prior experience and current information: Bayes' theorem. In words it states the following: We try to understand the world given some sensory input, the probability of some world state Θ given our incoming data D . We cannot access this directly, but we can simulate or calculate $p(D|\Theta)$ (the likelihood): the probability that we get this incoming data, given all possible states of the world as we currently understand it. We combine this likelihood with $p(\Theta)$, the prior, the probability that a certain state of the world could occur.

The formula is simple:

$$p(\Theta|D) = \frac{p(D|\Theta) * p(\Theta)}{p(D)} = \frac{\textit{likelihood} * \textit{prior}}{\textit{normalization}}$$

If we adopt a Bayesian notion, everything we learn about the world is assumed to be coded in a prior and a likelihood function. Learning the probability of certain states in the world will update our prior function. Learning about the mechanisms of the world

²The algorithmic level (how is it calculated?) is right in between the computational level (what is computed/the goal) and the implementation level (how is it done in the brain)

will also change our likelihood function. Integrating both, prior knowledge and current data will lead to our subjective probability distribution of what is out there in the world.

Let us take a simple example of inferring whether a dark silhouette on a foggy morning is a human or a lantern. For the likelihood function, the brain would need to calculate how likely it is, to see such a dark silhouette given that it could be a human and given that it could be a lantern. This is a type of forward simulation. The prior would encode that e.g. in the early morning it is much more likely to see a lantern than a human (in a different context, such as a rural environment, these prior probabilities could be drastically different). Combining both using Bayes' rule would allow the brain to perform a (Bayes optimal) decision on whether to perceive the silhouette as a human or a lantern (in this case, if the likelihood would show no preference, i.e. 50:50, but the prior would prefer a lantern, the brain should decide to perceive a lantern).

It is far from clear, whether the brain actually implements and uses Bayes' rule. There are many examples for which perceptual decisions are non-optimal (for recent reviews see Rahnev and Denison, 2018; Rahnev, 2017). But in other areas optimality has been shown, for instance in motor related areas (Darlington et al., 2017). It is widely accepted, that using a full probability distribution over all possible states in the world is impossible because there are simply too many possible states. It is even difficult to imagine, that the brain holds a single continuous probability distribution, without somehow approximating it. Luckily, there are many heuristics and shortcuts to approximate Bayes-like behavior. One popular example for an approximate Bayesian inference scheme is variational Bayes. There, the posterior is approximated by a simpler function that can be easily summarized (for instance, a Gaussian distribution, ranging from negative infinity to posi-

tive infinity, is completely described by only two parameters, the mean and variance). In higher dimensions, only the means and a covariance-matrix would suffice to describe the whole posterior distribution.

It is currently up to debate, whether the brain uses some kind of (variational) Bayesian inference (K Friston, 2010). The discussion is convoluted by an unclear test of what represents a Bayes optimal observer, and further by heuristics that can look and act like Bayes in a limited set of situations, but are not full Bayes (Rahnev, 2017; Goeke et al., 2016). To better understand how Bayesian inference could be implemented in the brain, we will have a look at the algorithmic level and discuss predictive coding.

1.3 Predictive coding

As we saw before, feedback ingrains every area of the brain. Therefore, it is only reasonable to think of possible computations that such a structure allows for. A quite popular and influential idea is predictive coding (Srinivasan et al., 1982; Mumford, 1992; Rao and DH Ballard, 1999; K Friston, 2005). The central idea to this theory is that "feedback connections from a higher- to a lower-order visual cortical area carry predictions of lower-level neural activities, whereas the feed-forward connections carry the residual errors between the predictions and the actual lower-level activities" (Rao and DH Ballard, 1999).

The brain is often assumed to be generative, that is, the brain is actively generating a representation of the outside world (K Friston, 2005). This allows for an efficient computational code if predictive coding is employed. When the predictions of the world are good, less activity and thus less energy needs to be used to perform the needed computations. There exists ample evidence that can be interpreted as a predictive code in the brain (Huang and Rao, 2011), for instance spatial and temporal predictions in the retina (Huang and Rao, 2011) or the LGN (Dan et al., 1996). In higher areas, much evidence is compatible with predictive coding, but direct tests are quite difficult to perform (Aitchison and Lengyel, 2017).

1.3.1 Predictive Coding and the Bayesian brain

We will now discuss, how predictive coding and the Bayesian brain are related. As we saw previously, predictive coding can be understood as an algorithmic representation of cortical processing

(Aitchison and Lengyel, 2017). The idea of the Bayesian brain is, that the brain uses a probabilistic framework and performs Bayesian inference for perceptions about the world. It is possible to combine these two ideas (Aitchison and Lengyel, 2017) with the following rationale: Predictions are the integral of posterior of latent states of the world. That is, the predictions are directly represented in the posterior $p(\Theta|D)$. Following the saying "the posterior of today is the prior of tomorrow", one can use the posterior as a prior distribution for the next time step³. The amount of updating between the prior and the posterior (that is, the difference of the likelihood and the prior) is the prediction error. In other words, one has an expectation, a prediction about the future (the prior), incoming information is evaluated in the generative model (likelihood), there is potentially a mismatch between prior and likelihood (prediction errors) and only this prediction error needs to be known to the higher level area (which produced the prior) to calculate the new expectation (the posterior, or prior of the next step). Such a scheme might seem unlikely to be represented directly in the brain (see e.g. Kwisthout and Rooij, 2013, for a discussion on possible intractabilities), but recent modeling work with spiking model networks show that predictive coding in the Bayesian brain is possible (Boerlin and Denève, 2011). One last point: though it seems quite reasonable to combine predictive coding and the Bayesian brain, one could have predictive coding in the brain without Bayesian principles (Aitchison and Lengyel, 2017) and vice versa. In chapter 3 and 4 I discuss two experiments that make use of the predictive coding framework and which show how predictions and expectations are a hierarchical system (Ehinger, König, et al., 2015) and influence our perception (Ehinger, Häusser, et al., 2017).

³It is usually not made explicit what timescale is meant. Posterior updating can happen over many different timescales

1.3.2 Prediction and action

In recent years, several theories have emerged that ultimately try to combine the predictive coding theory with the Bayesian brain in order to explain phenomena discussed in the philosophy of mind. Especially enactivism (E Thompson and Varela, 2001), the extended mind (Clark, 2013), and developments based on the sensory-motor account of visual perception (Seth, 2014; O'Regan and Noë, 2001) are making use of this "predictive brain". What is strictly common in all of these accounts is the focus on actions (see also AK Engel et al., 2013). Indeed, in the popular Bayesian brain hypothesis, popularized by Friston (K Friston, J Kilner, et al., 2006; K Friston, 2010; McGregor et al., 2015), there is only a conceptual distinction between actions and perceptions.

Eye movements are the perfect example for this, as they represent actions and completely determine the input for visual perception. K Friston, Adams, et al. (2012) suggest that percepts are hypotheses, and that "saccades [can be used] as [perceptual] experiments" to test these hypotheses. As I will show in the next paragraphs, using eye movements as actions that test hypotheses (percepts) has many benefits but also challenges. Some of the shortcomings I will address in a newly developed paradigm in chapter 5 (see also Ehinger, Kaufhold, et al., 2018). The primary requirement, actual prediction errors over eye movements, are tested in chapter 4.

1.3.3 Eye movements as perceptual experiments

In the following I want to show my motivation to use eye movements in several of my works. Besides the above mentioned enactive theoretical foundation, there are many pragmatic reasons

why one should be interested in eye movements. As shown in the beginning of the chapter, by moving the eye, the retina can sample a large part of visual space in a short time. Moving the eye is a necessity, as only a small part of our visual scene can be perceived in great detail (at the fovea) with everything else blurred to some degree (the peripheral vision). Because eye movements are the most common action we perform throughout our lives, even more often than our heart beats, they are likely quite optimized for optimal decisions where and when to look next.

Eye movements have many beneficial properties that allow us to understand the brain in great detail. First, eye movements are essential for many different processes. For example the control of eye movements have large influences on language processing and reading (Rayner, 2009), attention (Shepherd et al., 1986), eye movements are impaired in psychiatric disorders Hutton and Ettinger, 2006 and as we show are a prominent decision making process (Ehinger, Kaufhold, et al., 2018). Second, they are the prime example of the overt attentional system. Where and how long we look at things usually implies uncertainty of the respective foveated stimulus and attentional focus. And third, as already mentioned, they occur in high number. Because each eye movement gives us a new data observation, the high rate allows us to average over many trials and oppose the low signal-to-noise ratio of e.g. EEG measurements (see chapter 3).

1.3.4 Predictions over eye movements

We constantly move our eyes and thus, the visual snapshot of the world. Nevertheless, we have a very stable percept of the world. One mechanism that could explain this stability is based on predictive remapping. That is, there is a temporal prediction component in the visual cortex that allows us to integrate current

peripheral with future foveal content (Crapse and MA Sommer, 2008). Indeed, trans-saccadic remapping has been observed in behavior and physiology (Rolfs et al., 2011; Melcher, 2008; He et al., 2018). Recent evidence suggests, that this remapping process occurs only for attended locations and not the whole visual field (Szinte et al., 2018). Should forward mapped prediction be violated, then we would expect larger prediction errors to occur (assuming predictive coding). It has not been shown, whether prediction errors over eye movements actually occur. We conducted an experiment (Chapter 3) to find out whether prediction errors due to violation of predictions over eye movements can be elicited and recorded in the human EEG. We introduced an additional interaction about how the peripheral prediction was formed, either veridically through peripheral information, or through fill-in in the blind spot a, by definition, unreliable prediction.

1.3.5 Predictions based on fill-in in the blind spot

One way to investigate the generative models and possible prediction errors is offered by the physiological blind spot. As introduced before, the physiological blind spot covers the area of the eye where no photoreceptors exist, which nevertheless has a percept. This percept is called fill-in. The process that generates this fill-in is still unclear, but the most likely explanation is based on feedback signals from extrastriate cortex (Komatsu, 2006). If we combine a filled-in percept with the idea of predictive coding, this would mean that we can have predictions based on filled-in content.

In chapter 3, I will discuss an experiment that tries to answer the question whether we make use of those internally generated,

filled-in predictions in a different way, than normal predictions based on veridical information.

1.4 Randomization and the quasi-experimental nature of eye movements

"Correlation does not mean causation". But science would be quite uninteresting if we could not make statements on the mechanisms, the causes of the world. In order to make causal statements, one needs the additional procedure of randomization of subjects and experimental control over the independent variables of interest (for a formal causal calculus see the *do-calculus* by Pearl 2009, see also Cook et al. 2002). In many experimental designs, it is possible to generate factuais and counterfactuals, for example "intervention" and "no-intervention" (control). If all other factors which could possibly influence the conclusions are appropriately randomized, one can conclude causal relations (Cook et al., 2002). Of course, factuais and counterfactuals again imply that one can assert control over the system to be studied. It is not easily possible to do this in every system. In cases where it is not possible, only quasi-experimental designs can be used (Campbell and Stanley, 1963) and strong assumptions are needed to argue for causality. For instance in unrestricted viewing experiments, eye movements are not controlled by the experimenter. That means there is no direct control over fixation duration, saccade amplitude, fixation location or the geometry of the trajectories. That is, whenever we try to predict an experimental outcome by one of these variables, the results might be biased to some extent because the independent variable was not controlled for and there could be additional confounders present. For instance, the causal statement: "Longer fixation durations cause higher accuracy in a search task" is not a valid causal inference. There could be a hidden process that both prolongs fixation durations and cause higher accuracy at the same time. We would need to force subjects to look at certain search tar-

gets for a experimentally controlled time, thereby experimentally controlling fixation duration, to make such inferences. To conclude: In order to study causal relations of unrestricted eye movements we need new experimental designs that allow for experimental control of the saccade parameters of interest. In chapter 5 we will see a new paradigm that can be used exactly for this.

1.5 Ways to study causal interactions: Optogenetics

Transferring these thoughts to the brain, it becomes clear that if one can only experimentally control behavioral factors, it is difficult to assert the exact causal mechanisms of the ongoing neural processes. This situation could be improved enormously if we could act upon certain areas or cells of the brain directly. A game-changer in this regard has been optogenetics, first introduced in 2005 (Boyden et al., 2005; X Li et al., 2005). When shining light at optogenetical proteins, they can open or close ion channels and thereby activate or silence single neurons or whole areas. The impact of optogenetics on biology and especially the field of neuroscience is immense (Boyden, 2015; Fenno et al., 2011; Pama et al., 2013).

In order to use optogenetics, a number of different naturally light-active proteins (like the opsins in the retina) can be chosen from. Optogenetic agents are usually inserted in cells using a viral vector. This can become problematic when one needs to both silence and activate neurons. In such a case, two different opsins are needed. Two viral vectors are commonly used, but the chance that a cell expresses both proteins at the same time is small. Therefore new opsins are currently developed that combine both silencing and activation attributes depending on the wavelength of the light. One such protein, Melanopsin, naturally occurs in photosensitive retinal ganglion cells in the retina. Melanopsin is a g-coupled receptor that is silenced by reddish light, and activated by bluish light. Because daylight contains large proportions of blue light and the sun turns red in the evening, the direct connectivity to the hypothalamus and the influence on our circadian rhythm seems reasonable (Hankins et al., 2008; Do and Yau, 2010).

Besides that Melanopsin can activate and deactivate a cell depending on the wavelength of the activating light, it also stays activated for up to half an hour once activated. It can be therefore understood as a cellular "switch". We were recently involved in a study to systematically study the behavior of Melanopsin (Spoida et al., 2016) and developed a Bayesian generative model of the kinetics of Melanopsin as described in chapter 6.

1.6 Overview over this thesis

In the next chapter I will discuss a MEG study on how we form new categories from repeated exposure (Kietzmann, Ehinger, et al., 2016). In the third chapter, I will discuss a study on predictive coding, prior knowledge and the physiological blind spot (Ehinger, Häusser, et al., 2017). The fourth chapter directly addresses the question of prediction in a context of eye movements as actions using EEG (Ehinger, König, et al., 2015; König et al., 2016). The fifth chapter describes a study on information uptake, the relation to eye movements and experimental control of fixation durations (Ehinger, Kaufhold, et al., 2018). The sixth chapter introduces the methodological and statistical framework of Bayesian generative models based on the example of Melanopsin's kinetics (Spoida et al., 2016; Ehinger, Eickelbeck, et al., 2016).

Experience shapes computation: Learning new categories

” *The behavioral data indicates that the monkeys had reached a high level of categorization performance*

— **Reviewer #3**

Contributions

Neuroimage: Kietzmann, Ehinger, Porada, Engel, and König
2016

TCK, **BVE**, DP, AKE and PK conceived the study. TCK, **BVE** and DP recorded the data. TCK, **BVE**, DP and PK performed the analyses. TCK, **BVE**, AKE and PK wrote the manuscript.

Chapter quotes are taken out of context. I'm deeply grateful for the reviewers' time and work. Most of their comments were right on point and very helpful.

2.1 Layman's summary

Throughout our lives, we have to learn new categories. These categories help to make sense of the world, and offer us a way to organize our sensations. In our study, we let participants learn two new categories based of faces and measured their brain activity before, during and after learning.

At the beginning of learning, activity related to category recognition occurs in frontal regions of the brain usually associated with cognitive control, planning and rule-based learning. Only after a long learning period, we found this signal much earlier and in temporal areas of the brain. These areas are usually associated with stimulus feature analysis. Our study helps to resolve the mystery, of why some studies find frontal category learning and others temporal category learning. Our data suggest that in the beginning of learning, a flexible category extraction is used based on frontal areas. An example would be the difference between birds and mammals: they can be distinguished simply by knowing the fur to wings ratio. The second category learning type is more exhausting to learn, but arguably more automatic. A bird expert would probably recognize the differences between two bird species instantaneously and most likely, he will be one of a few people that can do it at all - he has had much more exposure to category examples. Thus the features of the category difference seem to be ingrained into his visual system. Only due to our very long training time, we were able to show that the computation of category extraction changes.

2.2 Extensive training leads to temporal and spatial shifts of cortical activity underlying visual category selectivity

The human visual system is able to distinguish naturally occurring categories with exceptional speed and accuracy. At the same time, it exhibits substantial plasticity, permitting the seamless and fast learning of entirely novel categories. Here we investigate the interplay of these two processes by asking how category selectivity emerges and develops from initial to extended category learning. For this purpose, we combine a rapid event-related MEG adaptation paradigm, an extension of fMRI adaptation to high temporal resolution, a novel spatiotemporal analysis approach to separate adaptation effects from other effect origins, and source localization. The results demonstrate a spatiotemporal shift of cortical activity underlying category selectivity: after initial category acquisition, the onset of category selectivity was observed starting at 275ms together with stronger activity in prefrontal cortex. Following extensive training over 22 sessions, adding up to more than 16,600 trials, the earliest category effects occurred at a markedly shorter latency of 113ms and were accompanied by stronger occipitotemporal activity. Our results suggest that the brain balances plasticity and efficiency by relying on different mechanisms to recognize new and re-occurring categories.

2.3 Introduction: Category Learning

One of the most essential tasks of our visual system is to make sense of the complex signals it receives from the world around us. A central aspect of this is the ability to group objects into various categories, allowing for considerable simplification, generalization and supporting higher cognitive function. To advance our understanding of the underlying cortical mechanisms, a large body of experimental work focuses on temporal aspects of category selectivity, asking for the earliest point in time at which category information is extracted. As a result, we now have ample psychophysical and electrophysiological evidence that naturally occurring categories can be extracted in only little more than 100ms of processing (Sugase et al., 1999; J Liu et al., 2002; H Liu et al., 2009; Hung et al., 2005; Kirchner and Thorpe, 2006; Carlson, Hogendoorn, et al., 2011; Carlson, Tovar, et al., 2013; Cichy et al., 2014). However, apart from the necessity for fast and robust categorization of re-occurring categories, our ever-changing environment poses the additional challenge to retain considerable plasticity in order to support the rapid learning of entirely novel categories. Here, the study of naturally occurring categories provides only limited possibilities, as it focuses on categories with which we already have extended experience (for instance, all of us can be considered face- and house-experts, as these categories play a vital role in our everyday life). It therefore remains an open question, how cortical representations and the temporal dynamics of category selectivity develop from the initial learning of a category towards category expertise. To elucidate this issue, we performed a longitudinal study in which we investigated the impact of extended category training of two artificial visual categories in a parametric feature space on the visually evoked responses using a rapid event-related magnetoencephalography (MEG) adaptation paradigm. Adaptation paradigms, also known

as repetition-suppression and repetition-enhancement (Krekelberg et al., 2006; Segaert et al., 2013), are widely applied in the field of functional magnetic resonance imaging (fMRI; see Grill-Spector and Malach, 2001, for an adaptation review) and offer the advantage to bypass the limited spatial resolution of any imaging method by focusing on response-changes in neuronal subpopulations which are measurable in the average response of a pre-defined region of interest (ROI). Adaptation paradigms therefore have the potential to reveal differences in neuronal selectivity that would remain unnoticed in more traditional univariate designs. While the limits of spatial resolution are even more drastic in case of MEG/EEG, these methods offer the possibility to investigate cortical processes with high temporal resolution. A combination of a rapid, event-related MEG adaptation paradigm with perceptual category training is therefore a promising candidate to resolve changes in the temporal aspects of category processing, indicative of changes in the underlying cortical mechanism. An additional advantage of our longitudinal paradigm is a control for effects of low-level stimulus properties. Data recorded during a baseline session allowed us to exclude the possibility that potentially found category effects are an inherent low-level property of the utilized feature space. Such differences in low-level statistics have previously lead to considerable challenges in the interpretations of category effects in studies using naturally occurring categories (Thierry et al., 2007; Rossion and Jacques, 2008; Crouzet and Thorpe, 2011; VanRullen, 2011). We investigated the emergence and development of category selectivity by recording MEG data in a baseline session, prior to any category training, a second time after five training sessions, and a third time after extensive category training in 22 training sessions. Category selectivity was estimated by comparing the visually evoked responses to stimuli that were either preceded by a different adaptor stimulus from the same category (category-internal), or by an adaptor stimulus of a different category (category-external), while holding low-level stimulus differences constant. To analyze

the high-dimensional MEG data, a novel spatiotemporal analysis approach was employed. Building on the observation that true adaptation effects should occur in the same cortical regions activated previously by the adaptor stimulus, the analysis exploits the linear additivity of MEG sources in order to explicitly separate experimental effects, i.e. differences between category-internal and -external conditions, into adaptation- and other, non adaptation-related effects. In the adaptation phase of the experiment, each trial consisted of two stimuli: an adaptor and a repetition stimulus. Interpreting each MEG topography as a high-dimensional vector, the cortical response to the second stimulus can be understood as a linear combination of (a) a re-activation of the regions previously responding to the adaptor stimulus, as required for adaptation, and (b) other cortical regions, activated uniquely during the processing of the second stimulus. This led to the insight that the response to the first, adaptor stimulus can be used to decompose experimental effects, observed in the second response, into adaptation-based and non-adaptation effects. Importantly, the involved vector projection maps the 271 dimensional sensor space onto a single, yet highly informative subspace and thereby avoids problems of multiple comparison (see Methods and Materials for details). Using this approach to focus on effects driven by adaptation we observed a temporal shift in category selectivity from a latency of 275ms after initial category acquisition to only 113ms following extensive training. This speedup suggests a marked change in the cortical network mediating the categorization of visual input. Indeed, source analysis revealed an anterior-to-posterior shift of cortical activity from initial to extensive category training. While the time-window of category selectivity found after five training sessions exhibited stronger activation in the prefrontal cortex (PFC), the early category effects found after 22 training sessions showed increased activation in occipitotemporal regions. Previous theories on visual categorization viewed either PFC or regions in the ventral stream as the origin of category selectivity. Our findings now reconcile these

contrasting views by suggesting that both processes are likely to contribute to categorization at different stages of category learning. While PFC is involved in the categorization of rather novel and dynamic categories, extensively used categories seem to obtain a privileged status and are resolved faster relying more heavily on cortical resources in occipitotemporal cortex.

2.4 Methods: The MEG adaptation paradigm

2.4.1 Participants

Nine healthy, right-handed subjects (five female, aged 19-30) participated in the study. All subjects had normal or corrected-to-normal visual acuity, were naïve to the purpose of the study and gave written informed consent to participate. The experimental procedures were approved by the ethics committees of the University of Osnabrück and the Ärztekammer Hamburg. Each subject participated in a total of 23 experimental sessions (one baseline session and 22 training sessions). MEG data were recorded during the first baseline session, as well as after training sessions five and 22. The MEG recording from subject nine in session 22 was excluded from the analyses due to excessive noise in the data.

2.4.2 Stimulus Space

Similar to previous work with macaques (Sigala, Gabbiani, et al., 2002) and humans (Reed and MP Friedman, 1973; Nosofsky, 1991; Sigala, Gabbiani, et al., 2002; Kietzmann and König, 2010), category training was based on two artificial categories of Brunswik faces (Brunswik and Reiter, 1938), defined in a four-dimensional, parameterized stimulus space (Figure 2.1), also known as a factorial morphspace (Goldstone, Steyvers, et al., 1996; Gureckis and Goldstone, 2008; JR Folstein, Palmeri, and Gauthier, 2012). Two of the dimensions were category-relevant (eye height and eye separation), while the two others (mouth height and nose length) were assigned pseudo-randomly, ensuring that no stimulus clusters of the same category existed that could potentially render

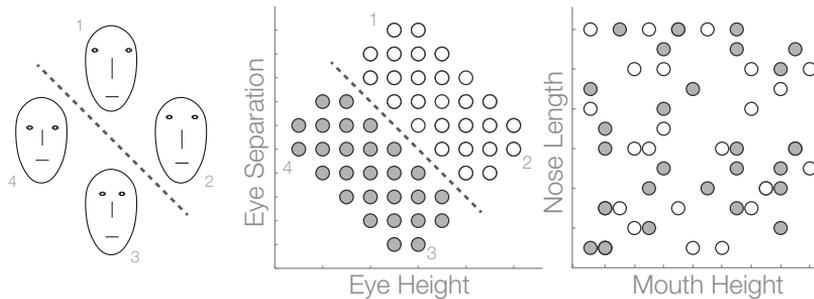


Fig. 2.1 Stimulus Space. Subjects were trained to distinguish two artificial categories of faces, defined in a four-dimensional parametric feature space. Two dimensions were category-relevant (eye height and eye separation), and two were irrelevant (nose length and mouth height). No single feature was decisive for the category of a given face, only the combination of features allowed for correct categorization. The category boundary was rotated by 90 degrees for every other subject.

these task-irrelevant dimensions informative. A linear category boundary split the category space of the two relevant dimensions in half, such that no single stimulus property was decisive for the category membership. This design is optimized to search for effects of category selectivity, since no linear re-weighting of singular input dimensions will lead to optimal training performance. The overall stimulus space consisted of 60 stimuli, six of which defined the respective category boundary and were not included in the training and testing. The final two categories comprised 27 stimuli each. Moreover, the category boundary was rotated by 90° for every other subject. The subjects were at no point in time instructed about the design of the stimulus space or the relevant category dimensions. Post training, no participant was able to verbally describe the relevant category rule. All stimuli shown during training and the MEG adaptation sessions were presented using the Psychophysics Toolbox 3 (Brainard, 1997; Kleiner et al., 2007) running under Matlab 2010a.

2.4.3 Category Training

In order to allow subjects to learn the two artificial categories of faces, they received category training in a total of 22 sessions with 756 trials each. Here, we largely followed our previous procedure (Kietzmann and König, 2010). In each training trial, subjects were presented with a single stimulus and were then asked to categorize it as either category A or B with their index- or middle-finger. Auditory feedback was provided as training signal. A high-pitch tone indicated a correct response, whereas a low-pitch tone and a forced break of two seconds indicated an incorrect response. To prevent a fixed association between the category membership and motor response, the finger used to indicate the category decision was switched three times across each training session. The subjects were notified of the switches.

2.4.4 Rapid Event-Related Adaptation Paradigm

To estimate the time-course of electrophysiological category effects, we used a rapid event-related MEG adaptation paradigm. This approach is similar to the more common fMRI adaptation (Grill-Spector and Malach, 2001) or repetition suppression / enhancement, and has only recently been introduced to the field of EEG and MEG (Marinkovic et al., 2003; A Harris and Nakayama, 2007; Caharel et al., 2009; Vizioli et al., 2010; Zimmer and Kovács, 2011; Huberle and Lutzenberger, 2013; Scholl et al., 2014). While fMRI adaptation paradigms are traditionally associated with effects of repetition suppression, repetition enhancement is now commonly observed across a wide variety of cortical regions (Krekelberg et al., 2006; Segaert et al., 2013). Especially for experiments investigating adaptation effects across time, a temporal sequence of early enhancement and late suppression has been reported (Marinkovic

et al., 2003; Petit et al., 2006), in line with the prediction of the more recent accumulation model of adaptation (James and Gauthier, 2006). Taken together, the repeated activation of neuronal populations can either lead to a suppressed or an enhanced response amplitude in fMRI and MEG/EEG.

During the rapid MEG adaptation paradigm applied here, participants performed a delayed match-to-category (DMC) task, indicating whether two subsequently presented stimuli belonged to the same or a different category (Figure 2.2a). To test for effects of neuronal adaptation, i.e. repetition enhancement and repetition suppression, we analyzed the magnetic fields evoked by the second stimulus, when either preceded by a different adaptor stimulus from the same category or an adaptor from a different category. In total, 432 trials were recorded per session and subject, including 216 trials with a category-internal and 216 trials with a category-external adaptor stimulus. The sequence of category-internal and category-external conditions was randomized across trials. To control for low-level feature differences in the two adaptation conditions, all category-internal and category-external trials were matched in distance and direction in the two relevant dimensions of feature space (Figure 2.2ab). This has the additional advantage that no linear re-weighting of the category-relevant dimensions can account for category selectivity (Goldstone, Lippa, et al., 2001; JR Folstein, Palmeri, and Gauthier, 2012), because all category internal and external adaptation trials will be affected likewise. Our setup is therefore tuned for observing effects of category selectivity, rather than attentional re-weighting of (single) features, required for later category extraction. During the randomized adaptation trials, a fixed mapping of experimental condition to motor response was prevented by switching the target keys for the two answers after half of the experiment. The structure of the adaptation trials was as follows. First, a fixation cross was presented for 800ms with an SOA of 100ms. Then, a first stimulus was presented for 500ms,

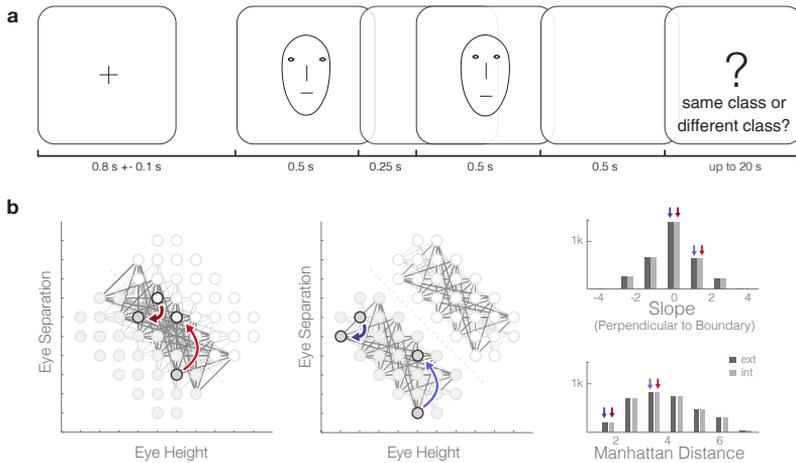


Fig. 2.2 Adaptation Paradigm. To test for electrophysiological correlates of category information, an adaptation paradigm was used. Each trial either crossed the category boundary (category-external) or stayed within a category (category-internal). (a) Temporal sequence of an adaptation trial. (b) Low-level properties of the category-internal and category-external adaptation trials were controlled by matching the distance and slopes of the corresponding stimulus-pairs. Exemplary trials are highlighted in color (category-external in red, category-internal in blue).

followed by an inter-stimulus-interval of 250ms. The second stimulus was again shown for 500ms. Finally, after an additional delay of 500ms a question mark was displayed on the screen, indicating to the subject that a response can be given (Figure 2.2a). This timed response was introduced to keep presentation of the second stimulus free of cortical activity related to the motor-execution.

2.4.5 MEG Acquisition

MEG data were recorded in a baseline session, prior to category training, as well as after five and 22 training sessions. The selection of five and 22 training sessions was based on previous work using a similar feature space, in which subjects were able to perform at >90% accuracy after only five training sessions (Kietzmann and König, 2010), while exhibiting high-level category

effects only after prolonged training. MEG data were acquired at 1200Hz using a CTF whole-head system with 271 axial gradiometers (CTF275, VSM MedTech). The position of the participants head was continuously recorded using three head localization coils (NAS/LPA/RPA). Moreover, a bipolar electrocardiogram (ECG) and an electrooculogram (EOG) with three channels were recorded. The EOG electrodes were placed below the eyes and on the forehead. The reference was positioned on the tip of the nose. The experimental stimuli were back-projected on a screen with a LCD projector (Sanyo XP51) at 60Hz refresh rate. The presentation distance was 60cm, leading to a display size of $2^{\circ} \times 3.3^{\circ}$ of visual angle.

2.4.6 Data Analyses

All analyses were performed using custom code in Matlab R2010a (Mathworks, Natick, MA, USA), fieldtrip (Oostenveld et al., 2011) and Brainstorm (Tadel et al., 2011).

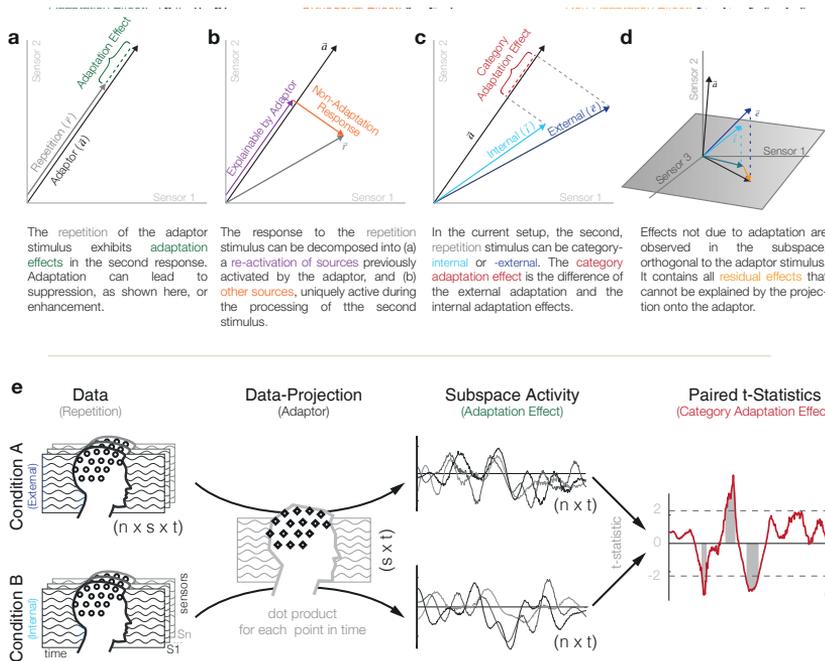
2.4.7 Preprocessing

After downsampling the data to 600Hz, artifacts due to muscle activity, sensor-jumps and extreme noise were first detected automatically using fieldtrip, followed by manual cleaning of the data. To account for sensor drifts, the data were high-pass filtered at 1 Hz. Moreover, frequencies above 100Hz and the artifactual frequency bands around 50Hz and 60Hz were excluded using a zero-phase Butterworth IIR filter. To remove eye-related and cardiac artifacts from the data, we used an automated procedure based on an independent component analysis. The underlying algorithm relies on a correlation-based and a weight-based artifact-metric computed for each independent component. Components surpassing a selected threshold were labeled as artifacts and re-

moved from the data. The optimal thresholds were determined automatically based on a receiver operator characteristic (ROC) analysis applied to a subset of the data for which two experts had classified components as artifacts. The resulting algorithm was able to detect 98.1% of the components tagged by the experts, with only 0.3% false positives (see supplementary material in Kietzmann and König, 2015). Without making use of additional eye-tracking data, our approach reaches performance levels of a state of the art algorithm for automatic artifact removal that require ground-truth eye-movement data (Plöchl et al., 2012). Finally, although our fully randomized design prevents systematic effects of head-position, we removed any residual effects from our data. We first extracted a six dimensional description of the head position and direction from the simultaneously recorded localization coils (NAS/LPA/RPA) and used this to regress out the effects of head-position (Stolk et al., 2013). All evoked potentials were baseline-corrected with respect to the 700ms fixation period prior to the presentation of the first stimulus.

2.4.8 Spatiotemporal Projection Approach: Separating effect sources

The current experiment makes use of a rapid event-related MEG adaptation paradigm. In each trial, two stimuli are presented: one adaptor and subsequently a second, repetition stimulus. Two experimental conditions are compared. The second stimulus can either be of the same category as the adaptor, or of a different category. Differences between these two conditions thus indicate category selectivity. Common to every adaptation paradigm, experimental effects can either originate from true adaptation, i.e. the differential re-activation of category selective regions (Figure 2.3a), or from other sources that are uniquely activated during the presentation of the second stimulus. As an example of the



The repetition of the adaptor stimulus exhibits adaptation effects in the second response. Adaptation can lead to suppression, as shown here, or enhancement.

The response to the repetition stimulus can be decomposed into (a) a re-activation of sources previously activated by the adaptor, and (b) other sources, uniquely active during the processing of the second stimulus.

In the current setup, the second, repetition stimulus can be category-internal or -external. The category adaptation effect is the difference of the external adaptation and the internal adaptation effects.

Effects not due to adaptation are observed in the subspace, orthogonal to the adaptor stimulus. It contains all residual effects that cannot be explained by the projection onto the adaptor.

Fig. 2.3 Spatiotemporal Projection Approach - Separating effect sources. A spatiotemporal projection approach was used to decompose the MEG response to the second stimulus, and thereby category effects, into parts due to a re-activation of the regions previously responding to the adaptor stimulus and activity in other cortical regions. (a) At any given point in time, the MEG activity pattern can be interpreted as a vector in high-dimensional space (for clarity, we here show a 2-dimensional example, without loss of generality). Adaptation effects are typically expected to lead to an altered amplitude in the second, repetition response, compared to the first, adaptor response. (b) The response to the second stimulus, however, is a linear combination of a repeated activation of the previously activated sources (as required by the adaptation paradigm), and other sources (e.g. response preparation, etc.). (c) Experimental effects can be decomposed in a similar fashion. To detect effects based on adaptation, the experimental conditions are projected onto the adaptor vector. (d) Non-adaptation based effects are found in the residual activity of both conditions, which reside in a subspace orthogonal to . If no sensible projection target can be defined a priori, a test can be performed for each sensor and time point, adjusting accordingly for multiple comparisons. (e) Adaptation-based effects are investigated across time, based on the projection-approach yielding activity traces for each subject, condition and session. These traces are subject to statistical analyses, highlighting temporal candidate clusters that show significant differences between category-internal and category-external trials. These candidates are then subject to subsequent analyses. Abbreviations: n, trial number; s, sensor number; t, time point.

latter, if category-internal and category-external conditions were reported unbalanced, via different hands, condition-dependent differences in the activations of the two motor-cortices would be expected. Such effects, despite originating from category-related signals, would not be due to adaptation (please note that this example is for illustration purposes only, as the current design balanced different motor responses across experimental conditions). In order to separate these different adaptation and non-adaptation effect sources, we here employ a spatiotemporal analysis approach that decomposes the MEG signals of the second stimulus into parts that are due to the re-activation of regions previously involved in processing the adaptor, and parts that cannot be accounted for by re-activation (Figure 2.3b). Effects due to adaptation. To focus on adaptation effects at a given point in time, we project the high-dimensional MEG response vector of the second stimulus (\vec{r}) onto the normalized adaptor response vector (\hat{a}):

$$\vec{r}_{\hat{a}} = (\vec{r} \cdot \hat{a})\hat{a}$$

The adaptation-based category effect (ace) is then computed as the difference in amplitude between projected category external (\vec{e}) and internal (\vec{i}) response vectors (Figure 2.3c), yielding a scalar estimate of the adaptation-based category effect:

$$ace = |\vec{e}_{\hat{a}}| - |\vec{i}_{\hat{a}}| = \vec{e} \cdot \hat{a} - \vec{i} \cdot \hat{a}$$

Applied for every point in time, this projection yields one-dimensional effect traces for every participant and session, which are subsequently subject to statistical analyses (Figure 2.3e).

To define the projection vector, \vec{a} , we here chose to use the group average response, evoked by the first, adaptor stimulus, as recorded in the baseline session (time window between 0 and 300 ms after stimulus onset, low-pass filtered at 35 Hz using a zero-phase Butterworth IIR filter, Figure 2.4). Using the same projection

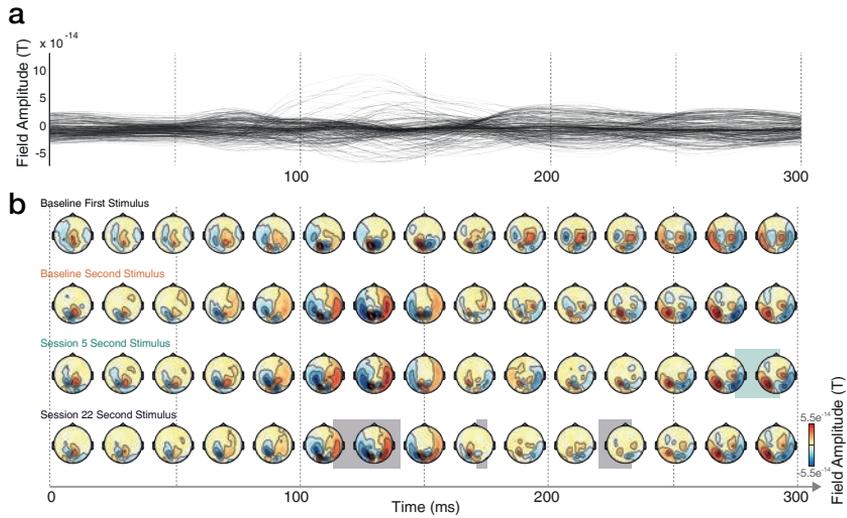


Fig. 2.4 Grand average responses to adaptor and repetition stimulus. (a) Butterfly plot of the first, adaptor stimulus. (b) Grand average response topographies of the first, adaptor stimulus during the baseline session (as later used for data projection), and for the second, repetition stimulus, shown separately for each session.

vector as basis for effect estimation has the advantage that it allows for comparisons of effect sizes across experimental sessions. If different projection vectors were used for every session, it would introduce unnecessary ambiguity as possible changes in effect-sizes might be merely due to changes in the underlying projection vector. To ensure that the current approach is appropriate, we performed a non-parametric cluster test based on an F-statistic in which we compared the responses to the first stimulus across all three sessions (baseline, session five, session 22) within the first 300ms of processing (the cluster-threshold was set to $p < 0.05$, cluster inclusion was at $\alpha < 0.05$). No significant differences were found, indicating that the same projection vector can be used across sessions. Finally, to be able to compare effects across time, the projection vector (\vec{a}) was normalized to unit length (\hat{a}) at each sample.

Non-adaptation effects. While the projection approach is straight-forwardly applicable to adaptation effects, as the adaptor stimulus uniquely defines the projection target, non-adaptation based effects can occur independently, rendering a projection approach inapplicable: given adaptation-based effects have been accounted for via projection, such residual effects can occur anywhere in the space, orthogonal to the adaptation vector. Thus, to test for such residual, non-adaptation effects, we statistically compared the previously unexplained parts of the external (\vec{e}) and internal (\vec{i}) vectors (shown as a projection onto the orthogonal plane in Figure 2.3d), which reside in the original 271-dimensional sensor space.

In summary, we use the response to the first, adaptor stimulus to decompose the response to the second, category-external or category-internal, stimulus into two possible sources: adaptation-based effects, and other, non-adaptation effect sources. In contrast to this, more traditional analyses of adaptation paradigms leave the origin of observed effects ambiguous, and thereby severely complicate their interpretation. In addition to this important advantage, the current projection method has further statistical benefits because adaptation-based effects can be investigated using a one-dimensional projection-signal, created from the original 271 dimensional data. This circumvents the problems of multiple comparisons occurring when all sensors are considered individually. Traditionally, this problem is solved by either selecting sensors of interest a priori (Rossion and Jacques, 2008), or alternatively by testing all sensors individually and afterwards controlling for multiple comparisons, for instance by applying non-parametric cluster-based correction methods (Maris and Oostenveld, 2007; Ehinger, König, et al., 2015). Unfortunately, both options are not without problems. An a priori, fixed selection of sensors is problematic, if it is unclear which sensors should be selected or, even more so, if sensors of interest change over time. The second solution, testing

all sensors and time-points individually while applying a cluster-based correction, provides the liberty of observing effect clusters anywhere in space and time, but at the cost of much decreased statistical sensitivity. These limitations are overcome by the current projection method. Another statistical consideration worth noting is that the data used for the projection vector are independent of the experimental data in question (comparing category-internal and category-external responses in the second stimulus). This avoids the (spatial) dangers of double-dipping in neuroimaging (Kriegeskorte, Simmons, & Bellgowan, et al., 2009). Notably, the current projection method is directly applicable to fMRI adaptation paradigms. In fMRI, the traditional use of ROIs, too, exhibits the problem that observed effects are intermixed and therefore cannot be unambiguously attributed to mechanisms of adaptation. That is, effects observed can either originate from adaptation, or other effect sources. A further benefit of our approach, exploited later in this paper, is that the effect decomposition allows for efficient source localization of MEG data. Effects due to adaptation can be interpreted as a differential activation of the regions contributing to the first, adaptor stimulus. Therefore, to source localize the effect, the data from the adaptor stimulus can be used. Put differently, the adaptor functions as a spatiotemporal localizer that explicitly focuses the analysis on stimulus-repetition effects. Thus, if an effect is found via projection, the same localizer determines the underlying sources. This approach resembles the standard methodology used for analyses based on independent components, for which effects are first investigated based on component activations and localized based on the component topography (Makeig et al., 2002; Pockett et al., 2007; Ehinger, Fischer, et al., 2014; Tsai et al., 2014). The assumption of our, and in fact any localizer approach, is that the same cortical processes are active during the trials used to define the localizer and the experimental trials of interest. While many experimental settings meet this assumption, adaptation paradigms are particularly suited for this approach. This is due to the fact

that they already presuppose the same cortical regions to be active during the processing of the first and second stimulus. The use of the first-stimulus response as a spatiotemporal filter for the evoked response to the second is therefore simply a consequent translation of the experimental paradigm to the analysis methodology. It should be noted that the projection approach, described so far, relies on the assumption that similar cortical regions are active at comparable latencies, because and are taken from the same point in time. If responses exhibit significant temporal shifts effects can potentially be missed. To partially counter this effect, we here low-pass filtered the adaptor stimulus, resulting in a temporally more robust fit. As a theoretical alternative, one could use the full response matrices (sensor x time) of adaptor and repetition to explicitly test for effects at all possible delays. Instead of an effect trace, this yields an effect matrix (M):

$$M = R \cdot A'$$

The diagonal of this matrix corresponds to a zero-delay and therefore to the effect traces used here. While able to detect effects at different latencies, this approach introduces a quadratic increase in the number of tests. The required corrections for multiple comparisons thereby decrease overall sensitivity.

2.4.9 Statistical Analyses

For statistical analyses of the adaptation effects, yielded through our spatiotemporal projection method, we computed the one-dimensional response traces for every participant, condition, and session (Figure 2.3e). Based on these signals, we then tested for training-induced category effects, following a two-staged approach. First, we temporally localize time-windows of interest, i.e. time-windows exhibiting significant category effects ($p < 0.05$) by performing a paired t-test at every point in time, contrasting

category-external and category-internal trials. Following this, we investigated, whether the observed category selectivity was indeed the result of category training. To accomplish this, we tested, for each candidate time-window, whether the respective effect is significantly larger after training compared to the difference observed in the baseline session (training interaction). This was accomplished by estimating the corresponding interaction effect size and its 95% confidence intervals. Corrections for multiple comparisons across time were performed at this final stage by applying a Bonferroni correction at the cluster-level, i.e. by enlarging the 95% intervals according to the number of clusters tested in each session. As a result of this statistical procedure, any cluster reported in the following will not only have shown significant category effects, but also a significant training interaction, verifying that the found effects are indeed caused by category training. Testing for a training interaction is an important additional prerequisite in investigations of developing category selectivity, as observed differences between category-internal and category-external conditions could also be an inherent property of the selected feature space and not the result of category training. This possibility is ruled out by the statistical procedure described. Summing up, we focused on adaptation-based effects by combining a spatiotemporal projection method with rigorous statistical analyses. This allowed us to overcome the need to use heuristics in selecting sensors and time-windows of interest, while controlling for multiple comparisons in space and time. The only free parameter of the overall approach for finding adaptation-based effects is the p-value for the selection of temporal candidate windows, which was selected to be $p < 0.05$. To test for non-adaptation effects, we performed a two-sided t-test for every point in time and space, and corrected the family wise error rate using a nonparametric, cluster-based permutation test (cluster inclusion threshold $\alpha < 0.05$, left- and right-sided cluster $p < 0.025$, respectively). This approach can find unpredicted effects, but at the cost of decreased statistical sensitivity. All analyses performed,

adaptation and non-adaptation, focus on the first 300ms of processing after stimulus onset, as this time-window approximately resembles typical fixation durations during free-viewing of natural scenes (Underwood et al., 1998).

2.4.10 Behavioral Relevance

To estimate the behavioral relevance of the observed category effects, we contrasted the effect size of adaptation trials in which the response of the subject was correct and trials in which an incorrect response was given. The reasoning of this approach was that if the found effects are behaviorally relevant, larger effects should be expected upon correct performance in the delayed match-to-category task. Similar to the statistical analyses of the training-interactions, we focused on clusters that exhibited significant category effects and training interactions, estimated the effect size and bootstrapped the respective upper- and lower bounds of the 95% confidence intervals (with replacement) while applying a Bonferroni correction for multiple comparisons at the cluster level. Matching the behavioral accuracy in the DMC task, on average 136 incorrect trials were compared to 277 correct trials in session five, whereas 94 incorrect trials were compared to 277 correct ones in session 22.

2.4.11 Source Analysis

To compare source activity on the cortical surface, we used the sLORETA algorithm (Pascual-Marqui, 2002), as implemented in the Brainstorm software (Tadel et al., 2011), on the adaptor stimulus data, which was used as projection target, to localize the adaptation-specific category effects. For every subject, we first segmented the individual MRI into white and gray matter using Freesurfer (Dale, BR Fischl, et al., 1999; B Fischl et al.,

1999). We then performed the source reconstruction based on each individual anatomy and aligned the results to MNI space (Colin27) using spherical averaging of the cortical surfaces. For statistical analyses, we contrasted the average source activity (L2-Norm) during the earliest time-window of category selectivity in session five (275-293ms) and session 22 (113-140ms) at every surface vertex and applied a clusterwise correction for multiple comparisons based on a nonparametric permutation test (Maris and Oostenveld, 2007). Only vertices showing $p < 0.05$ were included in the cluster estimates.

2.5 Results: Category selectivity at 116ms after extensive category learning

2.5.1 Behavioral Data on Category Training

Subjects were trained to categorize two artificial categories of faces defined in a parametric feature space (Figure 2.1). Training lasted for a total of 22 sessions consisting of 756 training trials each. In each trial, participants were required to make a category judgment for a given stimulus and received auditory feedback as training signal. Classification accuracy reached 89.4% after five training sessions, and 95.3% after training was completed in session 22 (Figure 2.5). At the same time, reaction times continued to decrease with training (from 679ms in session five to 538ms in session 22, $p < 0.01$ paired t-test). Thus, although high classification performance was reached already after five training sessions, the behavioral data indicate continued improvements over the whole training period.

2.5.2 Behavior during the Delayed Match-To-Category Task

The electrophysiological correlates of category effects were estimated using an adaptation paradigm in which subjects performed a delayed match-to-category task (Figure 2.2a). During the baseline session, and therefore prior to category training, the DMC performance of our subjects did not differ significantly from chance (49% accuracy, $p = 0.128$, t-test against a chance-level of 50%). This demonstrates that our artificial category structure is not an inherent property of the stimulus space. With training, DMC per-

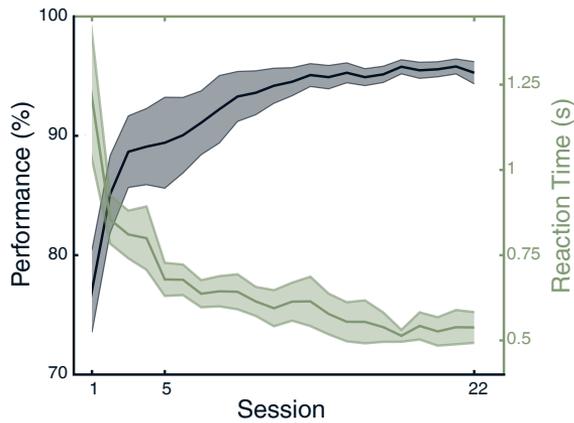


Fig. 2.5 Effects of Training on Performance and Reaction Times. Subjects received auditory feedback as a training signal, but no explicit information about the underlying feature space or category structure. Recognition performance and reaction times improved until session 22, illustrating continued training effects. Error bars depict SEM.

formance increased to 66.2% in session five and 76.0% in session 22 (see supplementary material in Kietzmann and König, 2015). A repeated measures ANOVA (with session (baseline, five, 22) and category membership (internal, external) as factors) revealed a significant main effect of session ($p < 0.01$, all pairwise comparisons are significant at $p < 0.01$, t-test, Bonferroni corrected), but no main effect of category membership ($p > 0.05$) and no significant interaction ($p > 0.05$). Thus, although there was an overall increase in task performance with training, there was no significant difference in the performance of the category-internal and category-external trials indicating that the task was equally demanding in trials of both conditions. Effects of condition difficulty can therefore not explain the categorical effects observed. The accuracy in the DMC task was lower than expected from the high training performance (95% training accuracy predicts around 90% accuracy for two consecutive decisions). This is in line with observations by Helie and Ashby (2012), who observed sub-optimal DMC performance even for comparably simple one-dimensional category boundaries.

Multiple reasons can account for this difference. First, a successful DMC trial requires, in addition to the correct classification of both stimuli, successful working memory encoding and retrieval, a successful category comparison, and a successful match to the correct motor response. Moreover, the electromagnetic shielding required for the MEG measurements required the use of a back-projected display with decreased contrast compared to the training monitor. Most importantly, our participants had considerably less experience with the structure of the DMC task, compared to the excessive amount of trials in the training paradigm.

2.5.3 Training-Induced Category Effects in Visual Responses (MEG data)

To test for category effects in the visually evoked responses, we compared the magnetic fields evoked by the second stimulus in the category-internal and category-external adaptation trials in the MEG adaptation paradigm, while controlling the low-level stimulus properties of the two conditions (Figure 2.2b). This indirectly tests for category selectivity, as differences between these two conditions will only be detectable if category-information is encoded in the underlying cortical activity. Importantly, adaptation paradigms were previously shown to result in both, effects of repetition suppression as well as repetition enhancement (Krekelberg et al., 2006; Segaert et al., 2013), depending on stimulus timing (James and Gauthier, 2006), effect latency (Marinkovic et al., 2003; Petit et al., 2006), and region of interest (Zago et al., 2005). For analyses of visually evoked responses, we employed a spatiotemporal projection approach that allowed us to focus our analyses on adaptation-based effects, and an unconstrained cluster-based analysis for effects that are not due to adaptation.

Focusing on adaptation-based effects first, we projected the evoked fields in response to the second stimulus onto the adaptor response, and thereby created activity traces for each session, subject and condition (category-external and category-internal). We then performed a paired t-test at every time-point to test for differences between category-external and category-internal conditions (positive t-values indicate a larger response for category-external trials). This provided us with temporal candidate clusters that exhibit significant category effects for every session. To ensure that the observed category effects were indeed the result of category training, it had to be shown that category effects were significantly larger post-training as compared to the baseline session. As a final step, we therefore estimated the effect sizes and confidence intervals of the training interaction for each temporal candidate cluster (Bonferroni corrected at the cluster level, thereby controlling for multiple comparisons). Only temporal clusters surviving this rigorous control will be reported in the following. The clusters reported will not only exhibit significant category effects, but also show significantly stronger category effects compared to baseline, indicating that the seen category effects are the result of category training.

We first analyzed the data from the baseline session. Here, we found no significant category effects (Figure 2.6a), confirming that the category structures used for training were not an inherent property of the used stimulus space. We then analyzed the data of the two post-training sessions five and 22, testing for category effects and training interactions. After five training sessions, the earliest significant, training-induced category effects were evident between 275-293ms (Figure 2.6b). With developing category expertise, however, a temporal shift in category effects was observed. After 22 training sessions, the earliest cluster exhibiting significant training effects occurred already after 113ms (lasting from 113 to 140ms). Additional time-windows of significant category

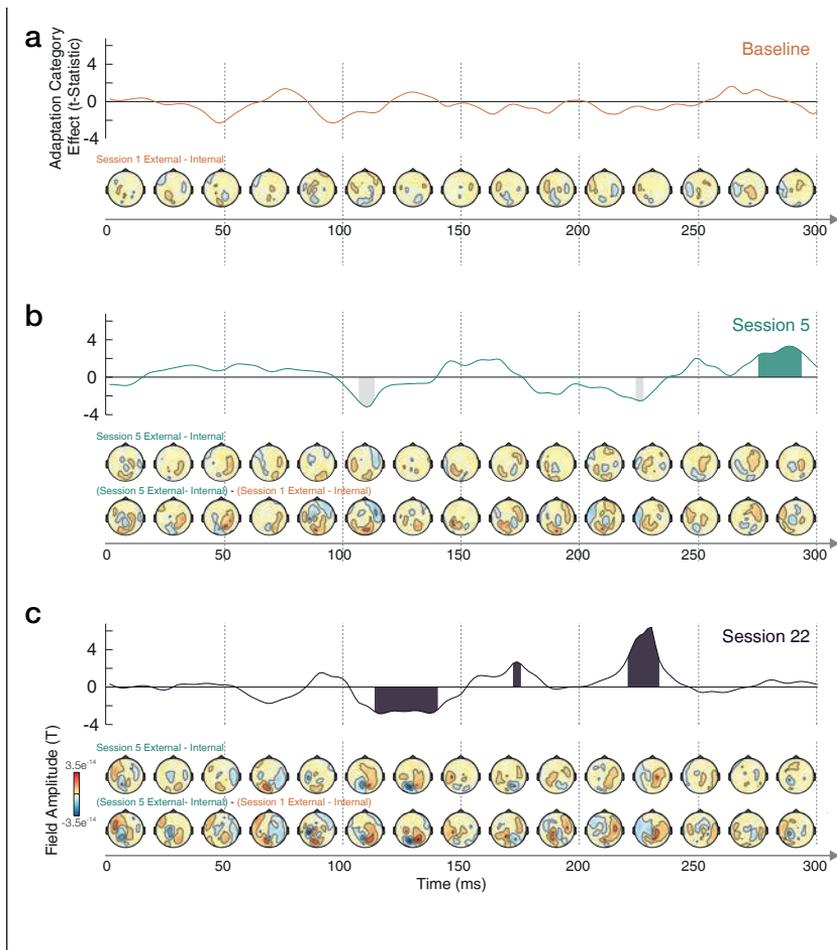


Fig. 2.6 Training-Induced MEG Category Effects across Time. The traces in each panel represent the adaptation-based t-statistics across time, as obtained from the spatiotemporal projection method, comparing category-internal and category-external trials. Candidate temporal windows during which visually evoked responses showed significant category effects and a significant training interaction are shaded in dark colors. Candidate windows exhibiting no significant training effects are marked in light grey. Together with traces of t-statistics, each panel shows the underlying effect topographies and training interactions (i.e. effect difference between baseline and post-training) where applicable. (a) During the baseline session, no significant category effects could be found. (b) After five training sessions, the first significant training-induced window of category selectivity is present from 275-293ms. (c) After extended category training in 22 sessions, the earliest training-induced category effects are present from 113-140ms. Additional clusters of significant training-induced category effects were found between 171-175ms and 220-233ms.

effects were found between 171-175ms and between 220-233ms (Figure 2.6c). The corresponding template and effect topographies are shown in Figure 2.4 and Figure 2.6, respectively. Importantly, the earliest cluster found in session 22 (113 to 140ms) is not only significantly different to the baseline session, indicating an overall training effect, but also significantly larger compared to session five (CI95 = [-8.47*10⁻¹⁵, -1.55*10⁻¹³]). The latter indicates that the extensive training between sessions five and 22 lead to the temporal shift in category effects. The observed speed-up of more than 160ms from session five to 22 is remarkable, as our subjects already categorized the stimuli at about 90% accuracy during training session five. Moreover, it is comparable to the observed decrease in reaction times of around 140ms from training session five to 22. As a necessary result of the close control of low-level feature differences (same stimulus-space distance and direction for category-external and category-internal trials), stimuli close to the category boundary were shown more frequently in the DMC task. Based on this, it could be argued that effects of long-term adaptation might specifically affect category external trials, thereby contributing to the early category-effects observed in session 22. Speaking against this possibility we found no significant differences during the baseline session. Furthermore, we report only clusters exhibiting both, significant category effects and training interactions to ensure that the effects reported are indeed the result of category training. Following adaptation-based effects, we tested whether category-specific effects exist that are not due to adaptation (Figure 2.6). Based on the residual data, i.e. the parts of the evoked fields that cannot be explained based on the adaptor response, we compared category-internal and category-external conditions, while correcting for multiple comparisons using a non-parametric cluster-based permutation test. This analysis revealed no significant effects of category selectivity (all cluster $p > 0.3$). This highlights the successful balance of motor-response mapping

across conditions and indicates that, in the current setup, category-selectivity was only observed adaptation-based.

2.5.4 Relation of Physiological Category Effects to Behavior

To test whether the observed adaptation-based category effects were behaviorally relevant, we compared the category effect sizes for successful and erroneous trials during the delayed match-to-category task. Again, we estimated the effect sizes and confidence intervals, while Bonferroni correcting for multiple comparisons at the cluster level. This analysis revealed significant differences for the earliest cluster in session 22, indicating the behavioral relevance of the effect. No other cluster in session 22 and five exhibited significant behavioral effects. Considering the absence of significant differences for session five, it should be noted that behavioral errors in the delayed match-to-category task can have various origins. Apart from the variability in the category signal, which is of interest here, they include subjects inattentiveness, errors in working memory and an incorrect mapping of the perceptual decision to the appropriate behavioral response. These additional sources of error significantly complicate the search for behavioral relevance, as they all do not predict differences in category-selectivity. Moreover, it is possible that effects of behavioral relevance occurred at an even later point, extending beyond the 300ms analyzed here.

2.5.5 Source Analyses

Following the analyses in sensor space, we tested whether the temporal shift in category selectivity observed between session five (275-293ms) and session 22 (113-140ms) is due to altered neuronal processing in the same cortical areas, or whether different

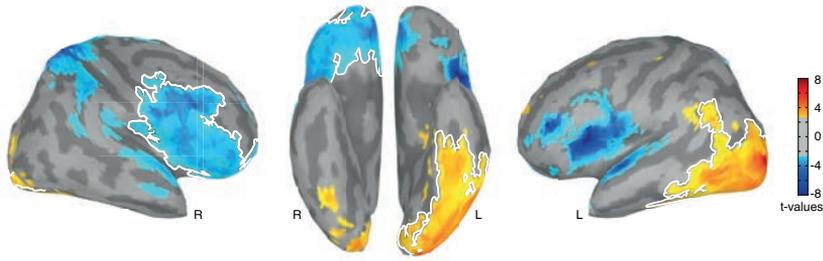


Fig. 2.7 Source Localization Results. Source activations of the earliest clusters of category selectivity in sessions five and 22 were contrasted. Shown are uncorrected t-values with a cutoff at $p < 0.05$. Blue regions show larger activity during the category-selective time-window in session five, red regions show larger activity in the early category-cluster effects in session 22. A white border highlights clusters after controlling for multiple comparisons (cluster-based permutation test).

sets cortical areas are activated during these two time windows of interest. To this end, we computed a standardized low resolution brain electromagnetic tomography (sLORETA, Pascual-Marqui, 2002) on the data of the adaptor stimulus, which underlies the observed category effects (see Materials and Methods for details). We estimated the average source activations during the two time-windows of interest and tested for significant differences based on a t-statistic, while controlling for multiple comparisons using a nonparametric cluster-based permutation test (Maris and Oostenveld, 2007) on the cortical surface. This analysis revealed that the previously shown temporal shift in category selectivity was accompanied by an anterior-to-posterior shift of cortical activation (Figure 2.7, positive t-values indicate a stronger activation in the early time-window observed in session 22). Although the source distributions exhibited considerable overlap, among others in parietal regions, the time-window of training-induced category effects in session five showed a significantly stronger activation in the ventrolateral and ventromedial parts of the PFC. In contrast, the cortical activation during the earlier time-window of category selectivity in session 22 exhibited significantly stronger activity in more posterior regions, including the occipitotemporal cortex. All results were reproduced in a separate source localization anal-

ysis based on dynamical statistical parametric mapping, dSPM (Dale, AK Liu, et al., 2000) instead of sLORETA. As stated above, the source analysis approach taken here specifically focuses on time-points at which adaptation-effects were observed. Using the adaptor response as basis, it highlights sources that are differentially activated at the two earliest significant timepoints in sessions 5 and 22. This assumes that the respective effects observed are the result of altered activity in the strongest sources. This is not necessarily true, as the projection approach can yield significant results that are driven by weaker sources and sensor-patterns. In the current case, however, this concern is not warranted, as the effect topographies nicely match the topographies of the adaptor stimulus. Moreover, the fact that no category-effects were observed in the residual activity indicates that the localized sources are the sole contributor to category selectivity in the current data.

2.6 Discussion

Previous work on naturally occurring categories has demonstrated that category information can be rapidly extracted from visually presented objects. It remained unclear, however, how the visual system copes with the challenge to reach such rapid recognition speeds while at the same time allowing for sufficient plasticity to encompass the fast learning of entirely new categories. Are the same neuronal mechanisms and structures involved in recognizing re-occurring and newly learned categories, or are they different? And, if they are different, are novel categories implemented differently with prolonged experience? Here we investigated these issues by extensively training nine subjects to categorize two artificial visual categories. During training, we recorded MEG data in a rapid event-related adaptation paradigm to investigate the emergence of category selectivity in visually evoked responses. Additionally, MEG data were recorded prior to category training to serve as a baseline. Using a novel data projection approach, which allowed us to separate adaptation-based and non-adaptation effects, we demonstrate the emergence and, following this, a temporal shift in category selectivity. The data recorded in the baseline session did not exhibit any category effects, indicating successful control for low-level stimulus properties. After five training sessions, the earliest training-induced category effects were found around 280ms of processing. With extensive training in 22 sessions, we observed a temporal shift in category selectivity. The first significant differences were now found about 160ms earlier, between 113 and 140ms. We then investigated whether the temporal shift in category selectivity was accompanied by changes in the spatial pattern of the underlying cortical activity. We compared the source activations during the two earliest temporal clusters of sessions five and 22 and found a significant anterior-to-posterior shift. While the cortical activity during the late category effects in session five

showed stronger signals in PFC, the early time-window of category selectivity in session 22 exhibited an increased activation in occipitotemporal regions. An interesting aspect of the results is that temporally late effects in session five and 22 exhibit positive t-values, suggesting a decreased response for category-internal compared to category-external trials and therefore repetition suppression. The earliest effect observed, cluster one in session 22, however, exhibits a reverse effect, indicating effects of repetition enhancement. This finding is in line with previous EEG adaptation experiments that demonstrated early enhancement, but late suppression effects (Marinkovic et al., 2003; Petit et al., 2006), and contributes to an ongoing debate about the mechanisms underlying differential repetition effects in electrophysiology and neuroimaging (James and Gauthier, 2006; Krekelberg et al., 2006; Segaert et al., 2013). Our finding of an early cluster of category selectivity, starting at 113ms and lasting until 140ms, is fully compatible with previous studies of natural categories in macaque and human. In the macaque, Sugase et al. (Sugase et al., 1999) recorded from inferotemporal cortex (IT) and observed a peak in category information after only 117ms of processing. In line with this, Hung et al. (Hung et al., 2005) demonstrated that relatively small numbers of randomly selected neurons in IT allow for reliable category decoding, peaking 125ms after stimulus onset. Interestingly, the authors also show decoding of low-level properties such as size and position of an object, arguing for residual retinotopic information in the neuronal response. This emphasizes the necessity to control for low-level stimulus properties and underlines the benefits of baseline measurements in category training. Finally, Freedman et al. (2003) applied a receiver operator characteristic approach to recordings from macaque IT and PFC. They showed that IT cells exhibited category selectivity after 127ms. In humans, electrocorticographic recordings provided direct evidence that natural categories can successfully be decoded at a mean latency of 115ms (H Liu et al., 2009). Remarkably, decoding was possible based on single trials, allowing for gener-

alization across rotation and changes in scale. In line with this, MEG recordings of human subjects provided evidence that visually evoked responses of houses and faces can be separated already at the time of the M100 component (J Liu et al., 2002). In the same study, a positive correlation of response amplitude and categorization performance was shown, indicating the behavioral relevance of the early category signals. Using a multivariate decoding approach, Carlson et al. (Carlson, Hogendoorn, et al., 2011) showed that it is possible to differentiate two visual categories (faces and cars) after 135ms of processing, even if the retinal locations of trained and tested stimuli were different. Similar results were later obtained using a wider range of categories (Carlson, Tovar, et al., 2013). Extending this approach, Cichy et al. (Cichy et al., 2014) performed a temporally fine-grained representational similarity analysis based on 92 object images and demonstrated successful decoding of different types of category selectivity at approximately the same latency. The authors furthermore showed a correlation between the brain responses in macaque and human, providing further evidence for a common representational space (Kriegeskorte, Mur, et al., 2008). These results of early category selectivity were extended to a more natural, cluttered stimulus set contrasting faces to other stimulus categories (Cauchoix et al., 2014). Finally, electrooculography (EOG) data provided by Kirchner and Thorpe (Kirchner and Thorpe, 2006) suggest that category information is present and behaviorally relevant after only 120ms of processing. However, it should be noted that all of the studies mentioned above either investigated neuronal responses to naturally occurring categories or did not include a pre-training baseline. Apart from the inherent challenges to differentiate category selectivity from systematic differences in the low-level statistics (Thierry et al., 2007; Wichmann et al., 2010; Crouzet and Thorpe, 2011; Rossion and Caharel, 2011; VanRullen, 2011), these setups do not allow for an investigation of emerging category selectivity with increasing category experience, which is the focus of the current study. Overall,

the neuronal mechanisms underlying the categorization of visual input have been in the focus of a lively debate over the recent years. A prominent view centers around the idea that category information is extracted by PFC (Serre et al., 2007; Cromer et al., 2010; Roy et al., 2010; Antzoulatos and Miller, 2011). Accordingly, neuronal selectivity in temporal regions is seen as merely providing a sufficiently complex vocabulary from which the category information can be flexibly read out. This view is consistent with the predictions of the two-stage model of perceptual category learning (Riesenhuber and Poggio, 2002), which hypothesized that neurons in IT obtain sharper tuning to re-occurring stimulus features, while regions in frontal cortex learn to associate these features with the corresponding category membership. In humans, experimental evidence supporting such division of labor was provided by Jiang et al. (2007). They showed that category training can lead to an increased shape selectivity in ventral areas whereas category selectivity was found only in the lateral PFC (but see Minamimoto et al., 2010). Moreover, there is evidence for enhanced shape selectivity in ventral areas in human and macaque (Sigala and Logothetis, 2002b; Freedman et al., 2006; Linden, Turennout, et al., 2010; Linden, Wegman, et al., 2013). Nevertheless, the large body of evidence for rapid category selectivity in IT, as reviewed above, supports a contrasting view according to which category information might already be extracted at the level of the temporal lobe (DiCarlo et al., 2012; Mur et al., 2012; N Liu et al., 2013). Closely mirroring this controversy, different labs have studied the cortical representations of spatial and motion-related categories in the parietal and prefrontal cortex, arriving at opposite conclusions. Whereas some observed stronger and earlier category signals in prefrontal compared to parietal cortex (Goodwin and Blackman, 2012; Crowe et al., 2013), others reported the reverse: earlier category-selective signals in parietal cortex preceding prefrontal category selectivity (Fitzgerald et al., 2012; Swaminathan and Freedman, 2012). Providing a potentially unifying solution to

these controversies, we have demonstrated here that prolonged category training can lead to a temporal shift in category selectivity, which is accompanied by an anterior-to-posterior shift in cortical activity. These data provide a cue as to how the brain could balance the need for robust and fast recognition of re-occurring categories while still allowing for considerable flexibility and rapid plasticity. Selectivity for novel categories relies more heavily on PFC and, as indicated by the long latency of the observed effect, potentially recurrent processing. Sufficient expertise with the categories, however, leads to changes in the cortical implementation of the trained categories, thereby allowing for a substantial speedup in processing times and emphasizing cortical processes in occipitotemporal regions. A comparable view was recently described by Seger and Miller (2010) who proposed that the brain might simultaneously implement fast and slow learning processes. Fast learning provides multiple advantages, such as increased flexibility and rapid adjustments, but at the cost of an increased risk of erroneous classification. Slow learning, on the other hand, is less error-prone but at the cost of extended training requirements. In line with this suggestion, Helie, Roeder, et al. (2010) trained participants in a rule-based categorization task and demonstrated an initial transition from subcortical to cortical areas, including PFC, and a second transition towards the premotor cortex with emerging automaticity. The current setup, using MEG, is not particularly suited to resolve subcortical activity. However, it is possible that a similar transition from subcortical to cortical areas also occurred in our participants during initial category training, potentially even earlier than our first post-training MEG recording. The question of the respective contribution of subcortical and cortical regions in category learning was recently addressed in the macaque (Muhammad et al., 2006; Antzoulatos and Miller, 2011; Antzoulatos and Miller, 2014), suggesting that the striatum is indeed involved during initial category learning, potentially entraining prefrontal circuitry. In line with this suggested learning transition, our results provide a potential

explanation as to why some previous studies did not see (early) category selectivity in temporal areas after category training (Jiang et al., 2007; S Li et al., 2007; Gillebert et al., 2009; Scholl et al., 2014). Apart from many differences between these experiments and our study, our data suggest that the extent of training is a decisive factor. Comparably short training times might only reveal rather late category selectivity in frontal regions, as observed in session five here, whereas prolonged training is required for early occipitotemporal effects. Another important difference is given by the type of category space used during training. Using psychophysical measurements Folstein et al. (JR Folstein, Palmeri, and Gauthier, 2012) demonstrated that factorial, but not blended morphspaces, lead to an enhanced discriminability of category-relevant feature dimensions, implying that studies using the latter (Jiang et al., 2007; Gillebert et al., 2009; Linden, Turennout, et al., 2010; Scholl et al., 2014), were less likely to observe category-selectivity in visual areas. This suggestion was corroborated by a follow-up fMRI study in which the authors demonstrated robust category signals in visual areas (J Folstein et al., 2012; JR Folstein, Palmeri, Van Gulick, et al., 2015). In line with such evidence for category-selectivity in visual areas, effects of expertise have been demonstrated in the FFA (Gauthier, Skudlarski, Gore, and aW Anderson, 2000), and the N170 ERP component (Tanaka and Curran, 2001). By contrasting correct and incorrect responses, we demonstrated significant behavioral relevance of the early category effects starting at 113ms in session 22. It has to be noted, however, that the time-points of category selectivity observed in sessions five and 22 do not necessarily mark the end point of the perceptual decision process. Successful performance in the DMC task requires the successful completion of additional processing steps, such as the successful comparison of the two shown categories and the mapping of the perceptual decision to the appropriate motor response. Moreover, effects of perceptual certainty (Philiastides and Sajda, 2006) and ongoing evidence accumulation (Donner

et al., 2009) can be expected to play a vital role in the perceptual decision process. While further experiments are required to fully disentangle the contribution of these different factors, we have shown here that the brain is capable of extracting visual categories based on two different modes. Novel categories are recognized late, involving recurrent processing and increased activity in PFC. This pattern of results is consistent with a re-labeling of existing visual features, which would allow the system to flexibly learn new categories and to quickly adjust to changing task-demands (Mc-kee et al., 2014). Extended category experience, however, leads to a significant speed-up in category selectivity, accompanied by increased activity in occipitotemporal cortex. This suggests that re-occurring categories are processed differently to allow for quick and reliable recognition. Taken together, our results suggest that the brain balances plasticity for acquisition of new and efficiency in processing of known categories by relying on different networks.

Predictions and Decision making: Expectations trump reality

” *It might be better to publish these results on face value, as they are, rather than trying to squeeze them into the procrustean framework of an inappropriate theory*

— **Three anonymous reviewers**

Contributions

eLife: Ehinger, Häusser, Ossandón, and König 2017
BVE JPO and PK conceived the study. **BVE** and KH recorded the data and performed the analyses. **BVE** JPO and PK wrote the manuscript.

3.1 Layman summary

We have seen in Chapter 2 how we can learn new categories. But in order to make sense of the world around us, we do not only need to categorize our sensory input of a single sense, but combine information from multiple sources while taking into account how reliable they are. For example, when crossing the street, category learning can tell us what a car and what a bike is. Usually we would rely more on the input of our eyes than on our ears. However, we can reassess the reliability of the information: on a foggy day with poor visibility, we might prioritize listening for the difference between the sounds a car or a bike make. How do we assess the reliability of information generated within the brain itself?

We are able to see because the brain constructs an image based on the patterns of activity of light-sensitive proteins in a part of the eye called the retina. However, there is an area on the retina where the presence of the optic nerve leaves no space for light-sensitive receptors. This means that there is a corresponding point in our visual field where the the eye and, therefore, the brain receives no visual input from the outside world. To prevent us from perceiving this gap, known as the visual blind spot, the brain fills in the blank space based on the contents of its surrounding areas. While this is usually accurate enough, it means that our perception in the blind spot is objectively unreliable.

To find out whether we are aware of the unreliable nature of visual perception in the blind spot, we presented volunteers with two striped stimuli, one on each side of the screen. The center of some of the stimuli were covered by a patch that broke up the stripes. The volunteers' task was to select the stimulus with uninterrupted stripes. The key to the experiment is that if the central patch appears in the blind spot, the brain will fill in the

middle so that the stripes appear to be continuous. This means that the volunteers will have to choose between two stimuli that both appear to have continuous stripes. If they have no awareness of their blind spot, we might expect them to randomly choose the right or the left stimulus. Alternatively, if they are subconsciously aware that the stimulus in the blind spot is unreliable, they should choose the other one.

In reality, exactly the opposite happened: the volunteers chose the blind spot stimulus more often than not. This surprising result suggests that information generated by the brain itself is sometimes treated as more reliable than sensory information from the outside world. Future experiments should examine whether the tendency to favor information generated within the brain over external sensory inputs is unique to the visual blind spot, or whether it also occurs elsewhere.

3.2 Humans treat unreliable filled-in percepts as more real than veridical ones

Humans often evaluate sensory signals according to their reliability for optimal decision-making. However, how do we evaluate percepts generated in the absence of direct input that are, therefore, completely unreliable? Here, we utilize the phenomenon of filling-in occurring at the physiological blind-spots to compare partially inferred and veridical percepts. Subjects chose between stimuli that elicit filling-in, and perceptually equivalent ones presented outside the blind-spots, looking for a Gabor stimulus without a small orthogonal inset. In ambiguous conditions, when the stimuli were physically identical and the inset was absent in both, subjects behaved opposite to optimal, preferring the blind-spot stimulus as the better example of a collinear stimulus, even though no relevant veridical information was available. Thus, a percept that is partially inferred is paradoxically considered more reliable than a percept based on external input. In other words: Humans treat filled-in inferred percepts as more real than veridical ones.

3.3 Introduction: Decision making and reliabilities of inferred stimuli

In order to make optimal and adaptive decisions, animals integrate multiple sources of sensory information across time and space. One of the prime examples of this is observed when animals are confronted with coherently-moving stimuli during random-dot motion experiments. In such experiments, performance and the corresponding neural activity vary proportionally to signal strength in a way that is consistent with the progressive integration of evidence over time (Shadlen, Britten, et al., 1996; Shadlen and Newsome, 2001). Besides temporal accumulation, sensory integration is also possible by combining the information from multiple sensory sources (Quigley et al., 2008; S Schall et al., 2009; Hollensteiner et al., 2015; Wahn and König, 2015b; Wahn and König, 2015a; Wahn and König, 2016). In the case of multisensory perception, several experiments have shown that integration often occurs in a statistically optimal way. This has been best demonstrated in cue-integration tasks in which humans perform as if they were weighting the different sources of information according to their respective reliabilities (Ernst and Banks, 2002; Alais and Burr, 2004; Körding and DM Wolpert, 2004; Tickle et al., 2016). This form of statistical inference has also been demonstrated for cortical neurons of the monkey brain, with patterns of activity at the population level that are consistent with the implementation of a probabilistic population code (Gu et al., 2008; Fetsch et al., 2012). In most of these sensory integration experiments, the perceptual reliability of different inputs is probed through quantitative manipulations of the inputs' signal-to-noise ratios (Heekeren, Marrett, Bandettini, et al., 2004; Tassinari et al., 2006; Bankó et al., 2011). However, some percepts are unreliable not because they are corrupted by noise but because they are inferred only from the context and thus intrinsically uncertain. This occurs naturally in monocular vision

at the physiological blind spot, where content is "filled-in" based on information from the surroundings. In this case, no veridical percept is possible at the blind spot location. Though changes in reliability due to noise directly result in behavioral consequences, the effects of the qualitative difference between veridical and inferred percepts that are otherwise apparently identical are unknown. We recently reported differences in the processing of veridical and inferred information at the level of EEG responses (Ehinger, König, et al., 2015). We demonstrated that a qualitative assessment of differences in reliability exists at the neural level in the form of low- and high-level trans-saccadic predictions of visual content. Notably, active predictions of visual content differed between inferred and veridical visual information presented inside or outside the blind spot. Although no difference was found between low-level error signals, high-level error signals differed markedly between predictions based on inferred or veridical information. We concluded that the inferred content is processed as if it were veridical for the visual system, but knowledge of its reduced precision is nevertheless preserved for later processing stages. In the present experiment, we address whether such an assessment of a dichotomous, qualitative difference in reliability is available for perceptual decision-making. Using 3D shutter glasses, we presented one stimulus partially in the participant's blind spot to elicit filling-in and a second stimulus at the same eccentricity in the nasal field of view outside of the blind spot. The subject's task was to indicate which of the two stimuli was continuously striped and did not present a small orthogonal inset (see Fig. 1A). Crucially, stimuli within the blind spot are filled-in and thus perceived as continuous, even when they present an inset. In the diagnostic trials, both stimuli were physically identical and continuous, and subjects were confronted with an ambiguous decision between veridical and partially inferred stimuli.

We evaluated two mutually exclusive hypotheses on how perceptual decision-making could proceed when confronted with an

ambiguous decision between veridical and inferred percepts. In the first case, we hypothesized that subjects are unable to make perceptual decisions based on an assessment of differences in reliability between veridical and inferred stimuli. Therefore, subjects would have an equal chance of selecting stimuli presented inside or outside the blind spot. Alternatively, it might be possible to use the information about the reduced reliability of filled-in information. In this case, we expect subjects to follow an optimal strategy and trust a stimulus presented outside the blind spot, where the complete stimulus is seen, more often than when the stimulus is presented inside the blind spot, where it is impossible to know the actual content within the filled-in part.

3.4 Results: A bias towards filled-in stimuli

We conducted five experiments (see Fig. 1 and the methods for a detailed description of the tasks). The first one tested the presence of a bias against the blind spot location; the other four experiments were replications of the first experiment with additional control conditions. In the first two controls, we tested the existence of biases between the nasal and temporal visual fields at locations other than the blind spot. In the third control, we tested whether an opposite bias existed when the task was reversed. The last experiment controls whether the observed bias could be explained by probability matching.

3.4.1 Experiment 1

In the first experiment, 24 subjects performed a 2-AFC task in which they had to indicate which of two stimuli was continuously striped instead of presenting a small orthogonal central inset (Fig. 1A, B). The stimuli were presented simultaneously in the periphery at the locations of the blind spots or at equivalent eccentricity on the opposite side (Fig. 1C, D). We used a 3D monitor and shutter glasses that allowed for the controlled monocular display of the stimuli. That means each stimulus was visible to a single eye only. There were always two stimuli, therefore, in a given trial either one or both eyes were stimulated (Fig. 1B). Importantly, subjects always perceived the two stimuli at the same locations, to the left and the right of the fixation cross. In this experiment there were perceptually ambiguous trials, where two continuous stimuli were perceived, and unambiguous trials where one stimulus contained a visible inset.

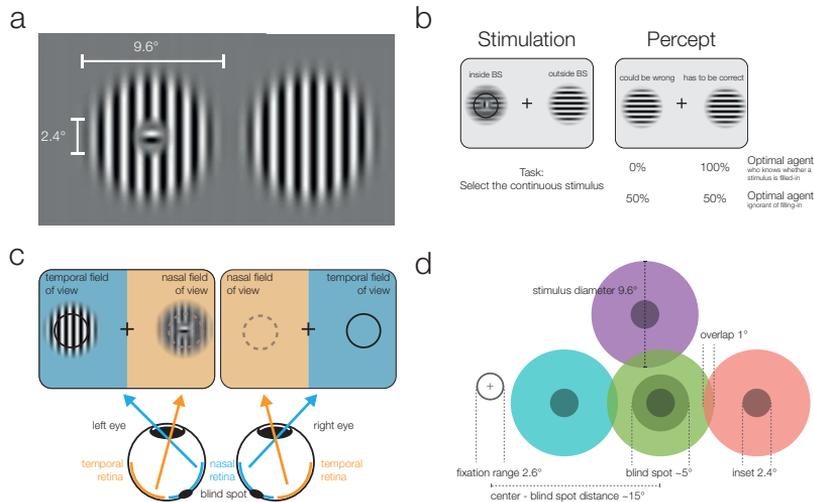


Fig. 3.1 a) Striped stimuli used in the study. The inset was set to 50% of the average blind spot size. The global orientation of both stimuli was the same, but in different trials, it could be either vertical (as shown here) or horizontal (not shown). b) Each stimulus was displayed individually either (partially) inside or (completely) outside the blind spot. This example presents an inset stimulus inside the subject's left blind spot. However, due to filling-in, it is perceived as continuous (right column). The task required subjects to select the continuous stimulus, and it was designed to differentiate between two mutually exclusive predictions: First, subjects cannot differentiate between the two different types of stimuli and thus answer randomly. Alternatively, subjects have implicit or explicit knowledge about the difference between inferred (filled-in) and veridical contents and consequently select the stimulus outside the blind spot in ambiguous trials. c) Two stimuli were displayed using shutter glasses. Each stimulus was presented to one eye only, and it is possible that both are presented to the same eye (as in the example depicted here). That is, the left stimulus could be shown either in the temporal field of view (nasal retina) of the left eye (as in the plot) or in the nasal field of view (temporal retina) of the right eye (not shown). In this case, the trial was unambiguous: The stimulus with an inset was presented outside the blind spot and could be veridically observed, therefore, the correct answer was to select the left stimulus. d) The locations of stimulus presentation in the five experiments. All stimuli were presented relative to the blind spot location of each subject. All five experiments included the blind spot location (green). In the second and fifth experiment, effects at the blind spot were contrasted with a location above it (purple). In the third experiment, the contrasts were in positions located to the left or the right of the blind spot. Note that both stimuli were always presented at symmetrical positions in a given trial, the position of the stimuli differed only across trials.

In the unambiguous trials, an orthogonal inset was present in one of the stimuli. Importantly, in these trials, the stimulus with the inset was outside the blind spot and therefore clearly visible. As expected, subjects performed with near-perfect accuracy (Fig. 2, unambiguous trials, blue data), choosing the continuous stimulus in an average of 98.8% of trials (95%-quantile over subjects [96.4%-100%]). There were two types of ambiguous trials. In the first type (Fig. 2, ambiguous control, red data), one of the following applied: both stimuli were continuous and appeared outside the blind spots in the nasal visual fields (Fig. 2, row 3); both were continuous and appeared inside the blind spots (Fig. 2, row 4); or one was continuous, the other had an inset, and both appeared inside the blind spots with the inset either in the left or the right blind spot (Fig. 2, rows 5 and 6). In the case when a stimulus with an inset was present, this central part was perfectly centered inside the blind spot (Fig. 1A), and in consequence was perceived as continuous due to filling-in. Thus, in all four versions, subjects perceived two identical stimuli, and there was no single correct answer. In this type of ambiguous trial, subjects showed a small global leftward bias and chose the left stimulus in 53.6% of trials (Fig. 2, continuous vertical line). In addition, no difference can be seen between the perception of pairs of filled-in stimuli and pairs of veridical continuous stimuli (Fig. 2, rows 3 vs. 4-6). This type of ambiguous control trial confirms that filling-in was accurately controlled in our experiment.

In the second type of ambiguous trials one stimulus was presented inside and the other outside the blind spot (Fig. 2, ambiguous diagnostic, data in green). This allowed us to test directly between two rival predictions: whether subjects will show a bias against the stimulus that is partially inferred (inset area inside the blind spot) and in favor of the veridical stimulus (in the opposite visual field), or no bias. Selecting the filled-in stimulus is a suboptimal decision because the stimulus presented partially in the blind

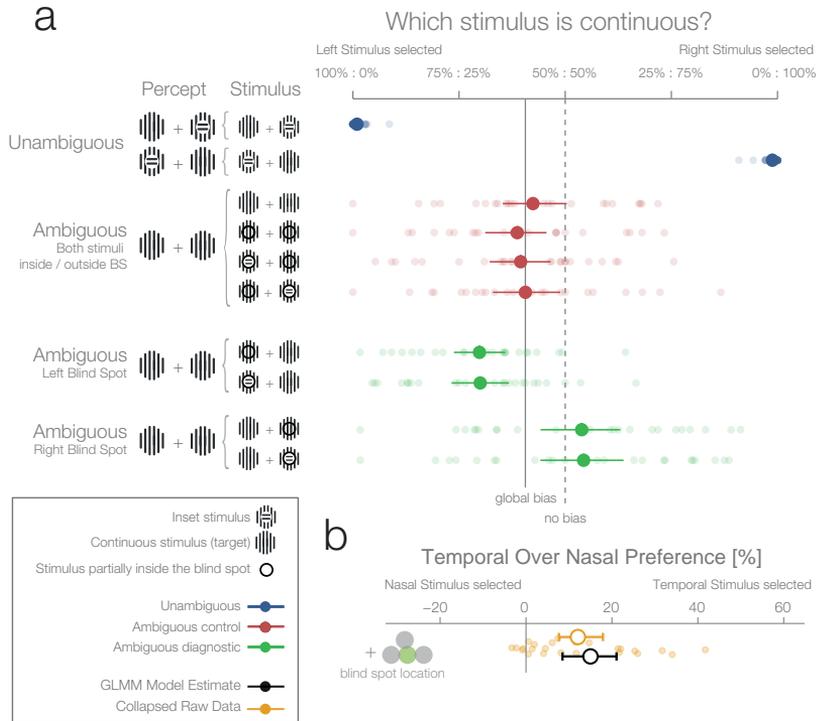


Fig. 3.2 a) The left column shows schematics of the actual stimulation and the associated percepts for the corresponding data presented in the right panel. A dark-lined circle, where present, indicates that the stimulus was presented in the blind spot and, consequently, an inset stimulus within was perceived as a continuous stimulus due to filling-in. The plot to the right shows each subject's ($n=24$) average response and the group average (95% bootstrapped confidence intervals, used only for visualization). The results from unambiguous trials (blue) show that subjects were almost perfect in their selection of the continuous stimulus when an inset was visible. For the first type of ambiguous control trials (red), both stimuli were presented either outside or inside the blind spot. Here, only a global bias toward the left stimulus can be observed (solid line, the mean across all observed conditions in red). Note that the performance when presenting an inset in the blind spot was identical to the one of presenting a continuous stimulus in the blind spot. The ambiguous diagnostic conditions (green) show the, unexpected, bias toward the blind spot (for either side). b) Statistical differences were evaluated by fitting a Bayesian generalized mixed linear model. In the model, the left and right ambiguous diagnostic conditions were combined in a single estimate of the bias for nasal or temporal stimuli (outside or inside the blind spot respectively). The plot shows the average effect of each subject (small yellow dots), the bootstrapped summary statistics of the data (yellow errorbar), and the posterior 95% credibility interval model estimate (black errorbar).

spot is the only one which could possibly contain the inset. This is explicit in the cases where an inset is shown in the blind spot but rendered invisible by filling-in (Fig. 2a, ambiguous trials with an inset stimulus). For analysis, we modeled the probability increase of choosing the right stimulus if the right stimulus was presented in either the temporal visual field of the right eye (blind spot) or the nasal visual field of the left eye (non-blind spot). A similar factor was used for the left stimulus. Subsequently, the two one-sided model estimates were collapsed to a single measure of preference for stimulus presented at the nasal or temporal visual field (outside or inside the blind spot respectively). As a model for inference, we used a Bayesian generalized mixed linear model. There were three additional factors in the model (handedness, dominant eye, and precedent answer) that are not the focus of the experiment and are thus reported in the methods section (see "Effects not reported in the Results section"). Figure 2a (ambiguous diagnostic, data in green) and 2b show that subjects indeed presented a bias. However, in contrast to our expectations, subjects were more likely to choose the filled-in percept with a 15.01% preference for stimuli presented in the temporal visual field (CDI₉₅ 8.49%-21.08%). In other words, when subjects had to decide which of the two stimuli (both perceived as being continuous, and in most cases actually physically identical) was less likely to contain an inset, they showed a bias for the one in which the critical information was not sensed directly but inferred from the surrounding spatial context. Remarkably, this result is at odds with both of our experimental predictions that postulated either no bias or a bias in favor of the veridical stimulus.

3.4.2 Experiment 2

The second experiment was designed to replicate the unexpected result of the first experiment and evaluate whether the blind

spot bias observed was due to systematic differences between nasal and temporal retinae. In experiment 1, we presented stimuli at mirror eccentricities inside and outside the blind spot, i.e. temporal and nasal respectively (see Fig. 1C). In experiment 2, we tested whether the bias in experiment 1 was specific to the blind spot location or related to known differences between the temporal and nasal retina (Fahle and Schmid, 1988). There is higher photoreceptor density (Curcio et al., 1990), spatial resolution (Rovamo et al., 1982), luminance discrimination (Pöppel et al., 1973) and orientation discrimination (Paradiso and Carney, 1988) at locations that project to the nasal retina (the temporal visual field where the blind spots are located). Thus, we repeated our experiment with a new group of subjects ($n=27$) and an additional experimental condition. In this new condition, the two stimuli were displayed at symmetrical locations above the blind spot (25° above the horizontal meridian; see Fig. 1D, purple location). The results of this second experiment replicate the observations of experiment 1 (Fig. 3A): subjects showed a bias for selecting the stimulus presented inside the blind spot (12.5%, CDI_{95} 7.35%-17.49%). However, subjects also presented a bias in the control condition, toward the stimuli presented in the temporal visual field above the blind spot (6.63%, CDI_{95} 0.77%-12.3%). The bias was nevertheless stronger inside the blind spot (paired-diff: 6.11%, CDI_{95} 1.16%-10.78%). In summary, additionally to the bias inside of the blind spot area, we observed that subjects also showed a smaller bias for stimuli presented to the nasal retina (temporal visual field).

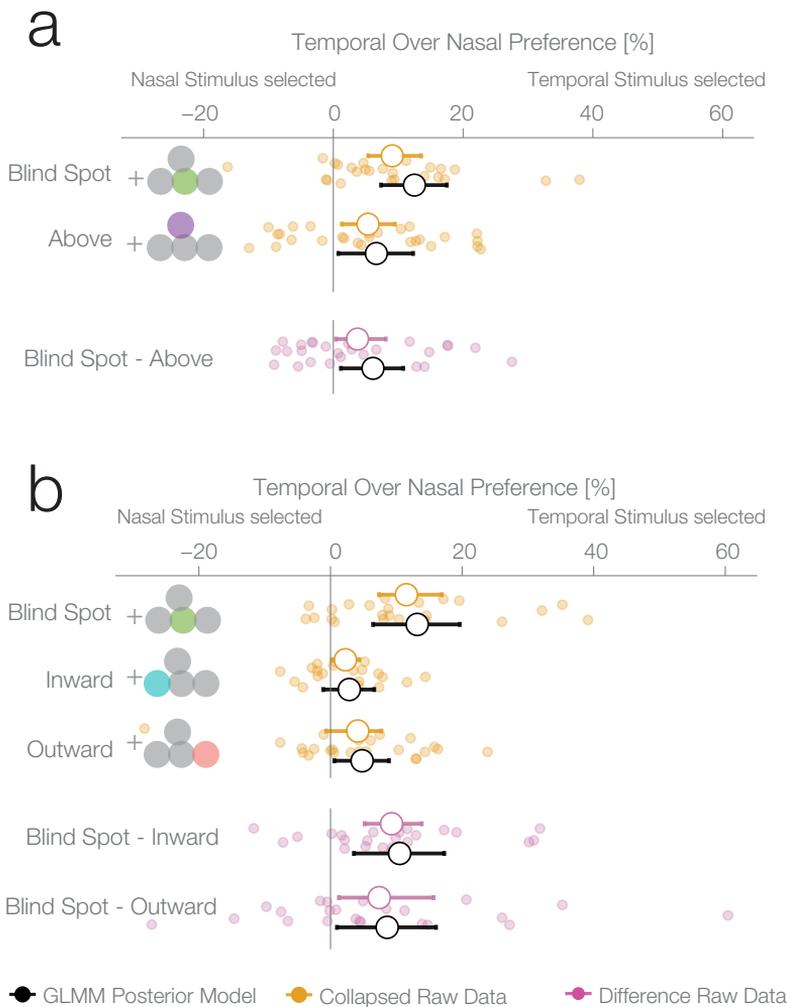


Fig. 3.3 Two control experiments were designed to test whether the observed bias for the blind spot could be explained by a general bias for stimuli presented in the temporal visual field. a) Results of experiment 2. In a given trial, stimuli were presented either at the locations corresponding to the blind spot or at locations above it. Results are presented as in figure 2b, with the addition of within-subject differences between blind spot and control locations (in purple). b) Results of experiment 3. In a given trial, stimuli were presented at the locations corresponding to the blind spot or at locations to inward (toward the fixation cross) or outward (away from the fixation cross) to it. Note that the blind spot effect is replicated in both experiments. In addition, both blind spot effects are larger than in any control location.

3.4.3 Experiment 3

To better delineate the distribution of bias across the temporal visual field and to clarify if the blind spot location is, in fact, special, we performed a third experiment on a new group of subjects ($n=24$). Here, we compared biases in the blind spot to two other control conditions flanking the blind spot region from either the left or the right (Fig. 3B). The blind spot location again revealed the strongest effect of a bias for the temporal visual field (13.18% CDI_{95} 6.47%-19.64%), while the locations inwards and outwards resulted in a 2.85% and 4.8% bias, respectively (CDI_{95} -1.1%-6.65%; CDI_{95} 0.58%-8.89%). The bias of both control locations was different from the bias of the blind spot location (BS vs. inward: 10.51%, CDI_{95} 3.55%-17.29%; BS vs. outward: 8.61%, CDI_{95} 0.98%-16.04%). In this experiment, as in experiments 1 and 2, we observed a bias that is specific to the blind spot region.

3.4.4 Experiment 4

The results of the three previous experiments suggest that subjects considered the filled-in percept a better exemplar of a continuous non-inset stimulus, in disregard of the physical possibility of the presence on an inset inside the blind spot. To confirm this, we performed a fourth experiment with a new group of subjects ($n=25$). This experiment was identical to the first experiment, except that in this case, the subjects' task was to choose the stimulus with an inset, instead of the continuous one. In this case, if a filled-in stimulus is indeed considered a more reliable exemplar of a continuous stimulus, the non blind spot stimulus should be preferred in the diagnostic trials. This was the case; subjects showed a bias for selecting the stimulus presented outside the blind spot (7.74%, CDI_{95} 1.56%-13.68%, Fig. 4a), thus resulting in the expected reversal of the bias pattern observed in the first

three experiments. This pattern is again suboptimal, since this time the filled-in stimulus is the one that could conceal the target. The result of this experiment indicates that the observed biases do not correspond to an unspecific response bias for the blind spot, and instead are a consequence of considering the inferred percepts as more reliable exemplars of a continuous stimulus.

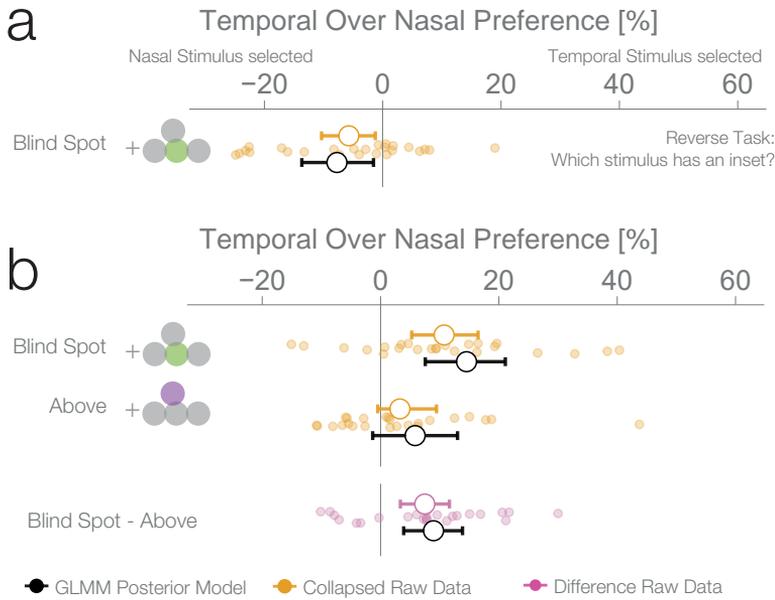


Fig. 3.4 a) Results of experiment 4. This control was the same as experiment 1, except that subjects have to choose the stimulus with an inset (instead of the continuous one). b) Results of experiment 5. This control was similar to experiment 2, except that no inset stimulus was ever experienced in the control location above in the temporal visual field.

3.4.5 Experiment 5

We performed a final control to evaluate whether the observed bias for a filled-in stimulus was not a result of subjects using a probability matching heuristic. It is possible that, in order to solve the ambiguous task, subjects used their knowledge of the rate of

appearance of continuous and inset stimuli at different locations as learned during unambiguous trials. As it is impossible to experience an inset in the blind spot, the base rate of continuous stimuli at that location is 1.0. Therefore, when confronted with two stimuli that are apparently identical, one inside and one outside the blind spot, subjects might just apply the base rate they have learned instead of relying on a perceptual estimate. If this is the case, subjects should show a bias for the location where they experienced exclusively continuous stimuli during unambiguous trials, which could result in a bias pattern similar to the one observed in experiments 1-3. To evaluate this alternative explanation, we performed a further experiment with the same group of subjects that participated in experiment 4. Experiment 5 was similar to experiment 2, with control trials presenting stimuli above the blind spot. However, in contrast to experiment 2, subjects never experienced an inset in the temporal field in the above positions during unambiguous trials (see Figure 1 - Figure supplement 1 for a detailed overview of trial randomization). This results in an identical base rate of occurrence of a continuous stimulus in the temporal field for both the above and blind spot locations. Consequently, if the behavior observed in the previous experiments was a result of probability matching, in this experiment we should observe the same bias at both the blind spot and the temporal field above locations. Subjects showed a bias for selecting the stimulus presented inside the blind spot (14.53%, CDI_{95} 7.56%-21.09%, Fig. 4b), replicating again the results of experiment 1-3. At odds with the probability matching hypothesis, the bias for the temporal field in the above location was only 5.84%, not different from 0 (CDI_{95} -1.33%-13.01%) and similar to what was observed in experiment 2. This bias was different from the bias observed in the blind spot (paired-diff: 8.95%, CDI_{95} 3.91%-13.85%). The same group of subjects participated in experiment 4 and 5, allowing us to make a within subjects comparison between the two tasks. Subjects' performance in these two tasks was negatively correlated ($r = -0.61$, $p = 0.002$, see Figure 4

- Figure supplement 1). Taking the task reversal of experiment 4 into account, this result indicates that subjects were consistently biased to consider the inferred filled-in stimulus a better exemplar of a continuous stimulus. The result of experiment 5 thus gives evidence that the bias for the filled-in stimulus was not a consequence of subjects matching the base rate of the occurrence of different stimuli during unambiguous trials.

3.4.6 Reaction time analysis

A bias for the temporal visual field, especially the blind spot, could also be reflected in the distribution of reaction times. We compared the reaction times of trials where subjects selected a stimulus in the temporal visual field against trials where the stimulus in the nasal visual field was selected. The reaction time analysis was not planned comparisons, thus, in contrast to the other analyses presented here, it is explorative. In the first experiment, we observed an average reaction time of 637 ms (minimum subject average: 394 ms, maximum 964 ms; Fig. 5). We used a linear mixed model to estimate the reaction time difference for selecting a stimulus presented inside the blind spot (temporally) against one outside the blind spot (nasally). In the first experiment, after excluding three outliers, we observed this effect with a median posterior effect size of 13 ms faster reaction times when selecting the blind spot region ($CDI_{95\%}$) 2-42 ms). The three outliers (on the right of the vertical dashed line in Fig. 5) were identified visually and removed because they were distinctively different from the rest of the population. The mean of the outliers was 5.2 SD away from the remaining subjects. The outliers were nevertheless in the same direction of the reaction time effect and did not change its significance (with outliers, 63 ms, $CDI_{95\%}$ 7-124 ms). However, faster reaction times while selecting the blind spot stimulus were not present individually in the other four experiments. The nomi-

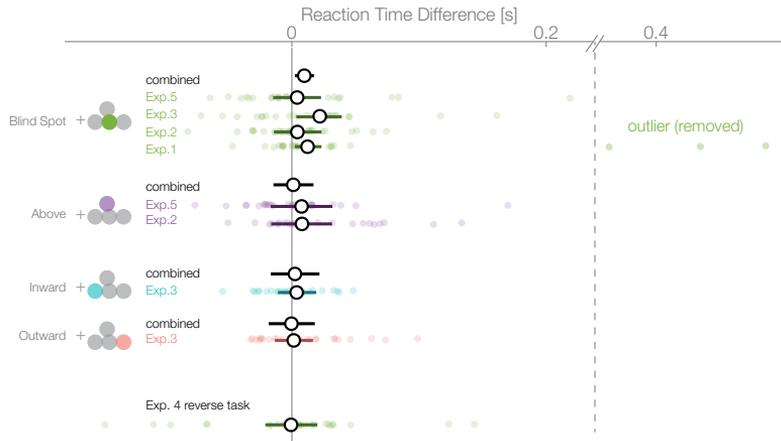


Fig. 3.5 Reaction times of trials where the nasal stimulus was chosen minus the reaction times of trials where the temporal stimulus was chosen. Single subject estimates and 95% CI posterior effect estimates are shown. The black (combined) estimate results from a model fit of all data combined, the individual confidence intervals represent the experiment-wise model fits. We observe a reaction time effect only inside the blind spot.

nal differences were in the same direction as experiment 1 but not significant (Exp.2: 4 ms, CDI_{95} -14-23 ms; Exp.3: 22 ms. CDI_{95} -3-39 ms; Exp.4: -1 ms CDI_{95} -20-21 ms; Exp.5: 4 ms CDI_{95} -15-23 ms). Non-significant results were obtained for the other locations tested (above Exp.2: 8 ms, CDI_{95} -38-53 ms; above Exp.5: 8 ms CDI_{95} -17-32 ms; outward: 2 ms CDI_{95} -13-16 ms; inward: 4 ms, CDI_{95} -29-37 ms). After combining all data (without experiment 4 as the task was reversed), we observed a reduced reaction time for decisions for the blind spot stimulus with 10 ms (CDI_{95} 2-17 ms) but not in any other location. We do not find this small bias in any experiment individually (except Exp. 1) but only after pooling over experiments and therefore, we should interpret it cautiously. In conclusion, subjects selected stimuli in the blind spot slightly faster than stimuli outside the blind spot. The same effect does not appear for the other temporal control locations.

3.4.7 Combined effect estimates over all experiments

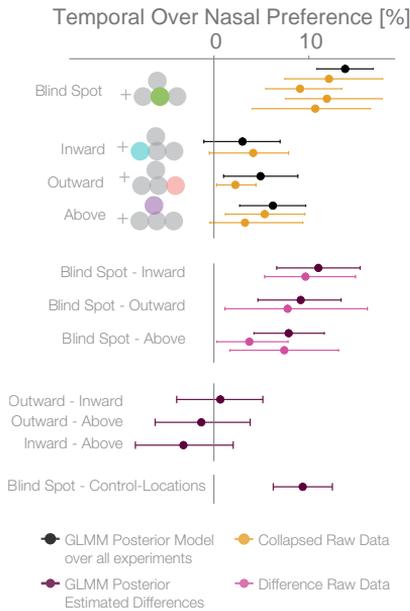


Fig. 3.6 Posterior GLMM-effect estimates of all data combined (black) except experiment 4 (inversed task). We also show for each experiment the 95% CI of bootstrapped means summary statistics of the data (yellow). Next, we show difference values between the blind spot and all other control locations (model dark, raw data pink). As discussed in the text, the control locations outward, inward and above do not differ (4th last to 2nd last row), and thus we compare the blind spot effect to all locations combined (last row).

For an overview of all experiments and the results of a Bayesian logistic mixed effects model that combines all experiments, see figure 6, Figure 6 - Figure supplement 1 and supplementary Table 1. In the combined model, we did not find any differences between the temporal field effects at locations other than the blind spots (Figure 6, 4th last to 2nd last row). That is, the temporal field effects of the locations inward, outward and above were not different from each other. For the sake of clarity, we combined these location levels. Keeping everything else constant, we observed that if we present one stimulus in the blind spot against the equidistant nasal location, subjects are 13.82% (CDI₉₅

10.84%-16.78%, t-test, $t = 8.7$, $df = 98$, $p < 0.001$) more likely to choose the stimulus in the blind spot. This bias is stronger than the effect observed elsewhere in the temporal field by 9.35% (CDI₉₅ 6.25%-12.47%; paired t-test, $t = 4.8$, $df = 74$, $p < 0.001$). In sum-

mary, subjects showed a robust bias for the blind spot locations that could not be explained by a non-specific bias for the temporal visual field. In conclusion, when confronted with an ambiguous choice between veridical and inferred sensory information, human subjects showed a suboptimal bias for inferred information.

3.5 Discussion

When confronted with identical physical stimulation, subjects showed a consistent bias for blind spot inferred percepts which was stronger than the bias at any other location in the temporal visual field. Why do subjects choose the blind spot location when it is objectively the least reliable? Our interpretation takes the results at face value: subjects must possess at least implicit information about whether a percept originates from the blind spot in order to show a bias for or against it. At the same time, the veridical information from the other stimulus is also available. This indicates that perceptual decision-making can rely more on inferred than veridical information, even when there is some knowledge about the reduced reliability of the inferred input available in the brain (Ehinger, König, et al., 2015). This is also supported by the results of the reaction time analyses that indicated a faster evidence accumulation for the inferred percepts. In other words, the implicit knowledge that a filled-in stimulus is objectively less reliable does not seem to be used for perceptual decision-making. This suboptimal decision between qualitatively different veridical and inferred inputs is in contrast to properties of standard sensory integration. There, reduced reliability derived from noisy but veridical signals results in a corresponding weighting of inputs and consequently in optimal decisions (Körding, Beierholm, et al., 2007). In the following, we discuss two potential explanations of this discrepancy of processing filled-in information and standard sensory integration. The first explanation focuses on physiological properties of neuronal and small circuits' response properties at and around the blind spot region. The second explanation addresses the conceptual level and uses the general notion of predictive coding. First, although the filled-in percept is by definition independent of the stimulus within the blind spot, it is nevertheless based on the information sensed by the region around the blind

spot in the nasal retina. We might assume that an area, e.g. in the nasal retina around the blind spot region, that has a lower contrast threshold also shows stronger neuronal signals for super-threshold stimuli. This could in principle lead to a filled-in stimulus with increased salience as compared to the veridical stimulus. Effectively, this explanation proposes that differences in physiological properties of nasal and temporal retinae are transferred to the filling-in process making it the "better" candidate stimulus in an ambiguous condition. Above we already introduced some evidence for psychophysical differences between the nasal and temporal visual field (Fahle and Schmid, 1988). There is also some evidence for the superiority of the blind spot in a Vernier task (Crossland and Bex, 2009). The areas around the blind spot showed greater performance compared to areas at similar eccentric locations in the nasal visual field. It is still unclear whether this goes over and beyond the aforementioned temporal/nasal bias. Unfortunately, this explanation runs into the problem that the sensitivity in the region corresponding to the blind spot in the other eye is also enhanced compared to regions at similar eccentricities (Wolf and Morandi, 1962; Midgley, 1998). This suggests that differences between the eyes in the area around the blind spot should be the smallest within the contrast between temporal and nasal retina. Moreover, we explicitly controlled for temporal-nasal differences in experiments 2 and 3, and found that it is not enough to explain the effect specific to the blind spot. Thus, an explanation of the observed effects based on known differences in retinal properties is currently tentative at best. An alternative explanation is based on the framework of predictive coding (K Friston, J Kilner, et al., 2006; K Friston, Adams, et al., 2012; Summerfield and Lange, 2014). Specifically, context information of static stimuli would be used to predict local stimulus properties leading to the phenomenon of filling-in. The predicted sensory input would then be compared to the incoming sensory input, and an error signal representing the mismatch would be returned. In the absence of veridical informa-

tion, no deviation and thus no error signal would occur. Effectively, the filled-in signal might have less noise. Reduced noise, in turn, results in a smaller prediction error and higher credibility at later stages. A faster reaction time to the filled-in stimulus compared to the veridical stimulus could suggest that the integration process is indeed biased with less noise. In summary, although the results reported here seem compatible with the predictive coding framework, this explanation presently remains vague and speculative. In conclusion, we find a new behavioral effect where subjects prefer a partially inferred stimulus to a veridical one. Though both appear to be continuous, the filled-in one could hide an inset and is, therefore, less reliable. In this perceptual decision-making task, subjects do not make use of high-level assessments about the reliability of the filled-in stimulus. Even more so, they prefer the unreliable percept.

3.6 Methods: An ambiguous 2-AFC task and a Bayesian mixed model

Many of the methods are taken from Ehinger, König, et al. (2015). All data and analyses are available at <https://osf.io/w-phbd>.

3.6.1 Subjects

Overall, 175 subjects took part in the experiments. Of the subjects, 32% (n=56) were removed due to the screening experiments described below. An additional 3% (n=6) were removed due to low performance (n=2, <75% in at least two conditions with a visible unique inset) or because they responded to the stimuli with the inset stimulus instead of the continuous stimulus (n=4). The experimental data were not recorded in 7% (n=13) due to eye tracking calibration problems (n=4) and other issues during data collection (n=9). The remaining 100 subjects were recorded and analyzed in the following experiments. For the first experiment, 24 subjects entered the analysis (average age 21.9 years, age range 18-28 years, 12 female, 20 right-handed, 16 right-eye dominant). 15 of these subjects participated in the EEG study reported by Ehinger, König, et al. (2015). In the second experiment, 27 subjects entered the analysis (average age 22.4 years, age range 19-33 years, 15 female, 25 right-handed, 19 right-eye dominant). In the third, 24 subjects entered the analysis (average age 21.9 years, range 19-27 years, 19 female, 23 right-handed, 16 right-eye dominant). In the fourth experiment, we report the results of 25 subjects (average age 22.1, range 18-35, 20 female, 24 right-handed, 14 right-eye dominant). In the last experiment, the same set of subjects participated as in experiment 4 with the exception of a single subject, who did not finish the both parts of the combined session with

experiment 4 and 5. All subjects gave written informed consent, and the experiment was approved by the ethics committee of the Osnabrück University. For the initial experiment, we set out to analyze 25 subjects. For the second experiment, we calculated a sample size of 18 subjects based on the results of experiment 1 in order to have a power of 90% (calculated with gPower, (Faul et al., 2009), matched pair means cohen's-d = 0.72, planned-power 90%). We disclose that the results of the initial analysis with this group were not conclusive about differences between the location inside and the location above the blind spot. Although the sample size was large enough to replicate the blind spot main effect, it was not adequate to find the difference between locations. Therefore, we decided to increase the number of subjects by 50% (n=9). For the third, fourth and fifth experiments, we used an empirical power analysis based on MLE of a linear mixed model in order to achieve 90% power for the smallest effect observed outside the blind spot. This resulted in a sample of 24 subjects.

3.6.2 Screening

As described above, many subjects failed a simple screening test. In this pre-experiment, we showed a single stimulus in the periphery either inside or outside the blind spot in the left or right visual field. In two blocks of 48 trials, subjects indicated which stimulus (no inset vs. inset) had been perceived. We thought of this simple experiment to evaluate our blind spot calibration method, as an inset stimulus inside the blind spot should have been reported as no inset. The first block was used as a training block. In the second block, we evaluated the performance in a conservative way. No feedback was given to the subjects. If the performance was below 95% (three errors or more), we aborted the session because the participant was deemed to be too unreliable to proceed further with our experiment. We analyzed the data of

those that failed the screening experiment, in four categories of failures that demonstrate the heterogeneity of subjects: Subjects reported inset when an inset was shown in the left blind spot (44%), or in the right blind spot (78%). Subjects did not report the inset of a stimulus presented outside the blind spot (37%), and subjects reported an inset, even though a continuous stimulus was shown (80%). The percentage represents how many subjects had at least one trial where a classification-criterion was fulfilled and thus do not add to 100%. The rates for subjects that did not fail the criterion were 16%, 21%, 13% and 22% respectively. The high percentage in the last category of removed subjects, in which they report an inset even though no inset was visible, strongly suggests that subjects failed the task not due to blind spot related issues, but due to inattention or perceptual problems. Even though we observe more wrong reports in the right than the left blind spot position, there was nevertheless no correlation with calibration position or size. Overall, 57% ($n=100$) of the recruited subjects passed this test and were admitted to subsequent experiments.

3.6.3 Eye Tracking, Screen, Shutter Glasses

A remote, infrared eye-tracking device (Eyelink 1000, SR Research) with a 500 Hz sampling rate was used. The average calibration error was kept below 0.5° with a maximal calibration error of 1.0° . Trials with a fixation deviation of 2.6° from the fixation point were aborted. We used a 24-inch, 120 Hz monitor (XL2420t, BenQ) with a resolution of 1920x1080 pixels in combination with consumer-grade shutter glasses for monocular stimulus presentation (3D Vision, Nvidia, wired version). The shutter glasses were evaluated for appropriate crosstalk/ghosting using a custom-manufactured luminance sensor sampling at 20 kHz. The measured crosstalk at full luminance was 3.94%. The

subject screen distance was 60cm in experiment 1, 2, 4, and 5 and 50 cm in the third experiment.

3.6.4 Stimuli

Modified Gabor patches with a frequency of 0.89 cycles/° and a diameter of 9.6° were generated. Two kinds of patterns were used (Fig. 1A): one completely continuous and one with a small perpendicular inset of 2.4°. For comparison, the blind spot typically has a diameter of 4°-5°. The Gabor had constant contrast in a radius of 6.3° around the center. This ensured the same perception of the continuous stimulus outside the blind spot in comparison to a filled-in stimulus, where the inner part is inside the blind spot. To account for possible adaptation effects, horizontal and vertical stimuli were used in a balanced and randomized way across the trials. Stimuli were displayed using the Psychophysics Toolbox (Brainard, 1997) and EyeLink Toolbox (Cornelissen et al., 2002). The stimuli were displayed centered at the individually calibrated blind spot location. The stimulus at the location above the blind spot in experiment 2 was at the same distance as the blind spot but was rotated by 25° to the horizon around the fixation cross. For the inward and outward condition of experiment 3, stimuli were moved nasally or temporally by 8.6°. Thus the stimuli had an overlap of only 1°. Less overlap is not possible without either cutting the border of the screen or overlapping with the fixation cross.

3.6.5 Task

After a fixation period of 500 ms, we presented two stimuli simultaneously to the left and right of the fixation cross. Subjects were instructed to indicate via button press (left or right) which stimulus was continuous. Each stimulus was presented either in

the temporal or nasal field of view. In some trials, the required response was unambiguous, when one of the stimuli showed an inset and the other did not (and the inset stimulus was presented outside the blind spot). In many trials (80% of all experiments and locations, 46% when the stimulus was shown above the blind spot in experiment 2), both stimuli were continuous and no uniquely correct answer existed (see Figure 1 - Figure supplement 1 for a detailed overview of the balancing). All trials were presented in a randomized order. If the subject had not given an answer after 10 seconds, the trial was discarded, and the next trial started. All in all, subjects answered 720 trials over 6 blocks; in experiment 1 the trials were split up into two sessions. After each block, the eye tracker and the blind spot were re-calibrated. After cleaning trials for fixation deviation and blinks, an average of 662 trials (90%-quantile: 585, 710) remained. For two subjects, only 360 trials could be recorded.

3.6.6 Bootstrap in figures

In several figures, we present data with summary statistics. To construct the confidence intervals we used bias-corrected, accelerated 95% bootstrapped confidence intervals of the mean with 10,000 resamples. Note that the summary statistics do not need to conform to the posterior summary estimates because they are marginals. Only the posterior model values reflect the estimated effect.

3.6.7 Blind spots

In order to calibrate the blind spot locations, subjects were instructed to use the keyboard to move a circular monocular probe on the monitor and to adjust its size and location to fill the blind spot with the maximal size. They were explicitly instructed to

calibrate it as small as necessary to preclude any residual flickering. The circular probe flickered from dark gray to light gray to be more salient than a probe with constant color (Awater, 2005). All stimuli were presented centered at the respective calibrated blind spot location. In total, each subject calibrated the blind spot six times. For the following comparisons of blind spot characteristics, we evaluated one-sample tests with the percentile bootstrap method (10,000 resamples) of trimmed means (20%) with $\alpha = 0.05$ (Wilcox, 2012). For paired two-sample data, we used the same procedure on the difference scores and bias-corrected, accelerated 95% bootstrapped confidence intervals of the trimmed mean (20%). We report all data combined over all experiments. In line with previous studies (Wolf and Morandi, 1962; Ehinger, König, et al., 2015), the left and right blind spots were located horizontally at -15.52° (SD=0.57° CI:[-15.69°, -15.36°]) and 15.88° (SD=0.61° CI:[15.70°, 16.07°]) from the fixation cross. The mean calibrated diameter was 4.82° (SD=0.45° CI:[4.69°, 4.95°]) for the left and 4.93° (SD=0.46° CI:[4.79°, 5.07°]) for the right blind spot. Left and right blind spots did significantly differ in size ($p=0.009$, CI:[-0.17°, -0.03°]) and in absolute horizontal position (in relation to the fixation cross; $p<0.001$, CI: [0.27°, 0.45°]). On average, the right blind spot was 0.36° further outside of the fixation cross. No significant difference was found in the vertical direction ($p=0.37$), but this is likely due to the oval shape of the blind spot in this dimension and the usage of a circle to probe the blind spot. These effects seem small, did not affect the purpose of the experiments and will not be discussed further.

3.6.8 GLMM analysis

We fitted a Bayesian logistic mixed effects model predicting the probability of responding right with multiple factors that represent the temporal over nasal bias and several other covariates

described below. Because we were interested in the bias between the nasal fields and the temporal fields of view, we combined both predictors for the left and right temporal (and nasal, respectively) locations and reported the combined value. Data were analyzed using a hierarchical logistic mixed effects models fitted by the No-U-Turn Sampler (NUTS, STAN Development Team). The model specification was based on an implementation by Sorensen and Vasisth (Sorensen et al., 2016). In the results section, we report estimates of linear models with the appropriate parameters fitted on data of each experiment independently. We also analyzed all data in one combined model: there were no substantial differences between the results from the combined model and the respective submodels (Appendix table 1). The models are defined as follows using the Wilkinson notation:

```
answer_Right ~ 1 + Temporal_Left*Location +
                Temporal_Right*Location +
                Answer{Right ,( t-1)} +
                Handedness_Right +
                DominantEye_Right +
                (1 + Temporal_Left*Location +
                Temporal_Right*Location +
                Answer_{Right ,( t-1)} | Subject)
```

And in mathematical terms:

$$Answer_{i,right} \propto Bernoulli(\Theta_i)$$

$$\theta_i = \text{logit}^{-1}(X_{within}\beta_{within} + X_{between}\beta_{between} + N(0, \tau X_{within}) + N(0, e))$$

Two factors were between subjects: handedness and dominant eye. In total, we have four within-subject factors, resulting in eight parameters: There are two main factors representing whether the left, and respectively the right, stimulus was inside or outside the

temporal field. Depending on the experiment, the main factor location had up to three levels: the stimuli were presented outward (Exp. 3), inward (Exp. 3), above (Exp 2, 5) or on (all experiments) the blind spot. In addition, we modeled the interactions between location and whether the left stimulus (and the right stimulus, respectively) was shown temporally. In order to assure independence of observation, an additional within-subject main factor $\text{answer}(t-1)$ was introduced, which models the current answer based on the previous one. In frequentist linear modeling terms, all within-subject effects were modeled using random slopes clustered by subject and a random intercept for the subjects. We used treatment coding for all factors and interpreted the coefficients accordingly. In the model, we estimated the left and right temporal field effects separately. For the statistical analysis, we combined these estimates by inverting the left temporal effect and averaging with the right temporal effect. We did this for all samples of the mcmc-chain and took the median value. We then transformed these values to the probability domain using the inverse-logit function, subtracting the values from 0.5 and multiplying by 100. All results were still in the linear range of the logit function. We calculated 95% credible intervals the same way and reported them as parameter estimates (CDI_{95} lower-upper) in the text. These transformed values represent the additive probability (in %) of choosing a left (right) stimulus that is shown in the left (right) temporal field of view compared to presenting the left (right) stimulus in the nasal field of view, keeping all other factors constant.

3.6.9 Reaction times

Initially, we did not plan to analyze the reaction time data. These analyses are purely explorative. The response setup consisted of a consumer keyboard. Thus delays and jitters are to be expected. However, with an average of 494 ambiguous trials per subject,

we did not expect a spurious bias between conditions due to a potential jitter. Reaction time data was analyzed with a simple Bayesian mixed linear model:

$$RT \sim 1 + \text{Temporal_selected} * \text{Location} + \\ (1 + \text{Temporal_selected} * \text{Location} \mid \text{Subject})$$

Only trials without a visible inset stimulus were used. Temporal selected consists of all trials where a temporal stimulus was selected. Because of the bias described in the results, there was no imbalance between the number of trials in the two conditions (difference of 10 trials bootstrapped-CI [-2, 23]).

3.6.10 Bayesian fit:

We did not make use of prior information in the analysis of our data. We placed implicit, improper, uniform priors from negative to positive infinity on the mean and 0 to infinity for the standard deviations of our parameters, the default priors of STAN. An uninformative lkj-prior (=2) was used for the correlation matrix, slightly emphasizing the diagonal over the off-diagonal of the correlation matrix (Sorensen et al., 2016; B Carpenter et al., 2017). We used six mcmc-chains using 2000 iterations each, with 50% used for the warm-up period. We visually confirmed convergence through autocorrelation functions and trace plots, then calculated the scale reduction factors (Gelman, Hwang, et al., 2014), which indicated convergence ($R_{hat} < 1.1$).

3.6.11 Posterior predictive model checking

Posterior predictive model checks were evaluated to test for model adequacy (Gelman, Carlin, et al., 2013). Posterior predictive checks work on the rationale that newly generated data based on

the model fit should be indistinguishable from the data that the model was fitted by originally. Due to our hierarchical mixed model, we perform posterior predictive checks on two levels: trial, and subject. In the first case, we generate new datasets (100 samples) based on the posterior estimates of each subject's effect. We compare the distribution of this predicted data with the actual observed values for each (Figure 6 - Figure Supplement 2 A). At the subject level, we draw completely new data sets, based on the multivariate normal distribution given by the random effects structure. We then compare the collapsed blind spot effect once for the newly drawn subjects with the observed data (Figure 6 - Figure Supplement 2 B). Taken together, these posterior predictive model checks show that we adequately capture the very diverse behavior of our subjects but also correctly model the blind spot effect on a population basis.

3.6.12 Effects not reported in the result section

Here we report the result of the covariate factor based on the combined model (all experiments modeled together). Note that the interpretation of such effects naturally occurs on logit-transformed values. Summation of different parameter-levels (as necessary for treatment coding) on logit-scale can be very different to summations of raw-percentage values. It can also be similar, close to the linear scale of the logit-transform, i.e. close to 50% (which we made use of for the blind spot effects reported at other points of the manuscript). We did not find evidence for a different global bias (main effect location) in any of the four stimulation positions tested here. The dominant eye factor had an 11.51% effect (CDI₉₅ 2.78% - 19.59%) on the global bias. Thus subjects with a dominant right eye also showed a preference to the right stimulus over the left one, irrespective of whether the stimulus was visible through the left or the right eye. We find a global bias

(the intercept, -26.75% CDI_{95} -38.18% - -9.29% , with treatment coding) toward choosing the left stimulus; this might reflect that in the first two experiments we instructed subjects to use the right hand, thus they used their index and middle fingers. In the third experiment we instructed subjects to use both index fingers, resulting in a decreased bias to the left, with a shift more to the right, and thus more to balanced answers, of 12.24% (CDI_{95} -1.98% - 24.16%). We did not find evidence for a bias due to handedness (7.71% , CDI_{95} -8.96% - 22.75%). There was an influence of the previous answer on the current answer. We observe a global effect of 7.86% (CDI_{95} 0.53% - 14.95%), which suggests that subjects are more likely to choose e.g. the right stimulus again when they have just chosen right in the previous trial. For this effect it is more important to look at random effect variance, which is quite high with a standard deviation of -31.4 (CDI_{95} 28.27% - 34.69%), suggesting that there is a large variation between subjects. Indeed, a closer look at the random slopes of the effect reveals three different strategies: Some subjects tend to stick the same answer, some subjects are balanced in their answers without any trend, and some subjects tend to regularly alternate their answers in each trial. Note that this behavior does not seem to influence any of the other effects: We do not see any correlation between the random effects, except for the correlation between the $n-1$ effect and the intercept (-0.55 , CI: -0.72 , -0.34). This correlation means that subjects who tend to alternate their keypresses will not have a strong bias in the intercept, or the other way around, subjects who press the same key all the time also have a bias towards this key. Other extended models we considered showed no effect when both stimuli were in the temporal field, nor any three-way interaction. Following standard procedures to avoid spurious effects of unnecessary degrees of freedom, we removed these variables from the final model.

Action and Expectations: Predictions over eye movements and filled-in percepts

” *I enjoyed reading this clever and compelling study of prediction error responses using a saccadic paradigm and monocular presentation within the blind spot.*

— Reviewer #1

Contributions

Journal of Neuroscience: Ehinger, König, and Ossandón 2015
BVE, PK, and JPO designed research. **BVE** performed research.
BVE and JPO contributed unpublished reagents/analytic tools. **BVE**
and JPO analyzed data. **BVE**, PK, and JPO wrote the paper.

Journal of Eyemovement research: König, Wilming, Kietzmann,
Ossandón, Onat, Ehinger, Gameiro, and Kaspar 2016
PK, NW, TCK, JPO, SO, **BVE**, RG and KK wrote the review manuscript.

4.1 Layman summary

In Chapter 3 we discussed how the reliability of a signal is incorporated into a perceptual decision. In this chapter we again look at the same question but from a different angle: What happens if we make an eye movement?

We perceive the world as a stable, continuous visual stream; we very rarely move the eye without a rough idea of what we will see next. Some of these rough ideas, also called predictions, will be more reliable, others less. If such a prediction was wrong we will be surprised. For example, if we make an eye movement and recognize that, what we thought was a dark bush, is actually a bear, we are surprised.

In this chapter we discuss an experiment that investigates whether we show differently strong surprise signals to wrong predictions if the reliability of the prediction is low or high. While recording brain activity, we sometimes exchanged the stimulus during an eye movement. Then, the peripheral prediction before the eye movement is violated and we hypothesized that the brain produces a surprise signal. Indeed our experiment showed that if the peripheral prediction is wrong, this results in a surprise signal. Next we changed how reliable the predictions are. For this, we used the same trick as in Chapter 3 - the blind spot - where the brain fills in the percept. Therefore, the information for the prediction is now completely generated by the brain, without additional incoming information. In our experiment we showed that this also influences our surprise signal. In contrast to Chapter 3, it seems we are less surprised when the prediction is unreliable.

4.2 Predictions of visual content across eye movements and their modulation by inferred information

The brain is proposed to operate through probabilistic inference, testing and refining predictions about the world. Here, we search for neural activity compatible with the violation of active predictions, learned from the contingencies between actions and the consequent changes in sensory input. We focused on vision, where eye movements produce stimuli shifts that could, in principle, be predicted. We compared, in humans, error signals to saccade-contingent changes of veridical and inferred inputs by contrasting the electroencephalographic activity after saccades to a stimulus presented inside or outside the blind spot. We observed early (<250 ms) and late (>250 ms) error signals after stimulus change, indicating the violation of sensory and associative predictions, respectively. Remarkably, the late response was diminished for blind-spot trials. These results indicate that predictive signals occur across multiple levels of the visual hierarchy, based on generative models that differentiate between signals that originate from the outside world and those that are inferred.

4.3 Introduction: Eye movements test visual hypotheses

The brain is likely to operate constructively, generating probabilistic models of reality that are in continuous testing against sensory inputs. Functionally, probabilistic models can successfully explain a large range of phenomena, like perceptual illusions (Weiss et al., 2002), spontaneous activity representing uncertainty (Fiser et al., 2010), and the optimal integration of multimodal signals (D Wolpert et al., 1995; Ernst and Banks, 2002; Körding and DM Wolpert, 2004). To find neural correlates of such processes, researchers look for patterns of brain activity compatible with probabilistic neural computation, in which predictive coding is one of the most popular models. In predictive coding, higher areas in a brain hierarchy predict the activity of lower areas by inhibitory feedback, while lower areas generate corresponding error signals in relation to their own feedforward inputs. In current formulations of predictive coding, the precision-weighted prediction errors are thought to be encoded predominantly in superficial pyramidal cells of the cortex (Feldman and KJ Friston, 2010; Bastos et al., 2012), and thus measurable by electroencephalography (EEG). Previous EEG experiments, which revealed neural signatures compatible with predictive coding, have mostly relied on passive tasks, in which the predictability of the stimuli is imposed externally. However, the predictive coding framework can also embrace predictions that are the consequences of agents' self-generated actions, in line with recent proposals of embodied cognition that emphasize the role of the body and self-generated action for perception (AK Engel et al., 2013). Eye movements can be considered as experiments in the visual domain, testing hypotheses about visual content through actions (K Friston, Adams, et al., 2012). Given that there is evidence for predictive coding in early visual areas for passive stimulation (Murray et al., 2002; Summerfield, Trittschuh, et al.,

2008; Alink et al., 2010; Kok et al., 2012), it is conceivable that the shifts of the visual input produced by eye movements could, in principle, result in predictable signals in all levels of the visual hierarchy. Moreover, active sensory predictions could also exist for signals that are generated in the absence of actual inputs. This occurs naturally in the retina's blind spot, which is demonstrable in monocular vision as a percept that is filled-in from the surroundings' content. We combined eye-tracking and EEG measurements to evaluate the existence and timing of predictive signals that are caused by human subject's eye movements. Crucially, in our experimental design we measured prediction error responses in the context of both veridical (precise) sensory information and inferred (imprecise) visual cues, presented outside and within the blind spot respectively. Two alternatives are conceivable in the case of blind spot stimulation: First, feedforward activity of neurons related to filling-in is taken by the brain as if it was actual input, and therefore, no differences should exist between prediction violations inside or outside the blind spot. Alternatively, within the brain's generative models, there is an expected uncertainty about the blind quality of filled in information; therefore, we would expect to see an attenuated error response when the violations were based upon imprecise filling-in, when stimuli were presented in the blind spot, relative to when they were not.

4.4 Methods: Gaze-dependent stimulus changes and inferred stimuli

4.4.1 Overview

To find signals compatible with predictive errors that can be differentiated from non-predictive remapping signals, we compared EEG responses to stimuli that were changed or unchanged during the saccade that brought it to the center of the gaze. Our study design has effectively two key factors (see Figure 1); namely, a stimulus change during the saccade (or not) and an initial presentation of the stimulus (pre-saccadic) within the blind spot (or not). The stimulus change involved rotating the inner segment of a circular grating to create an inner visual feature, in which the direction of the grating was orthogonal to the surround. Crucially, this visual feature (inset) was smaller than the blind spot resulting in perceptual filling-in when presented within the blind spot. We presented stimuli within the blind spot using monocular stimuli (by using shutter glasses), but alternatively to the right or left eye. This resulted in a design with four factors in total: stimulus Change (present or absent), Inset (present or absent), Blind spot (within or without) and Position (peripheral initial presentation right versus left).

4.4.2 Subjects

Fifteen subjects participated in the study (mean age: 22.5y [18-28], 1 of whom was left-handed, and 6 of whom had a left-dominant eye; 9 were female). All subjects gave written consent, and the experiment was approved by the local ethics committee. An additional 9 subjects were rejected before their EEG recording

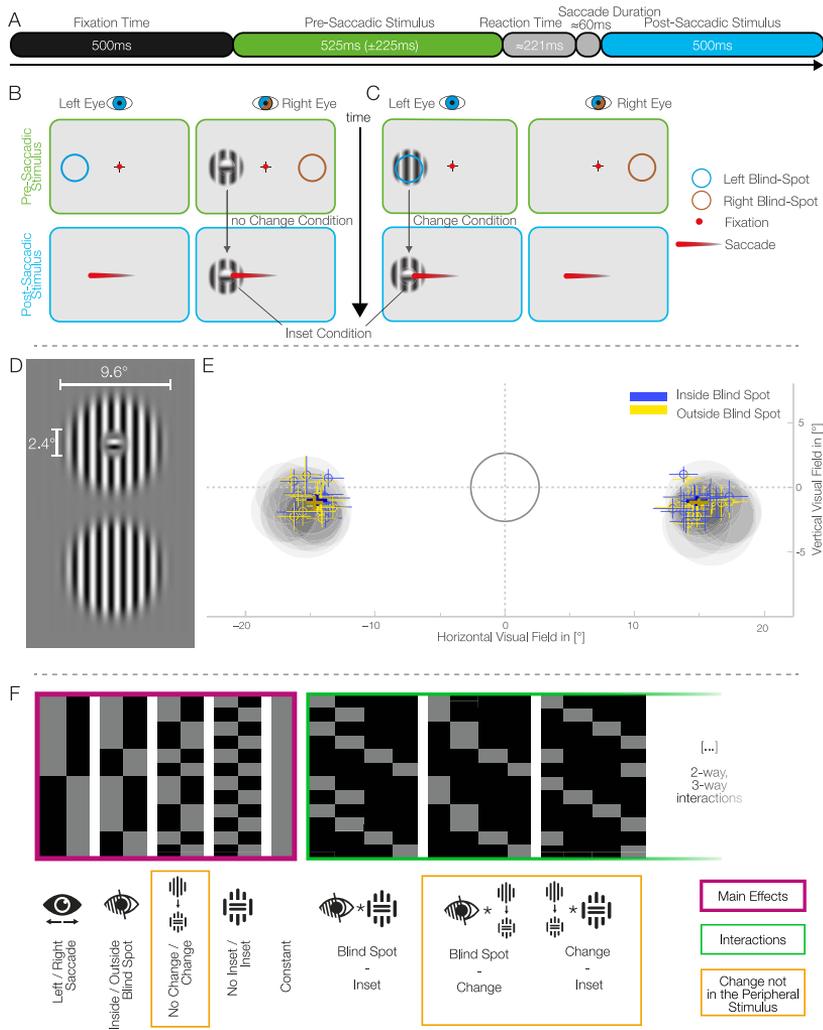


Fig. 4.1 A) Trial time course BC) Each set of two panels represents what is presented to each eye with the shutter glasses. After a fixation interval, a stimulus appeared monocularly in the periphery (upper panels). After the disappearance of the fixated crosshair, the subjects perform a saccade to the center of the pre-saccadic stimulus, which becomes the post-saccadic stimulus (lower panels). The colored circles represent the location of the blind spot in each eye and were not displayed on the screen. B) An example of a trial without change: the inset stimulus, presented outside the blind spot, does not change before and after the saccade. Importantly, presenting an inset stimulus inside the blind spot always leads to fill-in and therefore the perception of a continuous stimulus. We therefore recorded the inset no-change condition only outside the blind spot. C) A trial with change: the continuous stimulus, presented inside the blind spot, is exchanged during the saccade to an inset stimulus. D) Gabor patches used as stimuli. Horizontal stimuli were also used. The inset was set to approximately 50% the diameter of the blind spot.

Fig. 4.1 E) Calibration of the blind spot and saccades' end-points. The gray ring encloses the tolerance area for fixation. The gray discs represent the average calibrated blind spot sizes and locations for each subject. Bold crosses represent the winsorized average saccade end locations over subjects of both inside and outside blind spot trials (+/- winsorized STD). Small crosses show the same metrics for each individual subject. F) The design matrix used for the single subject GLMs. An overparameterized model of four main effects (purple), constant, and all interactions (green) were used.

either due to the screening procedure ($n = 4$, see below for criteria); technical problems ($n = 2$); incompatibility of lenses with the combination of shutter glasses and eye tracker ($n = 2$); or perceptual problems in the peripheral field of view ($n = 1$).

4.4.3 Materials

EEG:

Electrophysiological data were recorded using 64 Ag/AgCl electrodes with an equal-distance placement system (actiCap, Brain Products GmbH, Germany). Scalp impedances were kept below 5 kOhm. EEG data were sampled with 1000 Hz, using Cz as a recording reference, and the ground electrode was placed near Fz.

Eyetracking:

A remote, infrared eye-tracking device (Eyelink 1000, SR Research Ltd., Mississauga, Canada) with a 500 Hz sampling rate was used. The average calibration error was kept below 0.5° with a maximal calibration error below 1.0° . Trials with a fixation deviation of more than 4.6° from the fixation point were aborted.

Display:

We used a 24, 120Hz monitor (XL2420t, BenQ, Taipei, Taiwan) with a resolution of 1920x1080 pixels in combination with consumer-grade shutter glasses for monocular stimuli presentation (3D Vision, Nvidia, Santa Clara, USA, wired version). The shutter glasses were evaluated for appropriate crosstalk/ghosting using a custom-manufactured luminance sensor sampling at 20 kHz. The measured crosstalk at full luminance was 3.94%.

Change Latency:

As the main analysis of EEG data was about signals related to saccade-contingent changes, we needed to make sure that the stimulus change always occurred during the saccades. The online detection of a saccade by the eye-tracker took on average 27ms (SD: 1ms, 5/95-percentile: [22ms 35ms]) from the movement start, and the saccade duration was on average 60ms (SD: 4ms, 5/95-percentile: [48ms 80ms]). An additional 8.75ms (max: 11ms) delay occurred from the computer command to the actual stimulus change on the monitor. The slowest detection of a saccade (35ms) plus the maximum time it took to change the stimulus (11ms) was faster than the shortest saccade (48ms). Thus, the stimulus was always exchanged before the fixation onset. Reaction time from go signals to saccade start was, on average, 248ms (SD: 20ms).

Stimuli:

Modified Gabor patches with a frequency of 0.89 cycles/° and a diameter of 9.6° were generated. Two kinds of patterns were used (Figure 1D): one completely continuous and one with a small

perpendicular inset of 2.4° . For comparison, the blind spot typically has a diameter of $4\text{--}5^\circ$. The Gabor had constant contrast in a radius of 6.3° around the center. This ensured the same perception of the continuous stimulus outside the blind spot compared to a filled-in stimulus where the inner part is inside the blind spot. In order to account for possible adaptation effects, horizontal and vertical stimuli were used in a balanced and randomized way over trials. Stimuli were displayed using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007) and Eyelink Toolbox (Cornelissen et al., 2002).

4.4.4 Experiments

Calibration of Blind Spot

In order to calibrate the blind spots, subjects were instructed to use the keyboard to move a circular monocular probe on the monitor and adjust the size and location to fill the blind spot with maximal size. They were explicitly instructed to calibrate it as small as necessary to preclude any residual flickering. The circular probe flickered from dark gray to light gray in order to be more salient than a probe with constant color (Awater, 2005). All stimuli were presented centered at the respective calibrated blind spot location. In total, each subject calibrated the blind spot 30 times over two sessions. In line with previous studies (e.g., Wolf and Morandi 1962), the blind spots (left and right) were located horizontally at -15.4° (SD: 0.6°) and 15.7° (SD: 0.6°) from the fixation cross. The mean calibrated diameter was 4.9° (SD: 0.7°) for the left and 5.0° (SD: 0.5°) for the right blind spot. Blind spots did not significantly differ in size ($p = 0.061$, CI: $[-0.3, 0.0]$) but they did differ in absolute horizontal position (in relation to the fixation cross) ($p = 0.005$, CI: $[-0.5^\circ -0.1^\circ]$) with the right blind spot, on average, 0.3° further outside from the fixation cross. There was

no difference in the vertical position ($p = 0.87$, CI: $[-0.5^\circ \ 0.3^\circ]$). In summary, the properties of the subjects' blind spots were fully compatible with the previously reported values. One exception is that subjects calibrated the left blind spot, compared to the right one, closer to the fixation cross. Eye movements result in EEG artifacts that differ according to their kinematics. Although we used state-of-the-art procedures to remove this artifact (see below), it is important to evaluate the systematic differences in eye movement that could confound the analysis. The saccades did not differ in amplitude for saccades to the left against saccades to the right. There was a significant difference of amplitude between saccades inside the blind spot and saccades outside the blind spot of 0.3° (SD: 0.1° , $p=0.001$, CI: $[0.1^\circ, 0.4^\circ]$). However, this difference was small compared to the overall average saccade amplitude of 13.9° , and the stimulus size of 9.6° .

Screening Procedure

The screening procedure was used to ensure a normal fill-in, absence of problems in the peripheral vision unbeknownst to the subjects themselves, and the ability to sustain a high level of attention. A single stimulus, either continuous or with an inset, was monocularly presented in the periphery at the previously determined blind-spot location (inside the blind spot, temporally) or in the horizontally mirrored position (outside the blind spot, nasally). Subjects indicated via button press whether they perceived a stimulus without inset (left key) or a stimulus with inset (right key) stimulus. A total of 48 trials per block were shown, and they were fully balanced and randomized. We applied conservative criteria, requesting a 94% performance level in this simple classification task. If an inset stimulus was presented inside the blind spot and thereby eliciting fill-in, it was counted as a correct

trial when subjects answered that they perceived the stimulus as continuous.

Responses to saccade contingent changes outside and inside BS Concurrent EEG and eye-tracking recordings were performed allowing us to induce artificial mismatches between pre- and post-saccadic stimuli and thus evaluate the differences in EEG responses to saccade-contingent changes. At the start of the trial (Figure 1 A-C), the subjects were asked to fixate on a cross in the middle of the screen. A Gabor stimulus, the pre-saccadic stimulus, was presented monocularly either to the left or to the right eye and either outside the blind spot (nasally) or inside the blind spot (temporally). After an average of 525 ms (\pm 225 ms), the fixation cross disappeared, and the subjects had to perform a saccade to the center of the stimulus. We called the second stimulus, now in the center of the gaze, the post-saccadic stimulus. Two key factors were evaluated during the experiment. The first factor was Change. In order to induce a prediction error, we exchanged the stimulus during the saccade in half of the trials. This change was either from a stimulus with the inset present to one where the inset is absent or vice versa. Saccades were detected online when the gaze deviated more than 2.6° from the fixation cross. Saccade-contingent changes occurred equally often for movements to stimuli presented inside or outside the blind spot. The second factor, Inset, related to whether subjects saw a stimulus with or without an inset. Presenting the stimulus with an inset in the blind spot elicits fill-in and thus is perceived as a continuous stimulus, irrespective of the veridical physical stimulus properties. It is therefore impossible to perceive an inset stimulus when the initial pre-saccadic stimulus is presented in the blind spot; thus, we did not record trials in such a condition. Moreover, an inset in the later post-saccadic stimulus, when the previous periphery stimulus was inside the blind spot, can only co-occur with a change. In total, 2880 trials were displayed over two sessions with 10 blocks

per session. Each condition was displayed in a fully balanced and randomized way for each block.

4.4.5 Analysis

EEG Processing

Data were analyzed using MATLAB and EEGLAB (Delorme and Makeig, 2004). Data were resampled to 500 Hz and bad channels, which we identified by visual inspection (never more than one channel per subject) were excluded from further analysis and interpolated at a later stage (after data epoching, see below) using spherical interpolation. Signals were cleaned visually for coarse motor artifacts and signal drops. An independent-component analysis (AMICA, standard parameters as implemented in BCILAB version v12, Palmer et al., 2008) was applied on, only for this step, FIR high-pass filtered data (1 Hz, -6 dB cutoff at 0.5 Hz, 1 Hz transition bandwidth, FIRFILT, EEGLAB plugin). ICs were automatically screened for artifacts. For eye artifacts, we employed an automatic reliable algorithm (Dimigen et al., 2011; Plöchl et al., 2012) that removed, on average, 7.5 eye-artifact ICs (SD: 2.5, [lo:4 hi:18]) per subject. For muscle-artifact ICs, we correlated the spectrum of the ICs with a prototypical square-root spectrum commonly observed in muscle artifacts. A correlation higher than 0.7 was used to identify an IC as a muscle-artifact IC. We found, on average, 6.8 components per subject (2.3 [lo:0 hi:15]). All rejected ICs were also visually validated by inspecting the topographies, spectra and activation over time, and confirmed. Finally, EEG data were low-pass filtered below 50 Hz (-6 dB cutoff at 56.25 Hz, 12.5 Hz transition bandwidth) using a FIR filter. Data were cut into epochs of -300 ms to 500 ms with a baseline of -300 to -100 ms and aligned to two different events: the onset of the pre-saccadic stimulus and the start of the post-saccadic stimulus fixation. Trials

were excluded from further analysis in three circumstances: subjects made a saccade while they should have maintained fixation, the saccade endpoint was not in an area of 10° around the center of the stimulus, or the reaction time was greater than three times the standard deviation above the mean for each subject. After online fixation control and EEG cleaning, on average, 88% (4.9%, [lo:62.9% hi:96.5%]) of the trials entered further analysis. The complete eye-tracking and EEG datasets, and analysis scripts, are available upon request.

Statistics

We used robust statistics wherever possible. Robust statistics are more reliable in the case of small deviations from assumed distributions than their classical statistical counterparts (Wilcox, 2012, Chapter 1). If not stated otherwise, all reported descriptive values are 20% winsorized mean followed by 20% winsorized standard deviation in round brackets (Wilcox, 2012, Chapter 3.3) . When winsorizing, the upper and lower 20% of samples are replaced by the remaining most extreme values, and then the mean or standard deviation is calculated. The influence of outliers is thereby strongly attenuated. The median, arguably least affected by outliers, is equal to the most extreme winsorized mean (threshold of 50%), where all values, except the median value, are declared outliers'. Ranges are reported by [lo:X hi:Y]. We evaluated one-sample tests with the percentile bootstrap method of trimmed means (20%) with $\alpha = 0.05$ (Wilcox 2012, p.115-p.116). For paired two-sample data, we used the same procedure on the difference scores. We used bias-corrected, accelerated 95% bootstrapped confidence intervals of the trimmed mean (20%) and reported them in the text by (CI:[X Y]). All bootstrap tests and estimates were done with 10.000 resamples, except for the EEG multiple-comparison correction (TFCE) where we used only 1.000 resamples.

Single Subject GLM

Using a MATLAB toolbox suitable for mass-univariate generalized linear models (GLM; LIMO toolbox Pernet et al., 2011), GLMs were fitted on each electrode and time-point separately for each subject. The analysis of EEG data with GLMs, has previously shown advantages in terms of higher sensitivity and unbiased data-driven analysis (e.g. Rousselet et al., 2011; Dandekar, Privitera, et al., 2012) and is a standard application of statistical parametric mapping (Litvak et al., 2011). An overparameterized dummy coding with interaction comparisons was used for the design matrix (Figure 1F). The main factors used were stimulus Change, Blind spot (outside/inside), Position (left/right), and Inset (with and without). All possible interactions were modeled. The analyzed estimable functions are linear combinations of the parameters of the same experimental factor or interaction. For example, for the main factor Position, the estimable function was $\beta_2 - \beta_1$. We tested this statistically using Yuen's t-tests with corresponding H0-centered bootstraps over subjects. For an interaction example, Position \times Blind spot, the estimable function was $(\beta_{10} - \beta_9) - (\beta_{12} - \beta_{11})$ and tested by a Yuen's t-test with corresponding H0-centered bootstraps over subjects. For the pre-saccadic analysis, we did not model the change factor, as the stimulus was not exchanged in that time window of the analysis.

Group-Level Statistics

We used a standard threshold-free cluster enhancement (TFCE) measure of local responses and permutation testing (SM Smith and Nichols, 2009; C Pernet et al., 2015) to control the elevated familywise error ratio of the multiple electrodes*timeframes tests that were performed. In brief, TFCE bypasses the need to define an

arbitrary threshold for sample clustering by establishing the local support in space and time for every sample. This local support is given by the sum of all sections in time and space that are underneath it in other words, the sum of all samples that are not beyond and higher than any local minima that is between them and the sample under calculation. TFCE was calculated for the actual mean factors and for 1,000 bootstraps of centered subjects. For each of these bootstrapped samples of subjects, the maximum TFCE value across all samples in time and space is used to construct an H_0 distribution, against which the actual TFCE values were compared. Values above the 95th percentile were considered to be controlled for multiple comparisons at an alpha level of 0.05. The neighborhood distance was calculated on the default electrode locations. For a given model and data partition, the procedure described above controls for the familywise error rate resulting from fitting multiple GLMs to different electrodes and time points. We report only effects that extended over five samples (10ms) or more. Effect clusters of significant values were reported in the text with their respective timing and median TFCE-corrected p-value (\tilde{p}).

Shift of factor labels

Due to the nature of our factorial design, one can re-label the factors such that main effects and interactions are exchangeable. For example, a main effect of change can be regarded as an interaction between the presence of an inset before and after the saccade: An inset stimulus before the saccade combined with a change results in a no-inset stimulus after the saccade. Without the change, the stimulus before and after the saccade is identical. This is true in reverse for the no-inset stimulus before the saccade. We, indeed, found a Change x Inset interaction effect (Figure 4B) even before the saccade ended. We observed this pre-saccade-offset positive

effect from -64 to 6 ms (positive betas: -64 to 6 ms, $\tilde{p} = 0.019$, min-p = 0.004; negative betas: -60 to -34 ms, $\tilde{p} = 0.045$, min-p = 0.03, and from -30 to 0 ms, $\tilde{p} = 0.042$, min-p = 0.035). This seems puzzling at first, as the change of the stimuli occurred approximately 30 ms before the saccade offset, which is after this observed interaction of Change x Inset. The effect before saccade offset most probably resembles the main effect of the inset during the pre-saccadic stimulus stimulation and can be explained by this changing of factor-labels.

Incomplete Design

Due to the very nature of filling in at the blind spot, only an incomplete factorial design is possible: an inset stimulus inside the blind spot cannot be perceived. Therefore, we have to assume that the three-way interaction containing the Blind spot factor, the Change factor, and the factor Inset is negligible. We additionally confirmed all results in two reduced subset models, each containing a full-factorial design. For the first full-factorial model, we collapsed the change and inset factor in a combined factor and selected only trials that were available both inside and outside the blind spot. The second model excluded the factor Blind spot (and all of the data with a stimulus inside the blind spot) but included separate factors Change and Inset. This allowed us to confirm our full-factorial model with the limitation of two independent error terms for the two models, instead of one error term. All results were confirmed in the fully balanced designs.

4.5 Results: Early and late prediction errors and an interaction with inferred information

Fifteen subjects participated in a concurrent EEG and eye-tracking experiment. They were asked to perform an eye movement from the center of the screen onto a Gabor stimulus (Figure 1D) located in the periphery, in either the left or the right visual field (Figure 1A-C). The position of the stimulus was centered at either the left or right blind spot of each individual subject and the diameter of the stimulus itself was larger than the blind spot by a factor of approximately 2 (see Methods, Calibration of Blind Spot). By using a 3-D monitor with shutter glasses, it was possible to present the stimuli either in the blind spot or in the same location, albeit in the nasal non-blind spot field of the other eye. Uninstructed to the subjects, in half of the trials, the stimulus changed during the saccade. Post experiment debriefing established that all subjects became aware that the stimulus was sometimes exchanged. This consisted in a change of a small inset within the center of the Gabor patch (smaller than the blind spot) from no inset (a completely continuous stimulus) to the stimulus with an inset (during inside and outside blind spot trials) or vice versa (only during outside blind spot trials). This experimental manipulation allowed us to evaluate the presence of EEG responses compatible with error signals to changes in a stimulus that are contingent on subjects' actions. We analyzed electrophysiological responses that are produced after identical eye movements (Figure 1E) with the same amplitude, direction, and target location (see Methods, Calibration of Blind Spot) and that resulted in the foveation of identical stimuli, albeit between trials in which the stimulus was either changed or not changed during the eye movement. EEG responses were analyzed with mass-univariate general linear mod-

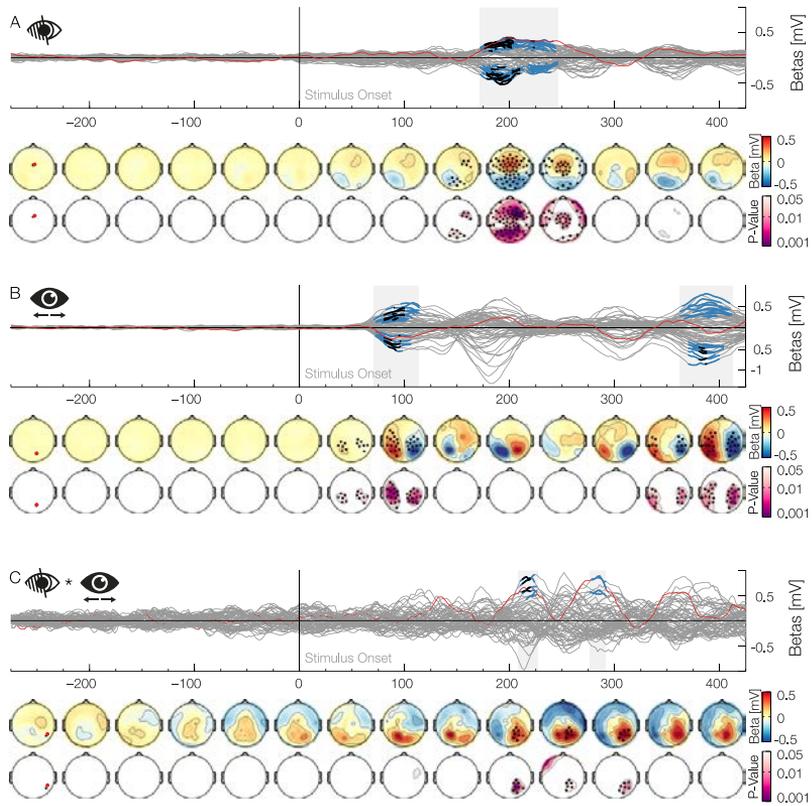


Fig. 4.2 Time-series plots of the EEG beta weights of the main factors or interactions for each electrode are shown (butterfly plot), aligned to the onset of the pre-saccadic stimulus. Blue marked latencies are significant under TFCE alpha of 0.05 and, therefore, are corrected for multiple comparisons over time points and electrodes. Black marked latencies are significant under additional Bonferroni correction for the testing of multiple factors in a model. This second procedure is overly conservative and only done to evaluate the robustness of the effects. The first row of the topographical plots represents the mean beta weights averaged over 50 ms bins. The second row depicts the minimal TFCE-corrected p-values over the same bin. Black marked electrodes represent significant channels. The location of the red highlighted channel is depicted in the first topographic plot. A) The main effect Blind spot depicts the difference of a pre-saccadic stimulus presentation inside and outside the blind spot, which is prominent 200 ms after stimulus onset. B) The main effect of Position shows a contralateral processing in occipital electrodes to a stimulus presented in the periphery. C) The Blind Spot CE Position Interaction depicts a lateral component of the effect shown in A).

els (CR Pernet et al., 2011) that were fitted to each electrode and time-point for each subject. The main factors used were as follows: stimulus Change (with/without), Blind spot (outside/inside), Position (left/right saccade), and Inset (with/without) (Figure 1F). On the group level, instead of a classical ERP analysis, we analyzed the estimable contrasts of the parameter estimates from the general linear model of the main factors and interactions. This corresponds to a summary statistic approach to random effects analyses. The reported values, termed beta values as a shorthand, are similar to difference values of raw ERP, but importantly, they take into account the variance of the other independent variables. We found clusters of significant effects associated with the main effect of interest, Change, at four distinctive latencies in the analysis of the data after the saccade offset. The following results are organized in three sections. First, we describe EEG effects of filling-in before a saccade was done. Second, we illustrate the effects related to late prediction signals post saccade. Finally, we describe early and middle latency effect of post-saccadic predictive signals.

4.5.1 Inferred information in the periphery

A stimulus presented in the blind spot elicits filling in, in which visual signals are inferred from surrounding information without a direct input from the retina or the outside world. Neural activity related to filling in has been described previously in areas V1 to V3 in neurophysiological studies with primates (Fiorani Júnior et al., 1992; Komatsu et al., 2000; Matsumoto and Komatsu, 2005) and in fMRI experiments with humans (Tong and SA Engel, 2001) but not in human electrophysiology. We found that, prior to the eye movement, after the onset of a peripheral stimulus, there is a main effect for the Blind spot factor (Figure 2A) from 172 to 246 ms (positive betas: 172 to 244 ms, median p -Value $\tilde{p} = 0.015$, min- $p = 0.001$; negative betas: 172 to 214 ms, $\tilde{p} = 0.011$, min- p

= 0.002, and from 216 to 246 ms, $\tilde{p} = 0.028$, min-p = 0.011; positive values represent a positive deviation from the average ERP over all conditions and vice versa). As the pre-saccadic stimulation was lateralized, we analyzed the main factor Position (Figure 2B) and found two significant effects. The first occurred immediately after the pre-saccadic stimulus onset from 70 to 114 (positive betas: 70 to 114 ms, $\tilde{p} = 0.015$, min-p = 0.003; negative betas: 72 to 106 ms, $\tilde{p} = 0.017$, min-p = 0.005). As to be expected, a stimulus in the right or left periphery elicits stronger activation in the contralateral occipital electrodes. A second effect can be seen from 362 to 408 (positive betas: 362 to 408 ms, $\tilde{p} = 0.022$, min-p = 0.01; negative betas: 368 to 404 ms, $\tilde{p} = 0.018$, min-p = 0.004). Importantly, processing of the blind spot filling-in could be lateralized as well. Therefore, we analyzed the interaction Blind spot CE Position (Figure 2C) and found two significant positive effects. The first was from 210 to 226 ms ($\tilde{p} = 0.022$, min-p = 0.003), and the second one was from 276 to 292 ms ($\tilde{p} = 0.03$, min-p = 0.011). The first of the lateralized effects started and ended during the main effect, whereas the second one started about 30 ms after the main effect. The overall blind spot results are in line with previous studies with intracranial recordings of V1 neurons that have receptive fields which include the blind spot, and that show differences in activation after 100 ms (see Figure 9B in Matsumoto and Komatsu, 2005) or after 200 ms (see Figure 9A in Komatsu et al., 2000). These two effects shown here establish an EEG correlate for the difference in visual processing of peripheral stimuli when they are veridical (outside the blind spot) or inferred (inside the blind spot).

4.5.2 Prediction signals over saccades

Our main goal was to find an EEG effect compatible with error signals related to the prediction of specific visual content across

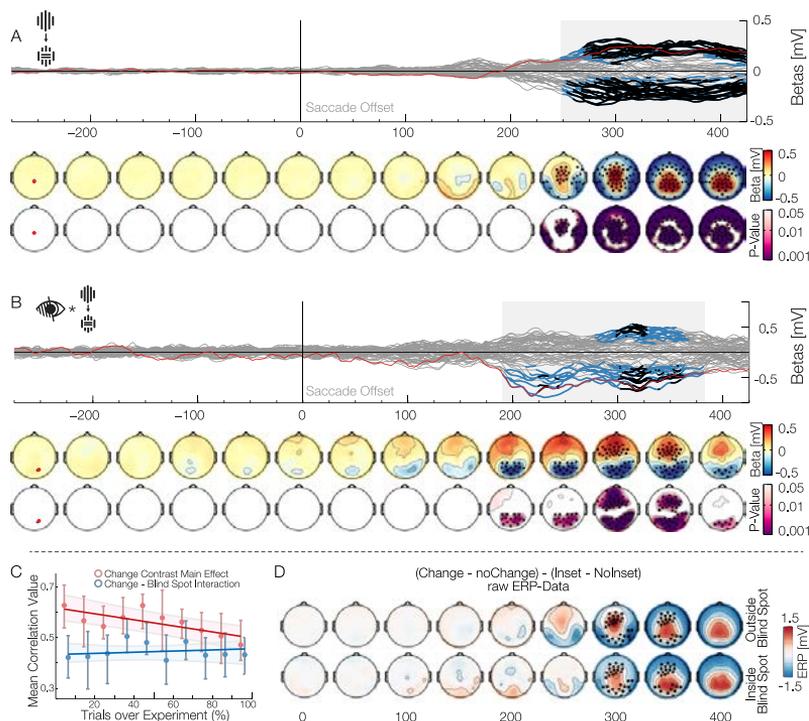


Fig. 4.3 Data aligned to saccade offset when the stimulus is foveated. A) The main effect Change is shown, comparing trials in which the stimulus remained the same, with trials, where it changed during the saccade. The effect resembles a prediction error in the form of a P3. B) The interaction Blind spot CE Change is shown. This shows a reduction of the prediction error described in A). C) Correlation of each subject effect-template with ERP data over 10 partitions of the experiment. The red correlation shows that the main effect of change habituates over the course of the experiment. The blue correlation shows no significant increase or decrease for the Change CE Blind spot interaction and thus stays stationary over the experiment. D) Raw ERP data of the interaction depicted in B). The upper row depicts outside the blind spot data and the lower row depicts inside the blind spot data. The difference of the change effect inside blind spot against outside blind spot was tested using a bootstrapped Yuen's t-test and corrected for multiple comparisons using TFCE. The significant electrodes and time-points can be seen as black dots. This confirms the reduction of the P3 inside the blind spot.

eye movements. In contrast to the results described above, all subsequent effects are for the post-saccadic stimulus. The violation of a predictive signal that is independent of specific low-level visual content (which would show in the interaction with Position or Inset) can be investigated by the main effect of the Change factor. Such a main effect was present from 248 to 498 ms after the end of the saccade (positive betas: 248 to 498 ms, $\tilde{p} = 0.001$, min-p < 0.001; negative betas: 250 to 482 ms, $\tilde{p} < 0.001$, min-p < 0.001). Taken together, the topographies and timing of these positive and negative effects were compatible with a P3 ERP (Figure 3A). The P3 component is usually found after infrequent or unexpected events, independently of sensory modality. Our data are therefore consistent with a high-level prediction error, associated with a prediction based on peripheral visual input and the subjects' eye movement. After establishing the presence of a signal compatible with post-saccadic prediction error, we investigated whether this error signal was different depending on whether the pre-saccadic visual input was veridical (outside the blind spot) or inferred (filling-in inside the blind spot). The Change x Blind spot interaction (Figure 3B) was significant from 190 to 382 ms (negative betas: 190 to 382 ms, $\tilde{p} = 0.007$, min-p < 0.001; positive betas: 276 to 368 ms, $\tilde{p} = 0.01$, min-p < 0.001). In order to understand the direction of the interaction effect, we additionally analyzed the raw ERP difference between the conditions change and no-change, once inside and outside the blind spot. To evaluate these data independently of effects due to the inset, we subtracted the inset difference ERP from the change conditions separately for trials inside and outside the blind spot (Figure 3D). The resulting ERPs show how the interaction modifies the P3 component in two different ways. First, and corresponding to the positive cluster of the interaction, there is a reduction of the anterior part of the P3 when the previously peripheral stimulus was shown inside the blind spot compared with when it was shown outside. This correspond to the P3a subcomponent, which is associated with orienting responses due to the process-

ing of unexpected, novel events (D Friedman et al., 2001; Polich, 2007). And second, corresponding to the negative cluster of the interaction, there is an increase in the posterior part of the P3 for blinds pot trials. This corresponds to the P3b subcomponent, which has been associated to several different processes like changes in episodic context (Donchin and Coles, 2010), the processing of statistical surprise (Mars et al., 2008; Kolossa et al., 2013) and the updating of perceptual evidence (O'Connell et al., 2012; Wyart et al., 2012; Kelly and O'Connell, 2013; Cheadle et al., 2014). In summary, this demonstrates that the brain treats violations of predictions differently, depending on whether they are based on external or inferred information. We controlled for non-stationarity effects, such as those due to learning, by investigating the Change \times Blind spot interaction over the course of the experiment. As a comparison, we looked into a similar correlation but with the main Change effect. The interaction is only estimable over multiple trials, therefore, we used trial partitions of the Change \times Blind spot interaction (the difference of differences) and the difference of change and no change for the change main effect. We partitioned the whole experiment into 10 parts and for each part calculated the corresponding ERP difference. We used the interaction beta weights (over time and electrodes) of each subject as a template and correlated these 10 parts with the template resulting in 10 correlation values for each subject. (Figure 3C). For the Change main effect, we observed a significant negative slope of correlation values against trial partition order (winsorized mean slope: -0.012, $p = 0.001$, bootstrap-CI: [-0.022 -0.005]). This indicates that the P3 amplitude diminishes over time because subjects get used to the experimental setting and become habituated to the saccade-contingent change (Ravden and Polich, 1998; Ravden and Polich, 1999). However, the slope over trials for the Change \times Blind spot interaction was not significantly different from zero (winsorized mean slope: 0.002, $p = 0.8$, bootstrap-CI: [-0.009 0.012]). Hence, subjects got used to the task and were less surprised by the saccade

contingent change, but there was no habituation of the Change CE Blind spot interaction, indicating that the modified error signal for filled-in content was a stationary effect. Besides the Change CE Blind spot interaction, we found one additional significant late interaction between the factors Change and Inset (Figure 4B). The interaction was significant from 334 ms to at least 500 ms (positive betas: 342 to 436 ms, $\tilde{p} = 0.016$, min-p = 0.005 and from 440 to >500 ms, $\tilde{p} = 0.035$, min-p = 0.019; negative betas: 334 to 428 ms, $\tilde{p} = 0.008$, min-p = 0.004, from 434 to 468 ms, $\tilde{p} = 0.036$, min-p = 0.013 and from 472 to >500 ms, $\tilde{p} = 0.039$, min-p = 0.026). We interpreted this interaction as a consequence of the imbalance in the experimental design (see Methods, Incomplete Design). We showed more continuous than inset stimuli in the periphery ($\frac{2}{3}$ to $\frac{1}{3}$) due to the physical limitation, as an inset stimulus in the blind spot is necessarily perceived as the stimulus without inset. Due to this imbalance, a change from continuous to an inset stimulus is twice as frequent as a change from inset to a continuous stimulus. This less frequent change, the one to a continuous stimulus, resulted in an increased response compatible with a higher surprise.

4.5.3 Middle and early prediction signals

The main and interaction effects presented above emerge late after the post-saccadic foveation of the stimulus and showed a topography that is consistent with a known high-level associative component. As such, those effects are unlikely to be related to trans-saccadic prediction signals of specific low-level visual input. We search for evidence for this kinds of predictions based on two criteria: effect latency and interaction with low-level stimulus features. Specifically, we would first expect that sensory error signals would be different whether the change was from pre-saccadic no-inset to post-saccadic inset stimuli or vice versa. And second, the

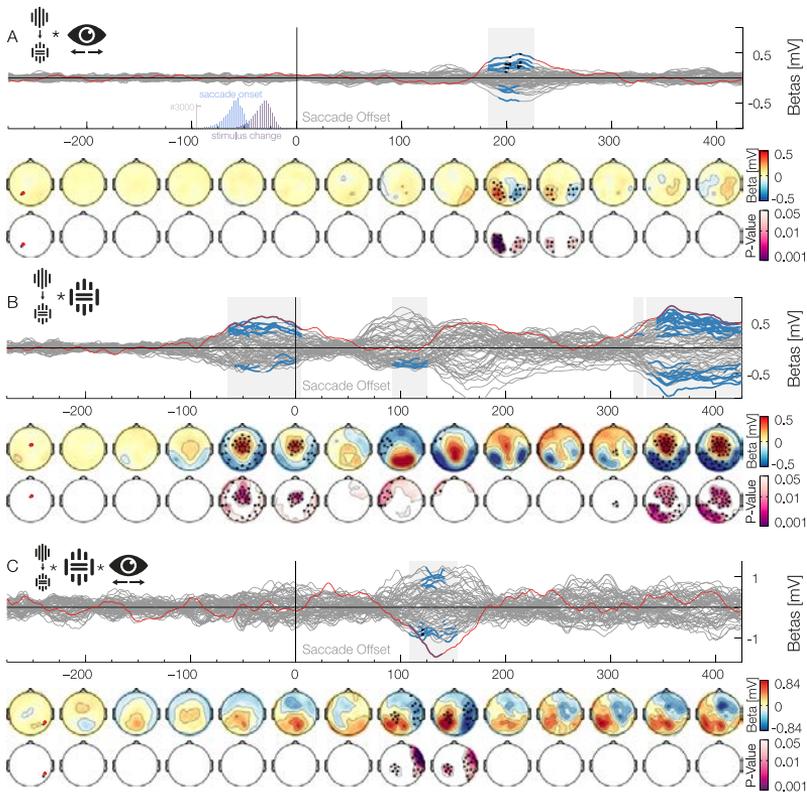


Fig. 4.4 Data aligned to the saccade offset when the stimulus is foveated A) The Change CE Position interaction shows a clearly lateralized prediction error at around 200 ms. The two histograms insets depict saccade onset and stimulus change, and show, that all changes occurred during the saccades. B) The Change CE Inset interaction shows three separate effects. Detailed descriptions are found in the results section. C) The Change CE Inset CE Position interaction shows an early prediction error that is lateralized and also dependent on the low-level stimuli properties.

pre-saccadic position of the stimulus should also have an effect due to the communication of prediction and error signals between unilateral hemispheric areas and the bi-hemispheric central representation of the post-saccadic foveal stimulus. The Change effect had a lateralized component, which was dependent on the stimulus location previous to the eye movement, even though the stimulus was then fixated centrally: we observed a significant Change \times Position interaction (Figure 4A) from 182 to 226 ms (positive betas: 182 to 226 ms, $\tilde{p} = 0.009$, min-p < 0.001; negative betas: 192 to 212 ms, $\tilde{p} = 0.029$, min-p = 0.015, and from 216 to 226 ms, $\tilde{p} = 0.039$, min-p = 0.027). Even earlier effects were observed in an interaction between Change and Inset factors and in the three-way interaction Position \times Change \times Inset. The interaction Change \times Inset (Figure 4B) was significant only for negative values, around 100 ms (90 to 126 ms, $\tilde{p} = 0.03$, min-p = 0.011). Lateralized effects of this two-way interaction were found in the Position \times Change \times Inset interaction (Figure 4C): we found a positive effect from 108 to 154 ms ($\tilde{p} = 0.03$, min-p = 0.001) and a similar effect, albeit negative, on the other hemisphere from 120 to 142 ms ($\tilde{p} = 0.035$, min-p = 0.006). Note that the interactions that include the factor Position result in topographic effect similar to the one observed for the pre-saccadic stimuli (Figure 2B), thus suggesting that the processing of post-saccadic foveal stimulus include extensive crosstalk with the areas that represented it before a movement was done. Altogether, these middle and early effects of the interactions between saccadic-contingent change, pre-saccadic position and low-level characteristics of the stimulus indicate the production of error signals to low-level visual predictions.

4.6 Discussion

We investigated the EEG correlates of prediction errors to changes in visual stimulation occurring during an eye movement. We found early (100 ms), middle (200 ms), and late (300 ms) latency surprise responses to changed stimuli. These responses can neither be explained exclusively in terms of remapping operations, as movement vectors were equivalent across conditions, nor by differences in the post-saccadic stimulus because we compared responses to identical stimuli. Early and middle latency responses were either lateralized and/or dependent on the Inset factor, suggesting that their sources are processes specific to the visual sensory domain. In contrast, the late latency response resembled a P3 ERP component, thus suggesting it has a source in a high-order process, possibly non-visual, that it is related to the occurrence of an unexpected event. Such dissociation between perceptual and associative predictions have been shown previously in human EEG, with comparable methods, topographies and time courses, but in the context of perceptual decision-making experiments (Wyart et al., 2012; Cheadle et al., 2014). The late response we observed was also present when the stimulus was located inside the blind spot, but it was reduced in amplitude compared to a response outside the blind spot. This modified response to a change in the blind spot indicates knowledge of the unreliability of the filled-in information, occurring only at a late stage of processing.

We observed early, middle, and late latency effects following a change of the stimulus during a saccade. Both early and middle effects are compatible with modality specific processes: they interact with the side of the field in which the movement is directed (early and middle component) and the orientation contrast of the stimulus (early component only). Previously, sensory error signals have been associated in EEG experiment with the mismatch nega-

tivity potential (MMN), both in the auditory (Garrido, JM Kilner, Stephan, et al., 2009) and visual domain (Stefanics et al., 2011). In the visual domain, orientation mismatch, comparable to our saccade-contingent change in orientation of the stimulus, result in MMN with a temporal and topographic profile similar to the early and middle latency change and change interactions seen here (Astikainen et al., 2008). These early effects are, however, unlikely to be related to processes occurring in the primary visual cortex. At 100 ms, the first feedforward-feedback sweep in V1 has already occurred (e.g. Hupé et al., 2000; Chen et al., 2007). Furthermore, the C1 visual ERP component, which originates in V1 (Di Russo et al., 2002), has a latency of 50-60 ms and peaks around 90 ms, which is mostly prior to the effects seen here. The absence of a very early error signal is in concordance with the current knowledge of the extent of remapping in visual areas. Responses related to remapping are present at high levels in the visual hierarchy and are almost absent in the primary visual cortex (Nakamura and Colby, 2002; Merriam et al., 2007), making it less likely that predictions related to eye movement reach this area. The prediction of foveal, high-spatial frequency content in the primary visual cortex would be, in any case, inefficient due to the limited spatial spectral resolution of peripheral information (e.g. SJ Anderson et al., 1991) and the limited accuracy of eye movements (e.g. Weber and Daroff, 1971). This favors a more restricted role of predictive coding for the primary visual cortex, in which only statistical regularities about the world and the effect of low-level spatial context are taken in account (e.g. Srinivasan et al., 1982; Rao and DH Ballard, 1999; Fiser et al., 2010), rather than an adaptive, all-encompassing process that also predicts specific content at all cortical areas. Nevertheless, the change-related effects starting at 100 and 200 ms are likely due to other stages of low- and middle-level visual processing. Previous studies that attempted to uncover error signals to an unpredicted sensory input have been inconclusive regarding how early, or how upstream, prediction signals are produced. In some

studies, only late signals related to prediction errors were found. For instance, in the modeling of mismatch negativity signals of auditory stimuli, only the late evoked responses, P3-like here, can be clearly attributed to inhibitory feedback prediction errors (Garrido, JM Kilner, Kiebel, et al., 2007). Similarly, a time-frequency analysis shows that high-frequency gamma differences, attributable to prediction errors, start only late, after 200 ms (Todorovic et al., 2011), which is the same latency for the effect that occurs in most ERP repetition suppression experiments (reviewed in Grill-Spector et al., 2006). In contrast, other experiments support early prediction errors. Neuroimaging studies show activity pattern in primary visual cortex consistent with predictive signals for pattern adaptation (Grill-Spector, Henson, et al., 2006; Summerfield, Trittschuh, et al., 2008; Egnér et al., 2010), apparent motion (Alink et al., 2010), and 3D grouping (Murray et al., 2002). In electrophysiological experiments, evidence for early predictions exist for the auditory mismatch negativity effect (Wacongne et al., 2011), and in the case of a reduction of EEG sensory responses due to self-stimulation (Martikainen et al., 2005). In this last kind of experiment the subjects trigger themselves the appearance of a standard stimulus (Schafer and Marcus, 1973) through a motor action coupled artificially, by experimental design, with the appearance of the stimulus. In the case of our experiment, the stimulus was always present, and the subjects' actions were directly related to the modality of stimulation. The stimulus changed its retinotopic location and resolution due to a shift of the visual field that follows over-learned sensorimotor contingencies expected for any eye movement. The absence of an early interaction between Change and Blind spot factors indicate that filled-in signals, which are inferred from neighbor inputs, are processed by visual areas as if they were the result of an actual input. Such interaction only emerges at a later stage in the form of a modulation of both anterior and posterior subcomponents of the P3. This supports the idea that the exclusively inferred quality of the signals from the blind spot is not lost, and it is taken

into account in higher-level associative areas. Even though most of the research showing visual-related P3 components has been done in conditions without eye movements, P3 responses have been recently described in experiment that permit them (Dandekar, Ding, et al., 2012; Kamienkowski et al., 2012; Kaunitz et al., 2014). In contrast to these experiments, in which an infrequent item or a search target produced the P3 component, here the P3 was elicited by a movement-contingent stimulus change. Whereas the MMN potential discussed above is related to low-level sensory processing, the P3 is considered a correlate of a high-level error signal. For instance, P3 responses seem to be a response associated to the processing of global deviants in a stimulus or event succession rather than to local deviants (Bekinschtein et al., 2009; Chennu et al., 2013), thus depending on the episodic context rather than in specific sensory features (Donchin and Coles, 2010). A formal interpretation of the P3 is that it corresponds to the processing of statistical surprise (Mars et al., 2008; Feldman and KJ Friston, 2010; Kolossa et al., 2013), and also, especially for the posterior subcomponent, to the update of perceptual evidence (O'Connell et al., 2012; Wyart et al., 2012; Kelly and O'Connell, 2013; Cheadle et al., 2014). The reduction of the anterior subcomponent (peaking at 300 ms) in blind spot trials is consistent with predictive coding simulations of attention (Feldman and KJ Friston, 2010). In these simulations, top-down estimates of reliability (precision) modulate the gain of prediction error units in lower regions of the visual hierarchy. This gain modulation would correspond to attention, where high-precision signals enjoy greater gain and the P3 represents a revision of these precision estimates. Changes in visual stimuli need to be associated with movements or transients to be detected (Grimes, 1996; Henderson and Hollingworth, 2013). Because this is prevented here by precise experimental timing, it is safe to assume that saccade-contingent changes, even if task irrelevant, are novel events that would result in a revision on the reliability of estimates of stimulus or event constancy, at least in

the local context of the experiment. However, in the case of information from the blind spot there is no such context that can change the intrinsic unreliability of filled-in signals, thus resulting in a reduced revision of conditional expectations about the stimulus being stationary. This means that an interpretation of our findings is that we attend away (ignore) visual information from the blind spot; thereby attenuating subsequent responses to violated predictions, when resampling the visual scene.

A modulation of the posterior subcomponent of the P3 in blind spot trials can be observed as well. In contrast to the modulation of the anterior subcomponent, this difference is not fully consistent with the temporal progression and topography of the change effect. We consider two alternative interpretations, not necessarily incompatible, of this posterior interaction. First, although the blind spot region remains unreliable, our experimental design could result in a revision, not of the precision estimates, but of the underlying model of the causes of sensory input from this region, normally inferred from the surrounding. In other words, in blind spot trials there is a revision not only of the expectation about the stimulus being stationary (albeit reduced in comparison to trials outside the blind spot in which the veridical pre-saccadic stimulus has a high reliability), but also of the filling-in model. The second interpretation follows the results of EEG experiments about perceptual decision-making (O'Connell et al., 2012; Wyart et al., 2012; Kelly and O'Connell, 2013; Cheadle et al., 2014), in which a similar posterior topography and time course to the one observed here is seen for the updating of perceptual evidence, indicating that the posterior cluster of the interaction could represent an updating of the perceptual evidence about the contents of the blind spot location. Crucially, these are updates of a decision signal instead of a perceptual one, and thus in our experiment, even in absence of an explicit task goal, would represent the accumulation of new evidence against the filling-in percept being a reliable model of

sensory input. There are two main conclusions of the present work. First, the sensory consequences of eye movements are actively predicted at multiple levels of the visual hierarchy. This occurs for the prediction of the actual visual content that is present before an eye movement, rather than only for the prediction of general statistics of visual content. Second, the prediction of content that is exclusively inferred differs between levels of the visual hierarchy. Low-level sensory areas process the stimulus as if it originated from an external input source. In contrast, in higher-level processing, the filled-in (and therefore imprecise) nature of the blind-spot information is taken into account. These results suggest that a hierarchy of predictions does not operate in a strictly successive way, in which prediction and errors necessarily propagate all the way down and up, respectively; low- and high-level predictions of the same content can be dissociated.

s

Action and Sampling: Eye movements as an exploration-exploitation dilemma

” *The hypotheses are a bit trivial but worthwhile testing.*

— Reviewer #2

Contributions

Journal of Vision: Ehinger, Kaufhold, and König 2018

BVE LK and PK conceived the study. **BVE** and LK recorded the data and performed the analyses. **BVE** and PK wrote the manuscript.

5.1 Layman Summary

In the previous chapters we saw, that whenever we perform an eye movement, we predict what the future percept will be. Moving our eyes and predicting the future input takes effort. The brain therefore developed efficient ways to select future points to move the eyes to.

The brain juggles two decision processes at the same time: when to stop collecting data from the current view and move the eyes, and where to move the eyes to. Researching these decision processes is difficult, because subjects decide in their own time when and where to make eye movements to.

In this chapter, we present a new paradigm that allows us to experimentally control both the where and when of eye movements, but we focused mostly on the when. We showed small parts of a stimulus to subjects, as if they were looking through a single hole of a Swiss cheese. After a controlled fixation time, we exchanged the hole and thereby forced subject to look at a different part of the stimulus. Thereby, we control both the where and when of eye movements and can examine the state of the decision processes. We could show two main influences on the time to move the eye. For one, the decision time depends on how much information was already extracted. If subjects looked at a stimulus already in great length, they are ready to move on earlier. Also, if we allowed eye movements on more than one future location, subjects showed a longer decision time to select between these options.

5.2 Probing the temporal dynamics of the exploration-exploitation dilemma of eye movements

When scanning a visual scene we are in a constant decision process regarding whether to further exploit the information content at the current fixation or to go on and explore the scene. The balance of these two processes determines the distribution of fixation durations. Using a gaze-contingent paradigm, we experimentally interrupt this process to probe its state. Here, we developed a guided-viewing task where only a single 3° aperture of an image ("bubble") is displayed. Subjects had to fixate the bubble for an experimentally controlled time (forced fixation time). Then, the previously fixated bubble disappeared, and one to five bubbles emerged at different locations. The subjects freely selected one of these by performing a saccade toward it. By repeating this procedure, the subjects explored the image. We modeled the resulting saccadic reaction times (choice times) from bubble offset to saccade onset using a Bayesian linear mixed model. We observed an exponential decay between the forced fixation time and the choice time: Short fixation durations elicited longer choice times. In trials with multiple bubbles, the choice time increased monotonically with the number of possible future targets. Additionally, we only found weak influences of the saccade amplitude, low-level stimulus properties, and saccade angle on the choice times. The exponential decay of the choice times suggests that the sampling and processing of the current stimulus were exhausted for long fixation durations, biasing toward faster exploration. This observation also shows that the decision process took into account processing demands at the current fixation location.

5.3 Introduction: Where and when to look at things?

Decisions are a central aspect of cognition. Because the world is noisy, evidence of the states of the world needs to be integrated over time (Vickers, 1970). Indeed, physiological studies using the random dot motion paradigm (Shadlen, Britten, et al., 1996; Shadlen and Newsome, 2001; PL Smith and Ratcliff, 2004; Gold and Shadlen, 2007) suggest such an evidence accumulation process. Furthermore, several other decision processes comply with such models. Examples are categorization (Heekeren, Marrett, Bandettini, et al., 2004), eye movements (Leach and R Carpenter, 2001), or self-initiated button presses (Schurger et al., 2012). Most commonly, these decision processes are modeled by using a biologically plausible drift-diffusion process (Ratcliff, 2001). The properties and neurobiological mechanisms of decisions are clearly a vast and active research field (Heekeren, Marrett, and Ungerleider, 2008).

While scanning a scene with eye movements, we need to decide when to move our eyes and what target to select for every single saccade. This is arguably the decision with which we are confronted the most often throughout our lives. Eye movements result in significant changes in signals to the brain, and they influence our conscious perception. The obvious way in which they influence our conscious perception is by changing the visual input, but also, more subtly, by the decision about what part of an environment to sample in the future (Kietzmann and König, 2015). A much-investigated example of this decision process can be found in the act of reading: To read this text, you continuously select the target of the next saccade. Extensive data and models have been published in this domain alone (Rayner, 1998; Rayner, TJ Smith, et al., 2009). The decision of where to look next is, of course, not

restricted to reading but occurs in all viewing behaviors, e.g., visual search (Najemnik and Geisler, 2005) or free viewing. Eye movements, i.e., the selection of the next fixation points, are evidently prime examples of decision processes.

Two primary decision processes occur in parallel: One decides when to look, the other where to look. A lot of work has been invested in understanding the selection of the next fixation location. Clearly, many different factors contribute to the decision of where to look next (Kollmorgen et al., 2010; König et al., 2016). Most notably, task-dependent factors (Buswell, 1935; Hayhoe and D Ballard, 2005; Rothkopf et al., 2007), stimulus dependencies (Einhäuser, Spain, et al., 2008; Foulsham and Underwood, 2008; Foulsham and Underwood, 2009; Koehler et al., 2014), and geometric dependencies of the trajectory (Motter and Belky, 1998; Hooge, Over, et al., 2005; Tatler, Baddeley, et al., 2006; Tatler and Vincent, 2009; Henderson and TJ Smith, 2009; Kaspar and König, 2011a) exist. For the future fixation location, some of these factors have been summarized in the saliency model (Koch and Ullman, 1985; Itti et al., 1998). In recent years, the performance of saliency models has slowly converged to the inter-individual noise ceiling. In other words, models based on features become as good as predictions based on other subjects, and therefore, only inter-individual differences remain to be explained (Wilming, Betz, et al., 2011; Kümmerer et al., 2015; Bylinskii et al., 2016). Furthermore, the concept of a saliency map is not just a computational convenience. Studies investigating neglect patients provide evidence for the existence of a saliency map in humans are more recent additions (Ossandón et al., 2012), presumably in the superior colliculus (White et al., 2017). Most models estimate saliency based on low-level stimulus properties, such as luminance, contrast, motion, or edges. In addition, high-level object-related features, such as the recall frequencies of objects in a scene, predict eye locations even better (Einhäuser, Spain, et al., 2008). In recent years, geometric

dependencies on the trajectory have increasingly been included in these models. Spatial bias (Tatler and Vincent, 2009), saccadic momentum (Posner and Y Cohen, 1980; Wilming, Harst, et al., 2013) and horizontal asymmetries (Ossandón et al., 2014) are incorporated to not only predict average fixation locations, but to model whole gaze paths (Schütt et al., 2017). Taken together, quite extensive literature and profound insights exist on the decision of where to look next.

The "when" question, i.e., the decision to initiate a saccade to a new target, determines the time available for the processing of the visual information available at the current fixation. Investigating this ongoing decision process solely based on the distribution of fixation durations is difficult, as selecting a new saccade target and extracting information from the current fixation are temporally overlapping processes (Findlay and Walker, 1999; Henderson and TJ Smith, 2009). Several clever paradigms have been established to dissociate these two factors. In gaze-contingent paradigms, such as scene onset delay or scene quality change paradigms, subjects explore an image, and at a critical fixation, the image is temporarily exchanged with a mask (Henderson and Pierce, 2008; Henderson and TJ Smith, 2009), or visually altered (Henderson, Olejarczyk, et al., 2014; Walshe and Nuthmann, 2014). These paradigms show that these changes influence the fixation duration immediately. They support the direct control theory (Gould, 1973; Rayner, 1978) of fixation durations. The direct control theory states that the processing difficulty should influence the fixation durations and thus the information content of the current fixation. A lower limit exists for how recent information can still influence the choice of what saccade to perform. The double-step paradigm (Becker and Jürgens, 1979) allows for the establishment of a minimal time of 80 ms when a saccade can still be reprogrammed (Findlay and LR Harris, 1984) depending on other factors for example task conditions (Walshe and Nuthmann, 2015). However, not all fixations

are under direct control. Henderson and Pierce (2008) found a different set of fixation durations where presumably the saccadic program could not be stopped. However, explaining this second set of fixation durations by a pure direct control theory might also be possible (Pannasch et al., 2011). These findings have been computationally modeled by using the CRISP model of fixation durations (Nuthmann, TJ Smith, et al., 2010). CRISP consists of three major components: A random timer that initiates saccade programs, a two-stage saccade programming step, and a modulation based on the current visual processing. Nuthmann et al. show that with this model, they can model a wide range of paradigms, including the mentioned scene onset delay paradigms (Nuthmann, TJ Smith, et al., 2010; Nuthmann and Henderson, 2012). For unrestricted eye movements, the literature is a bit sparser. The most comprehensive study comes from Nuthmann (2017), where she analyzed unrestricted eye movement data associated with multiple tasks for relevant factors for fixation duration. In another recent study, Einhäuser and Nuthmann (2016) analyzed the interactions of the "when" and "where" question in free viewing. In spite of this progress, the available literature for the "when" question is lacking compared with the literature for the "where" question concerning eye movements.

The ongoing foveal processing during a fixation is an example of an exploitation process. In contrast, the initiation of saccades and thereby the inspection of the environment exemplifies an exploration process (Gameiro et al., 2017). These together establish a dilemma, quite similar to the exploration-exploitation dilemma commonly observed in other disciplines (Daw et al., 2006; JD Cohen et al., 2007; Berger-Tal et al., 2014): It is a dilemma because at a given point in time, we can either exploit or explore, but not both. We are in a continuous decision process between exploring the image and exploiting the current view. It follows that the balance of these two processes determines whether we maintain the

current gaze location or initiate a new saccade and consequently the distribution of fixation durations. Compared with the traditional exploration-exploitation dilemma, knowledge gain continues throughout fixation. However, the formulation of eye movements as an exploration-exploitation dilemma places the focus on the ongoing decision process. The distribution of fixation durations is the main observable outcome of this dilemma.

In this paper, we investigate the exploration - exploitation dilemma in a guided viewing paradigm. During free viewing (as in unrestricted viewing), experimenters usually do not have direct control over fixation durations. They can usually bias the distribution but not control it. Of course, causal interpretations require controlled experimental interventions. Therefore, the ideal experimental conditions would allow us to control these factors by discretizing the temporal and spatial aspects of eye viewing behavior. Such a paradigm should allow one to precisely modulate all parameters of interest, i.e., fixation duration, the number of the possible next fixation targets, the saliency at the current and next locations, and the geometric features of multiple saccades. In this study, we made the first step toward such a paradigm. We used guided viewing, where subjects saw only a small aperture of a scene (a bubble) at a time at each fixation (Gosselin and Schyns, 2001; Kollmorgen et al., 2010). Following this, we exchanged the current bubble by up to five different bubbles at other parts of the underlying stimulus. The subjects then selected one of these bubbles via a saccade, and the unselected bubbles were removed. The participants explored the image by repeating this process. At the end of the trial, we used a memory task as a distractor task, but also to compare the task performance to trials where the bubbles are shown concurrently and not sequentially. The main benefit of this new paradigm is that it allowed us to disentangle the processing phase from the time to select a new fixation location.

We were interested in three predictions that the paradigm allowed us to test. First, it follows from the exploration-exploitation dilemma that after sufficient exploitation and analysis of the current view, the system will be ready to continue exploring the environment. Exploitation will finish, or saturate eventually, because information content of a single bubble is limited. If this process is stopped early, resources needed to continue exploration will still be used for exploitation and the readiness to perform a saccade should be low. Our paradigm allowed us to probe the dilemma: The display time of the fixated bubble was experimentally controlled. With a short display time, the visual system is still in the exploitation phase, and subsequent eye movements should be generated more slowly. Vice versa, with a long display time, the visual system is ready to, literally, move on, and saccades should be elicited more quickly.

Second, we tested whether saccade planning can be naively thought of as multiple evidence accumulators that race independently to a fixed threshold. The first to reach the threshold decided the time and place of the saccade. Multiple potential targets, and thus multiple accumulators, should result in a shorter delay to elicit a saccade. In our paradigm, we tested this by introducing an additional experimental manipulation: Instead of using the classical guided viewing paradigm, which operates with a single future target location, we allowed subjects to decide between multiple locations.

Third, we investigated how far these processes depend on the actual information being analyzed. We, therefore, used noise and urban images with low and respectively high information content. A reasonable assumption is that urban images contain more information that needs to be exploited. Thus choice times should generally be longer.

Summarizing, we hypothesized that longer forced fixation times on the current bubble lead to a shorter time to elicit a saccade

to the next target. Furthermore, we expected a decrease in the time needed to elicit a saccade with the number of possible future targets. Finally, we hypothesized that the choice time is positively correlated with the processing demand at the current fixation location.

5.4 Methods: A new gaze-dependent paradigm to control the where and when of eye movements

5.4.1 Subjects

In the primary study, 35 subjects participated (18-42 years, mean age: 24 years, eight male, three left handed, 14 left dominant eye). In a second experiment (an internal replication and a context experiment), we recorded 10 additional subjects (18-22 years, mean age: 19.5 years, one male, one left handed, one left dominant eye).

In the first experiment, we excluded five subjects. Four sessions had to stop early on the subjects' request; one other subject did not reliably look at the bubbles. In the second experiment, three subjects were excluded from further analysis. One subject stopped early, and two other subjects had excessive errors in eye tracker calibration and drift corrections. Each subject provided written informed consent. The ethics committee of Osnabrück University approved the experiment.

5.4.2 Apparatus and Recording

We used a 24" LCD monitor (Benq XL2420T) with a screen resolution of 1920x1080 pixels and a refresh rate of 120 Hz for presentation purposes. The participants viewed the screen from a distance of 80 cm. The participant's left eye movements were tracked at 500 Hz by using an EyeLink II system (SR Research Ltd., Mississauga, Ontario, Canada). We used a 13-point calibration with a mean validation error of $<0.5^\circ$ and a maximal validation

error of $<1^\circ$. Furthermore, we performed a drift correction before each trial.

5.4.3 Procedure

In every trial, the subjects explored a given image using small apertures, inspired by the bubble technique (Gosselin and Schyns, 2001): Subjects performed a drift correction and an additional subsequent fixation on a cross-shaped fixation point for 300-700 ms. Then, a trial followed with the display of one random single bubble (a sub-trial) visible for an experimentally specified time (forced fixation time) (Figure 5.1A). In this first sub-trial, we always used a single bubble. After the forced fixation time, the initial single bubble display was replaced with one to five new bubbles (one bubble in 51.6%, two in 25.8%, three in 12.9%, four in 6.5%, and five in 3.2% of fixations). This staggering was used to ensure a higher signal-to-noise ratio for the single bubbles, allowing a more certain estimation of the main effect of forced fixation time at cost of estimating the number of bubbles effect or possible interactions thereof. Out of a pool of precomputed bubbles, we randomly selected a set of new bubbles (choice bubbles), with a linear bias toward bubbles close to the current one. Thus, this procedure gave preference to bubbles that were closer to the currently displayed bubble. The subjects chose one of the new bubbles by performing a saccade toward it. The time needed to initiate the new saccade was termed the choice time and was our primary dependent variable. We used the phrase choice time instead of saccadic reaction time, as several bubbles offered a choice for the subject to make. Thus, the choice time included the time to select the new target as well as the time to initiate a saccade. During this choice time period, the current bubble was switched off; only the future bubbles were visible. After the saccade to the new target, this new target was displayed for the next forced fixation time. Concurrently with the

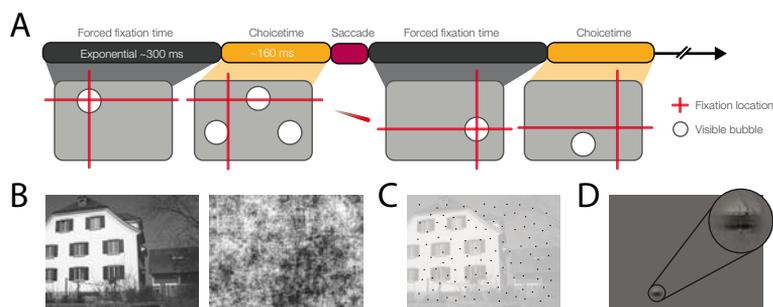


Fig. 5.1 A) Experimental Procedure. Subjects foveate a single bubble for the forced fixation time. It is sampled from an exponential distribution with a mean of 300 ms to ensure a flat hazard function. The next step removes the bubble and simultaneously displays one to five new bubbles (in this example three). The dependent variable is the time until a saccade is initiated. After the selection of one bubble, and most often before the saccade ends, the other bubbles are removed. The newly foveated stimulus is shown for a newly sampled forced fixation time. B) Example stimuli urban and pink noise. C) The sampling points for bubbles following the pseudo-random Poisson disc sampling algorithm. D) A single bubble as seen by the subject.

saccade detection, the other, non-chosen bubbles were removed from display. As subjects could in principle saccade to a distant bubble passing over bubbles in between, we used an algorithm based on the last three samples to estimate the saccade velocity. After the detection of a saccade, defined by a velocity of greater than $30^\circ/\text{s}$, we checked for the closest bubble only after the velocity was lower than $50^\circ/\text{s}$ again. These parameters resulted in the best time-delay/ spatial accuracy tradeoff. In the analyzed data, we observed that the other choice bubbles were removed in the online algorithm after on average 11 ms (90% range of -4 ms to 23 ms) after the end of the saccades as the eye tracker detected offline. Please note that minimizing this time was no easy feat, as the eye tracker had a delay of 5 ms to deliver the sample. The monitor took an additional 5 ms to switch to 95% target luminance, and with 120 Hz, we had (on average) 4 ms (+-4 ms) for the vertical retrace to start again. We think that an average delay of 11 ms to turn off the non-chosen peripheral stimuli is negligible. In any

case, a slower reaction time had an influence only on when the other bubbles were turned off. The selected bubble was constantly displayed until its forced fixation time expired. For the analysis, we used the forced fixation time starting at the end of the saccade as the eye tracker detected offline.

The distribution of forced fixation times followed an exponential function with a mean of 295 ms, which resulted in a constant hazard function. Thus, at any point in time of the trial, the probability that a new bubble would appear was held constant. Consequently, subjects could not anticipate when the currently fixated bubble would disappear and a saccade was necessary.

The subjects saw parts of the image at two times. One was the forced fixation time, when the time subjects foveated a single bubble. The second was the choice time after the forced fixation time when the subjects saw only stimuli in the periphery, and the foveated bubble was removed. The first, the summed-up display time over all fixated bubbles, was on average 6.1 s (95% quantile: [6.0s, 6.2s]) and was similar to the free viewing conditions used in other experiments. Thus, depending on the forced fixation times in one trial, subjects were presented on average 19.3 bubbles (95% quantile: [11, 26], SD of 4.5). The second part, i.e., when stimuli were visible only in the periphery (the summed-up choice times), took on average 3.2 s (95% quantile: [2.7s, 3.9s]). Taking the first part together with the second, we obtained an average of 9.3 s (95% quantile: [8.9s, 10.0s]) during which subjects saw the content of the stimuli.

5.4.4 Stimuli

Each subject completed 128 trials in total. 96 trials consisted of the guided viewing bubble paradigm; the other 32 trials belonged

to a static image condition with a different research question in mind. In this static condition bubbles were not shown subsequently. They were shown simultaneously instead and subjects were allowed to explore the bubbled-stimulus at their own pace (these trials are discussed further in the results section under "2 AFC Memory-Task"). If not explicitly stated otherwise, we only discuss data of the 96 trials with the guided viewing bubble paradigm. In half of all trials, we used grayscale urban images (1280*960px). These were photographs of Zürich city and surrounding cities and had been used in several previous studies (Wilmington, Betz, et al., 2011). In the other half, we used grayscale pink noise images with the same luminance distribution as the urban images (SHINE-Toolbox, Willenbockel et al., 2010). The bubbles (Figure 5.1 B-D) were Gaussian patches with a diameter of 3° , acting as apertures on the underlying image (Kollmorgen et al., 2010). Within each bubble, a fovea filter (Loschky et al., 2005) and a Gaussian filter were applied to imitate the visual acuity of the human eye and smooth the transition to the background. The fovea filter ensured that all information of a bubble could be extracted with a single fixation at the center. That is, there was no information gained by exploring a bubble by repeated saccades to different locations within it. A Poisson disk sampling algorithm placed bubbles pseudo-randomly on each image with a minimum distance constraint (Bridson, 2007). The minimum distance of two bubble centers was set to the bubble radius. Thus, the maximal area overlap of two bubbles was smaller than 15%. In contrast to a random placement from a uniform distribution, this algorithm avoided clusters or holes during sampling (Figure 5.1C). Successive bubbles were enforced to be without any overlap. We used on average 101 bubble locations (range 91-109) on each image.

Subjects were instructed to look at the center of the bubbles. The distribution of the distance of fixations to the bubble center (Figure 5.2A) shows that subjects fixated closely to the center. The

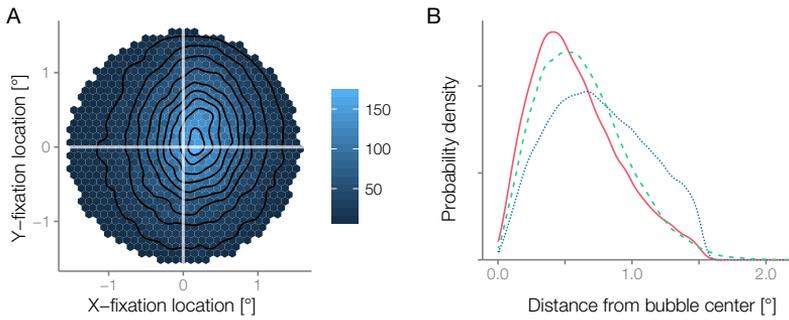


Fig. 5.2 A) 2D density plot of the number of fixations. The center is close to the origin. B) Kernel density of distance to bubble center. Blue/Dotted: first fixation, red/solid: first re-fixation, green/dashed: theoretic distribution based on the Gaussian bubble mask

dashed green line (Figure 5.2B) shows the density of distances when sampling from the Gaussian visibility mask. The distance to zero reflects that if one draws two samples (x and y) from a Gaussian, it is unlikely that both of them will be close to zero. Refixations (red) lead to a distribution that supersedes this density (Figure 5.2B). From that, we conclude that subjects targeted the center of the bubbles and performed corrective saccades (see also Kollmorgen et al., 2010 for similar results).

5.4.5 2-AFC memory Task

After each trial, the participants performed a 2-AFC memory task. We presented two bubbles simultaneously next to each other. One was chosen randomly from the pool of possible bubbles in the previously displayed image but was not necessarily shown during the trial. The other bubble was chosen randomly from any another image. The participants indicated via a button press which of the two bubbles could have been part of the image they saw. There was no time limit. The aim of the task was twofold: to ensure that subjects processed the image content in some depth and to use it as a motivation and distractor task.

5.4.6 Data Analysis

Data Preprocessing

We based our analysis on the premise that subjects fixated on the displayed bubbles ($n = 67454$). Thus, we removed all sub-trials where they did not fixate on the bubble, where the saccade detection algorithm failed, or where the calibration was not good enough to directly detect which bubble was fixated on. We excluded sub-trials, with a fixation outside of the currently displayed bubble during the forced fixation time ($n=11390$, 16% of total) and where there was no direct saccade to the next bubble but an intermediate one in-between bubbles ($n=22613$, 33% of total). Further, we excluded sub-trials, where the planned forced fixation time was different from the observed forced fixation time by an arbitrary threshold of more than 40ms ($n=6147$, 9% of total). Such a discrepancy could arise when the online saccade-detection algorithm detected the saccade earlier or later than the more sophisticated offline saccade detection. We used the observed forced fixation time in all analyses, which is the actual time in which the subjects saw the stimulus. In total, we removed 30185 sub-trials. Thus, all in all, 55.3% sub-trials remained (mean per subject 1062.3, range 806-1283). Sub-trials with more than one fixation inside the currently displayed bubble (25.2% of the remaining bubbles) were kept in the analysis. A reanalysis of choice times of sub-trials with only a single fixation did not change the results.

In order to remove extensive outliers that would strongly influence our analysis, we used an outlier detection algorithm (Leys et al., 2013) based on three times the median deviation (MAD) distance from the median for each subject (with a constant factor of 1.4826 included; thus, it defaults back to σ in the normal case). This procedure removed an additional 5.6% of the sub-

trials and is below the 10-15% recommended criterion (Ratcliff, 1993).

Linear Mixed Model

Fixation durations and, in our case, choice times are dependent on a multitude of factors. These can be categorical (e.g., the type of scene, urban against noise), or in our study most often continuous (e.g., the amplitude of the previous saccade, the angle between saccades, or, in our case, the forced fixation time). Either a repeated measures ANOVA or a linear mixed model could account for the repeated measures of subjects. However, a repeated measures ANOVA does not allow for continuous factors. Thus, we choose a Bayesian linear mixed model with uniform priors on all parameters.

The first set of predictors describes our main experimental manipulations. Two of the predictors are of primary interest in this study: The forced fixation time and the number of bubbles. The interaction between these two factors indicates a possible dependency between the underlying processes. In addition, we model the categorical stimulus type, urban images, against noise stimuli.

The second set of effects is correlative in nature and relates to the spatial relation of the fixated bubble to the bubble fixated on in the previous and next sub-trial. As we did not experimentally influence saccade trajectory, for example, by enforcing a fixed trajectory of bubbles, these spatial effects are correlative predictors and do not allow conclusions on causal relations. The angle in absolute monitor coordinates is a circular predictor. In order to model this dependency, we used a Fourier decomposition with one and two periods for the 360° of the predictors (thus, four predictors

in total, including two sines and two cosines). We modeled this effect once from the previous to the current bubble and once from the current to the next bubble. Also, we used the difference of the angles between the previous and the current bubble and the angle between the current and the next bubble. We also included the distances of the bubbles. To account for a central spatial bias, we included the standardized z-transformed x- and y-position and the squared, z-transformed x- and y-position of the bubbles in order to account for a quadratic decay toward the edges of the image, which is usually observed with spatial biases. Also, we model a term to describe the distance of a bubble to the center of the screen.

The third set of predictors incorporated temporal dependencies and sequential influences between the trials and bubble presentations. We used trial number, previous forced fixation time, and previous choice time as predictors.

Bayesian Model Fit

We analyzed the data using hierarchical logistic mixed effects models fitted by the No-U-Turn-Sampler in STAN (Homan and Gelman, 2014; B Carpenter et al., 2017). For the model specification, we followed the implementation by Sorensen (Sorensen et al., 2016). In maximum likelihood linear modeling terms, all within-subject effects were modeled using random slopes clustered in subject and a random-intercept for subject. We estimate all covariances between the random effects with an LKJ-Prior ($\nu=2$), which slightly emphasizes the diagonal over the off-diagonal. We used treatment coding for all categorical factors and interpret the coefficients accordingly.

We used six MCMC-chains with 1000 iterations each, with 50% of the iterations used for the warmup period. We visually con-

firmed convergence through autocorrelation functions and trace plots. Furthermore, we calculated the scale reduction factors (Gelman, Carlin, et al., 2013) and ensured the recommended criterion for convergence ($Rhat < 1.1$). To control for an adequate model fit, we calculated 1000 posterior predictive draws and plotted the median and 95th percentile together with the raw data. The posterior predictions matched the data well, and the model seems to be adequate for our inferences. When displaying the data and posterior predictions in their quantile ranges, the posterior predictive checks showed that our model did not capture all features of the data but merely the central tendency. We find a mismatch in the 2.5% quantile and 97.5% quantile. There, the raw data have higher choice times than the posterior predictive. This results from the skewed distribution of choice times on the subject level (Figure 5.3). But, most importantly, the posterior predictive median and mean value fit the data appropriately.

Reported Statistics

For Bayesian posterior predictive checks, we report 95% credible intervals (CI) using the Cousineau correction for grouped data (Cousineau, 2005). For posterior parameter estimates, we report 95% posterior CI. For other reported data, we use 95% bias-corrected and accelerated bootstrapped confidence intervals of the mean with Cousineau correction for grouped data where applicable (Cousineau, 2005). Cousineau correction adjusts confidence intervals of repeated or grouped estimates by first subtracting the total average of each subject before calculating the variance.

Software

Experimental software was written in PyGame (<http://pygame.org>). We processed data in MatLab (The MathWorks, Inc., Natick, Massachusetts, United States), Python (<http://python.org>), and R (R Core Team, 2013). We acknowledge the use of other packages used to analyze the data (Wickham, 2007; Wickham, 2015; Bates, 2010; Walt et al., 2011; Carr et al., 2015; Pinheiro et al., 2016). All experimental scripts, analysis scripts, and data are publically available under <http://osf.io/ba2pn>.

5.5 Results: Longer exploitation time and less available options result in shorter decision time

In this study, we use a new paradigm to break down the free viewing paradigm into controlled subprocesses. While several experimental paradigms exist to experimentally manipulate fixation durations, here, we directly control them using forced fixation periods. Furthermore, we control the participants' saccade trajectories with the use of small bubble-like stimuli that act as an aperture on the underlying image. These bubble stimuli allow us to guide the participants in their exploration.

Thirty-five subjects viewed 96 images each (plus 32 images intermixed with other trials, see Results 2AFC-Task for a description), and we analyze 35105 fixations based on the guided viewing paradigm. The mean of the median choice time, that is, the duration until subjects initiated the next saccade, was 156.7 ms (95 percentile range: [123.1 ms, 210.3 ms], Figure 5.3). As expected for fixation durations and reaction times, our choice time distributions are skewed toward the right tail. We used a Bayesian linear mixed model to explain the variations of the choice time. The units of the linear model parameters were chosen to be intuitive and to facilitate comparison between the effect sizes. We either used the average values of the predictors (246 ms based on the average fixation duration in parts of a big free viewing data set (Wilming, Onat, et al., 2017)) or maximal values (the total number of trials, 96 from this experiment plus 32 intermixed ones, resulting in 128 trials and 19 bubbles per trial average). We additionally grouped the predictors into three distinct classes: Experimental factors, correlative factors, and sequential effects. Experimental factors are controlled, randomized, and balanced. Correlative factors were

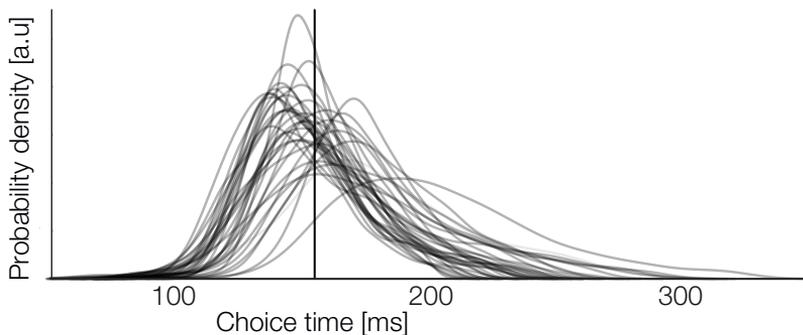


Fig. 5.3 Kernel density estimates of choice times for each subject. The horizontal line depicts the mean of the subjectwise-median choice times.

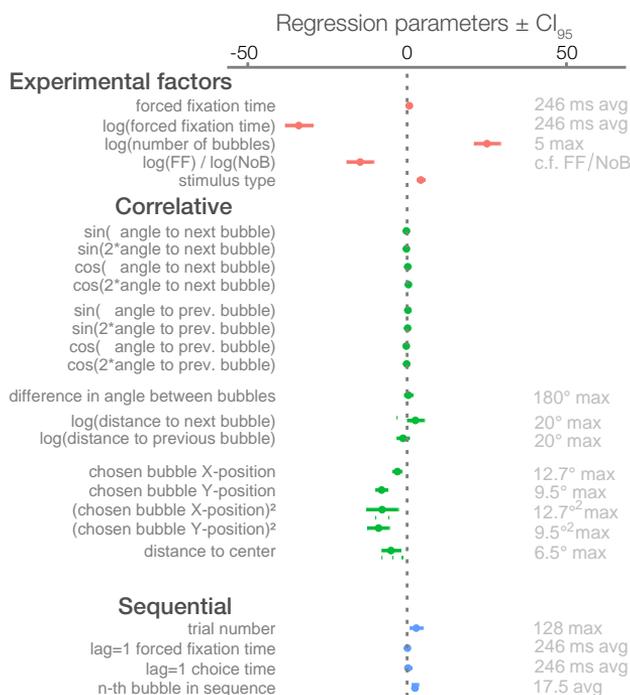


Fig. 5.4 Linear mixed model parameter estimates and Bayesian 95% posterior credibility intervals. We state the parameters in intuitive units (third column) to aid interpretation. A 'max' depicts the highest observed value of this factors, and 'avg' the average of independent data (Wilming, Onat, et al., 2017).

analyzed based on the subjects' behavior. The sequential effects capture possible influences due to previous trials. We present the posterior marginal density results in Figure 5.4.

5.5.1 Forced Fixation Time

Our first prediction relates to the exploration-exploitation dilemma: Sampling the decision process early after fixation onset should result in longer reaction times than eliciting a new saccade after prolonged initial fixation. All subjects show an exponential decline in the choice time (Figure 5.5AB). For short forced fixation times below 100 ms, we observe choice times of around 175 ms. In contrast, in the case of long forced fixation times, choice times saturate around 150 ms. This compressive nonlinearity has an instantaneous rate of growth of -34.0 [$-38.3, -29.6$] (95% credibility interval, see Methods) times the forced fixation time normalized by an average fixation duration in free viewing images of 246 ms. We find an additional linear effect of forced fixation time with a slope of 0.7 [$-0.1, 1.5$]. This effect is small (a change of 0.7 ms in 246 ms, or a 3ms change in 1000 ms) and the parameter estimate contains 0. Therefore, we do not discuss the linear part of the forced fixation time effect further. Descriptively, we observe an increase in the choice times for forced fixation times larger than 1s (not shown). Due to the exponentially distributed forced fixation times, on average, only 2.4% of the forced fixation times are larger than 1s. Therefore, we do not feel confident in making estimates for values larger than 1s, which are unlikely to influence the model estimates. From the mixed model results, it is clear that the forced fixation time increases the reaction time: the longer the current fixation lasted the faster subjects responded.

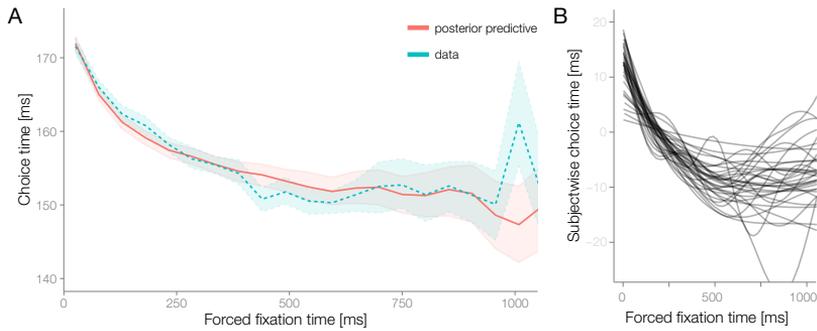


Fig. 5.5 Main effect of forced fixation time. Subjects viewed the bubbles for an experimentally controlled time, depicted on the abscissa. We record the time until they choose to perform a saccade to the next bubble on the ordinate. A) blue: data with bootstrapped 95% confidence interval of the mean using Cousineau correction. Red: posterior predictive distribution with 95% credibility interval. B) Single subject kernel density estimates of the data.

5.5.2 Number of Bubbles

The second experimental manipulation is the number of bubbles. After fixating a stimulus for the forced fixation time, the stimulus disappeared and one to five bubbles emerged in the periphery. The subjects then decided on one and performed a saccade onto it. The other bubbles then disappeared before fixation onset. Here, we observed a monotonic logarithmical increase of the choice time from one to five bubbles (Figure 5.6), with an effect size of 25.1 ms [21.0, 29.2] over five bubbles. While the logarithmical effect captured the data well, a categorical or piecewise-linear component (e.g. one bubble is unique, 2 to 5 bubbles are linear) seemed to capture the data equally well. Theoretic considerations for a possibly higher number of bubbles in future experiments make the logarithm the most reasonable choice with the best generalization. It is evident from the data that there is a monotonic increase in the choice time: the more bubbles that are available for a decision, the longer it takes for one to choose a bubble.

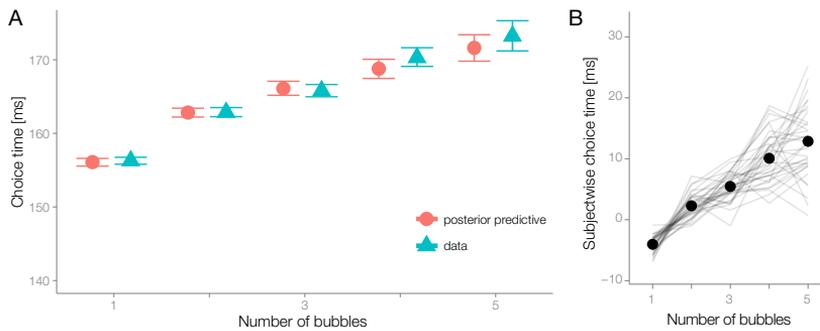


Fig. 5.6 We varied the number of bubbles from which subjects could choose. A) Main effect number of bubbles, blue: data with bootstrapped 95% confidence interval of the mean using Cousineau correction, red: posterior predictive distribution with 95% credibility interval. B) Distributions for individual subjects. Each subject showed an increase over the number of bubbles.

5.5.3 Interaction of Forced Fixation Time and Number of Bubbles

Next, we analyze the interaction between the forced fixation time and the number of bubbles (Figure 5.7) on the choice time. Here we find an interaction between the number of bubbles and the forced fixation time of -14.7 $[-19.0, -10.5]$. In Figure 5.7, we smoothed the display to get a more intuitive understanding of this relationship. An example to understand the effect size for this log-log interaction is helpful. The predicted difference between a forced fixation time of 100 ms and 1000 ms for a single bubble is 14.7 ms; for five bubbles, it is 20.0 ms. Thus, in this case, the interaction boosts the single bubble forced fixation effect by 40%. We acknowledge the small absolute effect size. Nevertheless, the credibility interval for this effect is far removed from 0. The interaction shows that the effects of the forced fixation and the number of bubbles on the choice time are dependent on each other: The more bubbles are available to choose from, the larger the effect of the current forced fixation time.

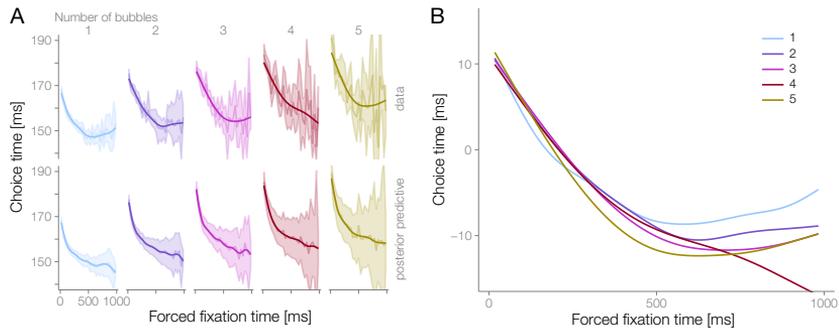


Fig. 5.7 A) The main effects of the number of bubbles (columns) and the forced fixation time (abscissa). The upper row depicts raw data with 95% bootstrapped mean with Cousineau correction and the lower row depicts the posterior predictive with 95% Cousineau credibility intervals. B) We show the interaction between the forced fixation time and the number of bubbles. The data from A, lower row, were further smoothed for this plot (generalized additive model with thin plate regression splines). And the mean of each curve was removed. This eliminates the main effect of the number of bubbles. A sharper decline results in trials with many bubbles to choose from compared to trials with a low number of bubbles.

5.5.4 Image Category

The last experimental effect is the image category. The bubbles of one trial were drawn from urban images in half of the trials and pink-noise images in the other half of the trials. We observed a main effect of 4.3 ms [3.1 ms, 5.5 ms], which means urban bubbles have a longer choice time than noise bubbles. This main effect is surprisingly small, given the vast differences in image statistics between the two categories.

5.5.5 Geometric Effects

We observed various correlative effects (Figure 5.8). These effects were not experimentally controlled or balanced and thus are correlative in nature. The first effect is the angle between the currently fixated bubble and the next bubble (Figure 5.8A, inset). We modeled this effect using four parameters describing

the sine and cosine of the base and two times the base frequency. This procedure allowed us to capture the dynamics of the circular nature of this effect (Figure 5.8A). The shape of the marginal effect display follows a smooth sinusoidal curve: Reaction times of saccades toward the upper part of the screen seem to be faster than toward the lower part of the display. But if we look at the model estimate, the maximal average difference (the amplitude), and thus the effect sizes of this summed curve, is only 0.9 ms [0.4 ms, 1.5 ms]. The mismatch between the marginal display and the parameter effect size is due to other predictors explaining a large share of the variance of the effect. We observed a similarly small effect of the angle between the previously fixated bubble and the currently fixated bubble on the decision process at the currently fixated bubble. Here the maximal difference over angles is 0.7 ms [0.3 ms, 1.1 ms]. This effect follows a double-u shape, with increased choice times for vertical saccades and faster choice times for horizontal saccades.

The difference between the two angles is shown in Figure 5.8B and is commonly referred to as saccadic momentum. In line with other research (Wilming, Harst, et al., 2013), return saccades appear slower than forward saccades, but in our study, this is visible only in the marginal plot. Taking into account other factors, the effect disappears. Whereas in free viewing Wilming et al observed effect sizes of around 45 ms, here we did not see a reliable effect, with on average 0.4 ms [-0.9 ms, 1.7 ms] for the maximal difference of 0° to 180° (see Discussion). The logarithmic distance to the next bubble (Figure 5.8 C) predicted the choice time with a maximal effect size of 2.6 ms [0.0 ms, 5.3 ms] between a theoretical saccade of 0.2° and the maximum of 30°. Similarly, it is unlikely that there exists an effect of the distance to the last bubble, with an estimate of -1.3 ms [-3.3 ms, 0.9 ms] for the maximal range. Next, we discuss the predictors for the absolute position of fixation. Without these predictors, the aforementioned nonsignificant effects all were

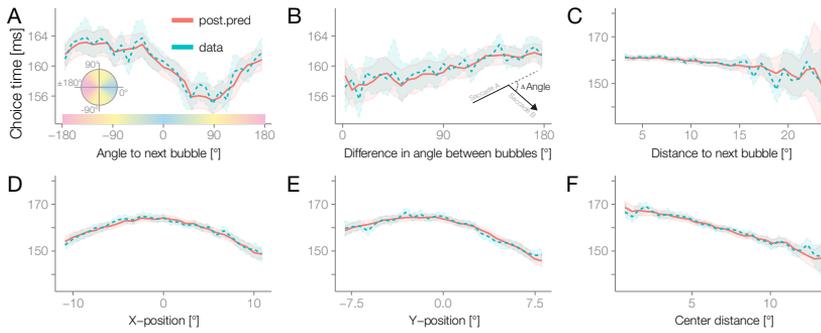


Fig. 5.8 A) Angle to the next bubble. B) Difference in angle between two bubbles, over termed saccadic momentum. C) Distance to the next bubble. D, E) Horizontal and vertical position of the bubble that has been selected (not the current foveated one), the distance of the selected bubble to the center of the image. C-F) Degrees depict visual angle.

significant (although small, model results not shown here). We modeled the absolute position of the chosen (thus future) bubble using a linear and a quadratic term to capture the symmetric and quadratic nature of this spatial bias. The horizontal position had a linear influence of -3 ms [-4.6 ms, -1.5 ms] for the maximal range. The vertical position had a bigger effect of -8 ms [-10.0 ms, -5.9 ms] for the maximal range (which in absolute measures is smaller than the horizontal position). We also observed quadratic effects for the horizontal position of -7.8 ms [-12.8 ms, -2.9 ms] and for the vertical position of -9.0 ms [-12.5 ms, -5.7 ms]. In addition to the absolute position, we also observed a linear distance-to-center effect of -5.1 ms [-8.0 ms, -1.8 ms]. These effects indicate a quicker reaction time the farther away from the center one's eyes rest.

5.5.6 Sequential Effects

The last effects we modeled are sequential effects: trial number, bubble number, previous choice time, and previous forced fixation time. The trial number influenced the choice time by 2.9 ms [0.9 ms, 4.8 ms] throughout the 128 trials of the experiment

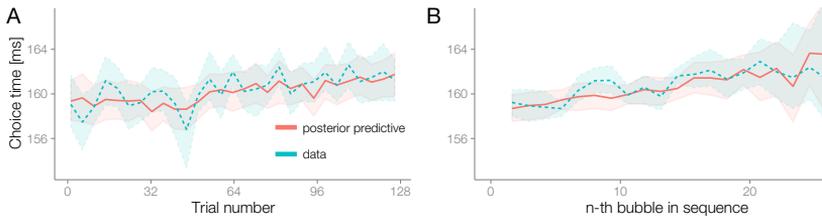


Fig. 5.9 A) Trial number. In total, 128 trials were used, but 32 intermixed trials were only related to the memory task. B) On average, 19 bubbles were shown in a single trial; this effect depicts a very small increase of reaction time during one trial. This concludes the correlative effects. Compared to the experimental effects, the effect sizes are very small.

(Figure 5.9A). The bubble number describes the n -th bubble a subject saw in each trial. It had, on average, an influence of 2.5 ms [1.5 ms, 3.5 ms] (Figure 5.9B) throughout one image, i.e., on average over 19 bubbles. The previous choice time had an influence of -1.9 ms [-3.0 ms, -0.8 ms] for an average fixation time of 246 ms. The previous forced fixation time did not seem to influence (0.01 ms [-0.2 ms, 0.3 ms]) the current choice time. All these effects are considered small.

5.5.7 2-AFC Memory Task

Subjects performed a 2-AFC task after completing the sequence of bubbles based on one image. They viewed one bubble taken from the previously shown image (but not necessarily a bubble shown during the previous trial) and another from a different image. The most informative difference happens when two urban bubbles need to be differentiated. Here we can see whether there is a difference in performance between this sequential task and the additionally recorded static condition. For two noise stimuli, performance was 52.7% (CI_{95} 49.7% - 56.0%, SD: 9.4%), as to be expected at chance level. For trials with one noise bubble and one urban bubble, performance was close to perfect, with 97.4% (CI_{95} 95.8% - 98.2%, SD: 3.3%). For trials wherein two urban bubbles

were compared in the memory task, subjects showed an average performance of 72.8% (CI_{95} 69.0% - 76.0%, SD: 10.6%).

This static image condition was recorded in addition to the regular trials interleaved in the same session. In this condition, a bubbled version of the image was visible, which consisted of 21 non-overlapping bubbles. The number was slightly higher than the on average 19 bubbles in the current study. The image was visible for 6 s, a similar time to the total viewing time in the regular trials. 16 noise and 16 urban images were shown intermixed with the sequential trials. Here we see very similar results: For two noise stimuli, performance was 53.6% (CI_{95} 47.9% - 60.6%, SD: 18.4%). For trials with one noise bubble and one urban bubble, performance was at 96.5% (CI_{95} 93.8% - 98.1%, SD: 6.2%). For trials with two urban images, performance was at 75.6% (CI_{95} 70.7% - 80.6%, SD: 15.1%). The confidence intervals for the sequential and static trials are largely overlapping. It is reasonable to conclude that performances in the two tasks do not differ and changing a static information uptake to a dynamic one does not introduce strong artefacts in the processing of information.

Second Experiment: Internal Replication and Context In light of the recent replication crisis in the (cognitive) psychological sciences subjects (Pashler and E Wagenmakers, 2012; Aarts et al., 2015), we strived to internally replicate our findings on an independent set of subjects. In addition, we were interested in what way context shown at the beginning of a trial influenced the choice times and integration performance. We performed a second experiment ($n = 10$) where we briefly flashed the entire scene for 92 ms (11 displayed frames at 120 Hz) at the beginning of half of the trials. Similar to the first experiment, subjects performed a drift correction before fixating on a fixation cross for 300-700ms. After, the whole scene was flashed for 92ms and then the experiment started with the first bubble (always a single bubble) at a location around the fixation spot (we used the same algorithm as in the first experiment, see Methods:

Procedure). This allowed subjects to extract the gist of the scene (gist condition) but did not allow making any saccades. It has been shown that at around 100 ms, subjects already extract the main features of a scene (Potter, 1976), even when the image was masked, which we did not do here. In the other half of the trials, the experiment remained identical to the original experiment to allow for within-subject comparisons. The two types of trials were randomly intermixed.

As can be seen in Figure 5.10, we replicated all effects we found in the first experiment. In addition, we did not observe an interaction with the gist of a scene in any of the factors. In the 2-AFC memory task, subjects performed at very similar levels: The case of two noise stimuli was not different from chance level (50.3%, CI_{boot95} : 44.5% - 59.3%). One noise and one urban bubble resulted in near perfect performance (98.3%, CI_{boot95} : 97.1% - 99.2%). And we found the same results as in the previous experiment. When subjects had to choose between two urban bubbles, there was no improvement in either gist condition (no-gist: 78.8%, CI_{boot95} : 70.4% - 84.2%, and with-gist 70.5%, CI_{boot95} : 55% - 79.0%). Nevertheless, the second experiment replicates all of our choice time results in a new cohort of subjects. Furthermore, briefly presenting the gist of a scene did not influence either the choice times or performance in the memory task.

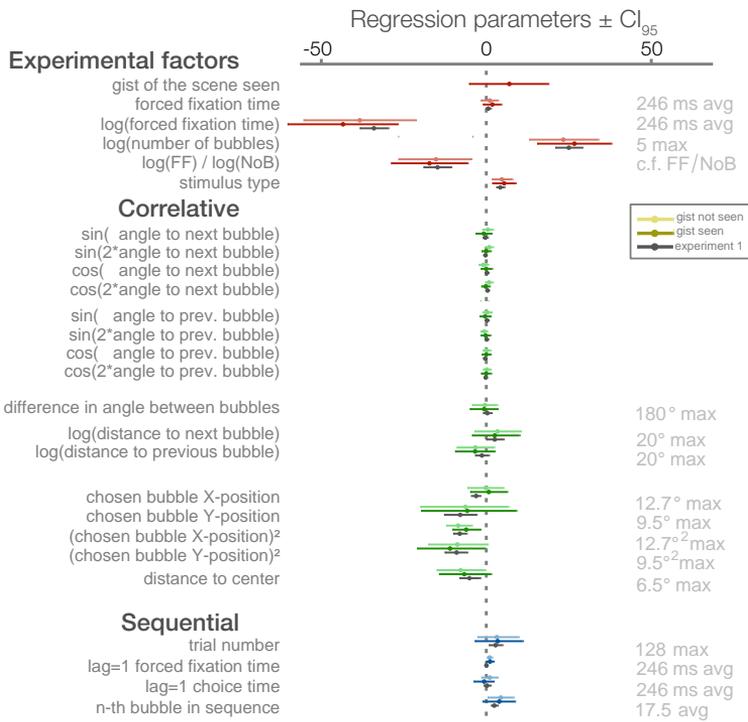


Fig. 5.10 The upmost interval of each posterior predictive triplet depicts parameter estimates when subjects did not see the gist of a scene (bright color). The middle interval (dark color) shows parameter estimates when the gist was shown for 92 ms before exploration started ($n = 10$, paired). The lowest interval (gray) depicts the parameter estimate shown in Figure 5.4 (different subjects, $n = 35$), wherein the gist was not seen either. Showing the gist does not induce any relevant difference in the within-subjects comparison or in comparison with the cohort of the previous experiment.

5.6 Discussion

5.6.1 Summary

The distribution of fixation durations can be described by an exploration-exploitation decision process. Here, we used a guided viewing paradigm to control the decision processes occurring during a fixation and dissociated it from the processing of the fixated location. We found an exponential decrease of the time needed to choose the next fixation target dependent on the time available for processing the stimulus at the current fixation location. This dependence provides evidence for the exploration-exploitation dilemma in the decision process. Secondly, we found a monotonic increase in choice time with the number of available saccade locations. These data indicate that potential future saccade targets are accumulating evidence in their favor in a dependent manner.

5.6.2 Exploration-Exploitation

Our first main result directly tested the exploration - exploitation idea: We described eye movements as an ongoing decision process between further exploitation of the current view and further exploration of new, unseen elements. In this study, we interrupted the subjects during the exploitation stage of the current bubble at unpredictable points in time and investigated the influence on the choice time. From the literature, we expected saccadic planning times to be between 100 and 175 ms (Rayner, 1998; JD Schall and KG Thompson, 1999), which is in line with the present observation of 150 ms. Please note that the choice time was measured from stimulus offset. It was not measured from the previous saccade offset, as at that time, no targets for further saccades were available. Thus, the intersaccade time amounts to an average of 450

ms. In principle, the choice time (starting with stimulus offset and target onset) could be constant and independent of the forced fixation time, ending with stimulus offset and target onset. Instead, we observed an exponential relationship of the choice time and the forced fixation time. This demonstrates that the choice time is dependent on the degree of processing at the current fixation location and gives support for the exploration-exploitation view.

5.6.3 Are There Alternative Explanations for the Exponential Decay?

In simple reaction-time experiments, Drazin (1961) observed an exponential decay when the foreperiod, the period between warning signal and go signal, was increased. This mimics our forced fixation effect and could be an alternative explanation. As discussed in Niemi and Näätänen (1981) and more recently using saccades (Oswal et al., 2007), the important factor influencing this effect is the predictability of the stimulus. The hazard function, the instantaneous probability that a foreperiod/forced fixation is ending, is essential here. With a nonuniform hazard function, as in Drazin, subjects can predict with higher certainty when a stimulus is going to appear, whereas when a nonpredictable, flat hazard rate is used (Mowbray, 1964; Baumeister and Joubert, 1969; Oswald et al., 2007), no effect of foreperiod on the resulting reaction time can be found. We used a uniform, thus unpredictable, hazard function. Therefore we think it is unlikely that our effect can be explained by predictability effects of reaction time. There are two limitations when comparing our study to previous studies. First, the foreperiods in previous studies are usually much longer (in the range of seconds). Second, all previous studies had a fixed, constant foreperiod. Our constant foreperiod, the saccade duration, is effectively very small. We do not think that these limitations can explain the exponential decay. A different potential confound is

based on refixation strategies. It is possible that the instruction to look at the center of the bubble initiated a refixation program at each fixation. This could result in prolonged choice times for short forced fixation times. This prolongation should be linear with a slope of -1 ms choice time per 1 ms forced fixation time. Because we observe an exponential decay in our data, we think that this confound cannot capture our results.

5.6.4 Related Paradigms

A paradigm that has a very similar procedure to the current study is the mask-onset delay paradigm (Rayner, TJ Smith, et al., 2009; Glaholt et al., 2012). Their goal was to measure the minimal time viewers need to see a scene at each fixation in order to obtain normal viewing behavior. In order to test that, Glaholt et al. masked the stimulus after a given time (50 ms, 75 ms, and 100 ms). This stops any further incoming information. Either the whole stimulus (Rayner, TJ Smith, et al., 2009; Glaholt et al., 2012) or a circular part around the current fixation was masked (Glaholt et al., 2012). In addition, Glaholt et al. (2012) also controlled for the sudden onset of the mask which could elicit saccadic inhibition. They measured the time until subjects started the next fixation. They found that in the full-scramble condition, subjects elicited similar saccades and reached the same performance as in free viewing only at 100ms viewing time in either of their tasks. Our results indicate that even more time is needed (at least 400 ms) until the effect of forced fixation time saturates for all fixations. We speculate that the information content of the momentary view and the task (Glaholt et al., 2012) are crucial factors here. For example, in an reading experiment, Rayner et al (Rayner, W Inhoff, et al., 1981) conclude that only 50-60 ms of visible stimulus is needed for fixation behavior to be indistinguishable from unrestricted information access. Another important factor may be local vs

global masking. In Glaholt et al. (2012), scrambling only the local mask reduced this effect. Here, 50ms was enough to extract the information to solve either task. It seems possible, though, that this task could have been solved without foveal information and only using peripheral information alone as proposed by Nuthmann (2014). Glaholt et al. (2012) also found a bimodal distribution of fixation durations, which they explain by a saccadic inhibition mechanisms due to the rapid onset of the stimulus mask. Saccadic inhibition (Reingold and Stampe, 2002) is a delay of saccade production that the onset of a (possibly irrelevant) stimulus has. Its most salient feature is a bimodal distribution in the fixation durations; shortly after the inhibitory stimulus fewer saccades are generated. In our case we did not observe this bimodality in the choice times (Figure 5.3) or when adding the forced fixation times to the choice times (not shown). Thus, saccadic inhibition cannot explain our results.

Another phenomenon related to the one reported here could be seen in the stimulus onset delay paradigm (Vaughan and Graefe, 1977; Vaughan, 1982). In this paradigm, subjects searched for a target at two fixation points. At each fixation there was a variable stimulus onset delay, replacing the fixation point either with a target or with a distractor. The time to elicit the next saccade can be seen mirrored in our choice time. Similarly, the stimulus onset delay is mirrored in the forced fixation time. If the stimulus was shown with a delay of 300 ms after onset, Vaughan found decreasing response times by 100-150 ms compared to immediate display. Thus, the delay speeds up the response time to the appearing stimulus. Vaughan (1982) discusses a possible explanation based on predictability, similarly to the foreperiod effect discussed above. It is likely that the crucial difference between the two tasks is that in our paradigm the stimulus is directly visible at fixation onset, then subsequently vanishes and the new target appears. Whereas in the stimulus onset task, the stimulus is not visible upon

fixation onset but becomes visible after the forced non-stimulation period. We conclude that due to the very different effect sizes (30ms in our study against 150ms), the difference in task and the unpredictable foreperiod in our study, it seems unlikely that the effect observed due to the stimulus onset delay is related to the exploration-exploitation effect we investigate in the present study.

5.6.5 Evidence Accumulation With Multiple Targets

Our second prediction relates to the evidence accumulation between multiple targets. We assumed one independent evidence accumulator for each target that race to a fixed threshold. The first to reach the threshold is selected. Our observed monotonic increase with the number of future target locations is incompatible with this explanation. This result is compatible with data and evidence accumulator models from Leach and R Carpenter (2001). To reconcile these data with such an evidence accumulator model, interactions between the integrators of different locations are necessary. This might occur through different thresholds for multiple future locations as can be found in monkey LIP neurons that decreased drift rates through lateral inhibition due to limitations in sampling capacities (R Carpenter and Williams, 1995; Leach and R Carpenter, 2001; Churchland et al., 2008) or more complex interactions with time-varying thresholds (Ludwig, 2009). A future analysis step with this paradigm is to fit drift-diffusion models taking into account the effects described here.

Van den Berg introduced two processes that are decisive for the duration of a fixation, one starting a new saccade and one staying at the current fixation. This is made explicit in visual search (Beintema et al., 2005; Berg and Loon, 2005). In their

model, both processes are explicitly modeled as two dependent integrators racing to individual thresholds that decide whether to continue exploiting the current view or go on and explore the scene. In visual search experiments, it seems as though the processing of the peripheral stimulus is secondary to the processing difficulty at fixation (Hooge and Erkelens, 1999; Wu and Kowler, 2013). The conclusion from these studies is that in visual search subjects do not bias fixation durations for better target selection, but only for discrimination of the target at hand. This is not to say that the peripheral information is ignored; it is still used for target selection, it just shows a weak influence on the fixation durations. It seems that the subsequent integration of foveal and peripheral information can occur independently (Ludwig et al., 2014), but not always so (VanRullen et al., 2004; Berg and Loon, 2005). Contrary to our initial prediction, we found only a very small effect of foveal processing in our experiment: a 4 ms difference between urban and pink noise images. On the other hand, the peripheral decision task showed a difference of up to 25 ms depending on the number of bubbles. Thus, in our experiment, it seems we have a reverse relation to observations in visual search (Hooge and Erkelens, 1999). Of course the task and general structure of the experiment are quite different. In Hooge et al., the target is available for as long as the subjects prefer, whereas in our study we forcefully interrupt the information extraction process. These results show how strong task dependency can influence the interaction and integration of peripheral and foveal information.

We observed a logarithmic increase between the number of possible targets and reaction times, which is similar to the "Hick's"-effect (Hick, 1952; Proctor and Schneider, 2017). But there is also evidence that saccades do not follow this rule: the "anti-Hicks" effect (Lawrence et al., 2008; Lawrence, 2010). This anti-Hicks effect is commonly observed when there are multiple possible target locations to choose from. These authors differentiate between

exogenous (e.g., prosaccades) and endogenous (e.g., antisaccades), where an anti-Hicks effect can be found in the former but not the later. We cannot replicate this finding here. It is well possible that selecting between multiple possible targets is an intrinsic task and our results are in agreement with earlier findings. We think there are two additional important differences: One is that subjects in our study could not predict how many bubbles would appear compared to a blocked design in Lawrence. A second difference is that in our study there is no correct stimulus, whereas in Lawrence there was always one correct target. We could reconcile this by a new experiment with a modified version of our paradigm presenting multiple bubbles but highlighting one as a target. In this case, an anti-Hicks effect could be observed.

A functional explanation can be seen by the "cost of a saccade" (see discussion in De Vries et al., 2016) Based on their ideas in the discussion section, a saccade has two consequences in our paradigm. For the duration of the saccade, saccadic inhibition suppresses information processing. In addition, in the case of multiple bubbles, a saccade also removes peripheral information that could have been further exploited, would the pre-saccadic fixation duration (the choice time) have been prolonged. Thus, longer choice times for trials with many bubbles would allow for longer exploitation of the peripheral content. It cannot be disentangled by this study alone whether longer choice times for more targets are based on a more difficult decision process or a more detailed peripheral information extraction of the stimuli, or both.

From our finding of a monotonic increase followed a new prediction for free-viewing paradigms. In free viewing, some parts of an image are preferably gazed upon. This can be quantified by empirical saliency, usually measured by the density maps of fixations over multiple observers. It has been recently shown

that locations with higher saliency have higher fixation durations (Einhäuser and Nuthmann, 2016). Given our findings and what has been found about the interaction of saliency and fixation durations, we predicted prolonged fixation durations if an image has multiple distinct peaks of empirical saliency against other locations without these distinct peaks. The peaks in saliency mimic multiple bubbles as shown in our study.

5.6.6 Why are there no stimulus content dependencies?

Contrary to our prediction, we found only a small effect of bubble content on the choice time. We think that this effect is due to chance and does not reflect a real effect. We imagine two explanations of why the effect could not be observed. It is possible, that there is truly no dependency between the information extraction process and the forced fixation effect. The effect then reflects a content-independent property of the system, in a sense a hardwired, probabilistic solution to the exploration/exploitation dilemma. Alternatively the time required to analyze pink noise and urban images may be comparable. This would be in line with high fixation durations during free viewing of pink noise stimuli (Kaspar and König, 2011b). Similar rates should lead to similar choice time distributions. Further comparisons with other stimuli, or modifications of bubble-saliency, could be used to understand this better. Computational Models of Fixation Duration Models of fixation durations were proposed in the reading literature (Reichle et al., 1998; Engbert et al., 2005). In recent years, fixation duration models for free viewing and visual search have emerged as well; for example, the CRISP (Nuthmann, TJ Smith, et al., 2010) and ICAT (Trukenbrod and Engbert, 2014) models. In these two models, fixation durations are modeled by two main components, a stochastic random walk for the timing of saccades and one for the

saccade generation. For the saccade generation part the authors used either a random draw from a gamma distribution (CRISP) or another stochastic random walk (ICAT). Our data could be used to test specific model assumptions made by CRISP and ICAT, in particular the ones relating to saccade generation.

Instead of truly generative models of fixation durations, as in the models described above, a simplified model using drift-diffusion modeling could be used. The LATER model (R Carpenter and Williams, 1995) is a popular race-to-threshold model for (saccadic) reaction times (Noorani and RH Carpenter, 2016) and has recently been applied to explain both fixation durations and locations (Tatler, Brockmole, et al., 2017) in scene viewing. Similar to the previous, more advanced models, LATER does not only model the average choice times, but the whole distribution. Differences between experimental manipulations could be caused, or be hidden in the shape of the response distributions (for a Bayesian hierarchical solution based on Weibull-functions see Rouder et al., 2005) and not captured by the central tendency measurement. The discrete notion of exploitation (fixation on a stimulus) and exploration (which stimulus to select) can help differentiate between drift rates of saccade initiation and saccade generation.

5.6.7 Comparison of Other Factors to Free Viewing

Many dependencies of fixation duration are described in the literature and consequently modeled in our study. Comparing our results to the literature, it is obvious that our observed effects are drastically smaller, and sometimes not existent. One example is the absolute angle of a saccade: saccades from the center to above the horizon are generally observed to be faster than saccades below it (for an overview see Greene et al., 2014). For example Heywood

and Churcher (1980)) reported an effect of 31 ms and Tzelepi et al. (2005) reported 27 ms. Our effect is estimated to be around 0.8 ms and thus does not exist in practical terms. Another example is the amplitude of the previous saccade, which has been described in a search paradigm by (Salthouse and Ellis, 1980) with 90 ms for 15°. In our study, we found only 2.6 ms. A third example is an angle to the previous fixation, termed saccadic momentum or inhibition of return (Dorris et al., 1999; AJ Anderson et al., 2008; Wilming, Harst, et al., 2013), which has been found to influence fixation duration with a piecewise linear relationship by approximately 70 ms per 180° (Wilming, Harst, et al., 2013). In our study we did not find evidence for such an effect. Note that some effects do increase in size when not taking into account shared variance with other predictors. This is evident from the mismatch between effect size and marginal data as plotted in the figures. Still, even interpreting the marginal effects, the same qualitative judgment of smaller effects than in free viewing persists.

Some of the relationships described in the literature seem to be task dependent. For example, in a recent study (Nuthmann, 2017) a saccadic momentum effect could only be found for memorization and aesthetic preference but only very weak during a search task. This could partially explain the mismatch of our effect sizes. Yet another factor is that in our paradigm, peripheral information is restricted and the controlled fixation duration might have a strong influence on the oculomotor system to show the same effects that can be seen in unrestricted viewing (Nuthmann, 2014). We propose a third, nonexclusive explanation: in our study, we separate the processing of the current input from the planning and execution of the next saccade. Also, our observed choice times are quite fast. In free viewing, these two processes occur in parallel. Thus, saccadic planning can start quite early in the current fixation. Consequently, in free viewing more time is available for other processes to intervene and modify either the duration of the

exploitation phase or the planning of the next saccade. To test whether the exploitation phase is modified in free viewing by these geometric factors, we propose a modification to our experiment: instead of hiding the currently exploited bubble and stopping the exploitation process, we ought to keep it together with the new target bubbles. Thus, exploitation can continue and is partially under the subjects control, and we expected to see effect sizes of similar size found in free viewing.

5.6.8 Comparison to Nuthmann (2017)

In a recent study by Nuthmann (2017), a complementary approach to this study was used. She analyzed fixation durations from free viewing with three different tasks using linear mixed models. Of course, the paradigm differs in many ways. Nuthmann used unrestricted free viewing with different tasks, but we used the combination of the bubble and guided-viewing paradigm, a restricted viewing task. Nuthmann focused on the influence of stimulus features on the fixation durations, we merely analyzed the difference between background images based on pink noise against urban images as our focus laid on the direct control of fixation durations. Both studies analyzed oculomotor effects but have a slightly different set of predictors. In this study, we additionally included the absolute angle and the absolute position of the fixation. In general, Nuthmann found very strong oculomotor effects over all tasks, with some exceptions. For example, Nuthmann observed strong saccadic momentum in two tasks, but much weaker momentum in the visual search task. After controlling for other oculomotor effects, the initially observed distance-to-center effect of Nuthmann vanished. This was not the case in our study. We found an effect of distance to center even though modeling very similar other oculomotor predictors. Building upon the work of Nuthmann on feature influences, our paradigm could be used

to directly check her image feature based findings on a causal level: one can modify the low-level features of certain bubbles. For example, in a new experiment one can experimentally modify the "clutteredness" of certain bubbles and then expect a modified choice time.

5.6.9 Alternative Statistical Models, and Model Critique

We analyzed the choice times using a Bayesian linear mixed model. Linear mixed models (Bates, Mächler, et al., 2014) and other hierarchical models (Gelman and Hill, 2007) are steadily replacing the need to use traditional ANOVA/ANCOVA (Bagiella et al., 2000; Quené and Van Den Bergh, 2004; Richter, 2006; Baayen and Milin, 2010; Kliegl et al., 2010; Nuthmann, 2017). On the other hand, statistical recommendations for applied mixed models are still being developed (Barr et al., 2013; Bates, Kliegl, et al., 2015) and sometimes puzzling results can be observed (Hodges, 2014). One common problem is that maximum likelihood estimates fail to converge, mostly due to the complexity of the covariance matrix between random effects (Barr et al., 2013; Bates, Kliegl, et al., 2015). Due to the high number of predictors our model is prone to this problem. Therefore, we used a Bayesian version of the mixed model (Sorensen et al., 2016) that allows us to put a small prior on the covariance matrix that de-emphasizes correlations between random slopes. This allowed us to fit the complex model and interpret the maximal instead of a reduced (possibly parsimonious) mixed model. In the case sufficient evidence from the data for a nonzero correlation exists, this will overrule this weak prior. This contrasts with forcing correlations to be zero as commonly done in more parsimonious models. It is currently up for debate whether it is possible to use those parsimonious models while still preserving correct type-1 error rates (Bates, Kliegl, et al., 2015; Matuschek

et al., 2015). Often only Bayesian mixed models allow estimating all random slopes in the mixed model to achieve a full statistical coverage when the model is sufficiently complex. Without all random slopes, posteriors, standard errors, or confidence intervals being too narrow, resulting p values are too small, type-1 errors are too high, and thus the effect estimates are too liberal (Schielzeth and Forstmeier, 2009; Barr et al., 2013; Sorensen et al., 2016). After the model fit, we used posterior predictive model checks to see whether the model adequately captures the data. As expected from a Gaussian fit of skewed data, the model checks revealed that we adequately fitted the mean of the distributions, but failed to model the tails appropriately. A possible enhancement for future analysis is to model the choice times using skewed normal, lognormal, or mixed normal distribution to reflect the most extreme quantiles.

5.6.10 The Bubbled Guided-Viewing Paradigm as a Research Tool

We introduced a new paradigm to investigate fixation durations in free viewing. There are some benefits over current free viewing paradigms, but also shortcomings. One problem with this paradigm is that we do not have a satisfying explanation of why the geometrical correlative effects are so much smaller than in free viewing. This could be a potential problem in studying any geometric effect with this paradigm. It might be that using a less naïve forced fixation duration distribution could alleviate this problem. Second, due to strict criteria, we had to remove $\approx 45\%$ of the data. The biggest factor was intermediate fixations between two bubbles. Further analysis is needed to determine whether better online fixation detection algorithms or higher accuracy can improve upon this. On the other hand, the benefits are quite clear: we have a direct window to the decision-making mechanisms during free viewing. It controls the time a stimulus is available, how many and where

the next stimuli are displayed, and thus, gives valuable variance estimates to constrain computational models. The paradigm is easily expandable to other spatial sampling mechanisms. In this case, we used a Gaussian distance measure, but it is straightforward to relate the geometrics of exploration to real free-viewing scan paths of subjects recorded beforehand. An alternative could be the use of empirical saliency maps that mean sampling the points as bubbles that are most likely fixated by other subjects. In addition, it is not clear how the number-of-bubbles effect generalizes to more bubbles, but there are two likely candidates: an asymptotic behavior or an inverse cubic function. Summarizing the benefits, we can state that we have a new tool to experimentally probe the exploration-exploitation state of the system that allows for close control and is flexible in its sampling scheme.

5.6.11 Conclusions

Here we developed a new paradigm that allows for experimental control over fixation durations and exploration behavior. We observed that we could selectively interrupt exploitation behavior and confirm predictions from the exploration-exploitation idea. In addition, we show a monotonic increase of choice time with the number of future targets.

Bayesian Generative Models and Optogenetics: Modeling the Kinetics of Melanopsin

” *I think the model is dispensable.*

— Reviewer #1

Contributions

bioRxiv: Ehinger, Eickelbeck, Spoida, Herlitze, and König 2016
BVE and PK conceived the model. **BVE** performed the analyses. KS and DE recorded the data. **BVE** and PK wrote the manuscript.

Current Biology: Spoida, Eickelbeck, Karapinar, Eckardt, Jancke, Ehinger, König, Dalkara, Herlitze, and Masseck 2016
KS, DE, OAM, DJ, **BVE**, and SH conceived experiments. KS, DE, TE, RK, and OAM performed electrophysiological experiments and cell-based assays. DD and MDM produced AAVs. KS, DE, SH, OAM, MDM, and PK wrote the manuscript.

6.1 Layman Summary

Two things are very useful in order to understand a complex system: 1) Direct control over all of its parts, and 2) a non-black-box model of the system explaining and predicting its future behavior. Both things are difficult to obtain for the brain. It is very hard to directly control which brain areas or neurons are active and it is also difficult to develop cognitive models of behavior. In this chapter, I will approach this issue by developing analysis methods based on a smaller and less complex neuronal system than the whole brain at once.

To tackle the first problem of direct control, optogenetic methods have been heavily developed in recent years. Optogenetic proteins allow direct activation and silencing of target neurons through colored light. In this chapter we study one of these protein, Melanopsin.

To deal with the second problem, the non-black-box representation, we developed a model of Melanopsin's functioning, implemented in the framework of Bayesian Modeling. This allows us to estimate the unknown variables that influence how Melanopsin works directly from the observed data from cellular recordings. I used Bayesian Models in the previous chapters as well, but merely as a statistical approximation. In this chapter I will use them as a model able to predict and generate new data. With this approach, I am able to simulate the behavior of different Melanopsin types under completely new conditions. I will show, in a tutorial-type of way, how one can build, fit and enhance such a Bayesian model.

6.2 Understanding melanopsin using Bayesian generative models - an introduction

Understanding biological processes implies a quantitative description. In recent years a new tool set, Bayesian hierarchical modeling, has seen rapid development. We use these methods to model kinetics of a specific protein in a neuroscientific context: melanopsin. Melanopsin is a photoactive protein in retinal ganglion cells. Due to its photoactivity and signaling kinetics, melanopsin has recently become attractive as an optogenetic tool and an important component in the elucidation of neuronal interactions. Thus it is important to understand the relevant processes and develop mechanistic models. Here, with a focus on methodological aspects, we develop, implement, fit and discuss Bayesian generative models of melanopsin signaling dynamics. We start with a sketch of a basic model and then translate it into formal probabilistic language. As melanopsin occurs in at least two states, a resting and a firing state, a basic model is defined by a non-stationary two state hidden Markov process. Subsequently we add complexities in the form of (1) a hierarchical extension to fit multiple cells; (2) a wavelength dependency, to investigate the response at different colors of light stimulation; (3) an additional third state to investigate whether melanopsin is bi- or tri-stable; (4) differences between different sub-types of melanopsin as found in different species. This application of modeling melanopsin signaling dynamics demonstrates several benefits of Bayesian methods. They directly model uncertainty of parameters, are flexible in the distributions and relations of parameters in the modeling, and allow including prior knowledge, for example parameter values based on biochemical data.

6.3 Introduction: Why should we develop Bayesian generative models of Melanopsin?

Time-varying data can be analyzed with a multitude of statistical methods. Integrating ordinary or partial differential equations is one of the major tools in the natural sciences. For example in order to analyze the morphology of an action potential we could model the rise and fall by a system of two coupled differential equations. In a linear approximation this results in two exponential functions, where the time-constants of the exponential describe the rise and fall. Alternatively we could use the more complex Hodgkin-Huxley model (Hodgkin and Huxley, 1952). This system of equations does not only better describe the data, but allows a direct interpretation of model variables in terms of molecular and cellular properties. Furthermore, in many experiments, multiple factors influence the dependent variable concurrently and the process of interest is non-stationary. In that case, extracting single time constants can be biased and unable to explain the data. And consequently the mechanistic model should be preferred. The benefit of such generative models is the ability to generate 'fake-data' using previously fitted parameters. It allows to predict unseen data and simulate experiments where, for example, some of the parameters were changed. Thus, the first step to analyze time-varying data, is to develop a formal mechanistic model of your data.

Once we specified the model, we need to estimate the values for the parameters based on measured data. A solution to such systems of differential equations is most commonly in the form of maximum likelihood estimates, i.e. the one parameter set so that the occurrence of the data as observed is most likely. While often used, another approach has important benefits and improvements:

Bayesian parameter estimation. It allows us to directly estimate parameter uncertainties, interpret them intuitively as probabilities about parameters conditioned on the data and we are able to seamlessly include prior knowledge. Due to these benefits, Bayesian parameter estimation has seen a strong comeback and is becoming ever so popular (Cronin et al., 2010; Ghasemi et al., 2011).

In order to use Bayesian estimation we need to understand three concepts: the likelihood, the prior and the posterior. The likelihood tells us how likely it is, that our data are generated by a given set of parameter-values. The prior tells us, how likely certain parameter-values are in the first place. Thus if we a-priori know that a receptor has a certain time-constant from previous experiments, we can directly incorporate this knowledge in our current model-fit and adequately influence the posterior of the time-constant parameter and all other co-dependent estimates. The posterior of each parameter is the distribution that shows us how probable each parameter-value is, given our data and prior knowledge, thus a combination of prior and likelihood. In the end we do not only get a single best-fitting parameter value, but a distribution. Thus in addition to the most probable parameter value, we estimate the uncertainty of the parameters, the probability distribution. A broad probability distribution indicates that we cannot estimate the parameter well: neighboring parameter values have a similarly high posterior probability. But a thin distribution indicates that the parameter can be estimated with high precision. Furthermore, dependencies between several parameters might be complex, but can be modelled by these methods. With Bayesian methods we can flexibly use generative models and, importantly, the posterior probability can be interpreted as uncertainty of a parameter, a straight forward and often implicitly used interpretation.

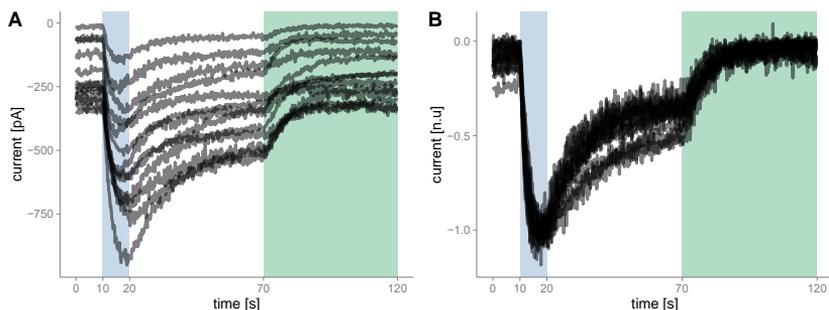


Fig. 6.1 A) Raw data of hOpn4L patch clamp recordings. hOpn4L was expressed in HEK 293 cells which express GIRK1/2 subunits. The GIRK-mediated K^+ -currents were sampled at 50-200Hz. Blue light (470nm) activates current influx, green/yellow light (560nm) deactivates the influx. B) Data were resampled to 5 Hz. We then normalized the range by mapping the 95% percentile of each cell between 0 and -1.

As an example to guide this paper we use patch clamp recordings of cells expressing melanopsin, a photosensitive opsin-type occurring naturally in the retina. In mammals it is expressed in intrinsically photosensitive retinal ganglion cells (ipRGCs) which project to the suprachiasmatic nucleus and influence the circadian rhythm (Hankins et al., 2008; Do and Yau, 2010). A melanopsin expressing ipRGC will increase firing frequency if photons of a certain wavelength activate the protein. Melanopsin is activated using blue light (470 nm) and can subsequently be deactivated using green-yellow light (560 nm). Melanopsin variants differ in their activation/deactivation kinetics. Mouse melanopsins' (mOpn4L) activation is sustained, once activated it stays activated for several seconds to minutes, whereas human melanopsin (hOpn4L) shows only transient activation upon light-stimulation (Spoida et al., 2016). Melanopsin presumably occurs in two states, the M (active) and R (resting, inactive) states (for a review see (Schmidt and Kofuji, 2009), but see (Emanuel and Do, 2015)). Activating the protein with blue light increases the probability of the R-state to change its configuration to the active M state. Concurrently, a constant transition-probability from R to M and M to R exists that leads the cell to an equilibrium distribution of melanopsin

in M and R state configurations. Here, we use data from patch clamp recordings in human embryonic kidney cells (HEK) stably expressing GIRK 1/2 subunits (Spoida et al., 2016), where GIRK channels are activated by human and mouse melanopsin using blue light and subsequently deactivated with green-yellow light (Figure 1).

In this paper we develop a Bayesian mechanistic model of melanopsin and discuss the implementation of the model, the inverse fit, model checks, the interpretation of the parameters and how we can exchange parts of the model in a modular way to improve our understanding and design new experiments.

6.4 Method and Results: From Model Building, Parameter Estimation and Diagnostics

6.4.1 Model Building

It is helpful to start with a graphical model representation (Figure 2 A). In this paper we loosely follow the model notation in (MD Lee and EJ Wagenmakers, 2014). Once the graphical model is specified, it can be directly implemented into a Bayesian programming language. In the graphical model (Figure 2 A) all parameters that change over time are shown inside the time point box and indicated with time-indices. The main parameters are the proportion of firing (M) and resting (R) states. In every simulation time step t_i there is a certain probability to switch states from M to R: $p(MtoR)_t$. This transition probability is influenced by a constant rate C_{MR} and a green-light dependent rate L_{MR} . Of course the light dependent rate is only taken into account, when there is green light, thus we need a dummy-coded green light variable L_G with 0 when there is no light, and 1 when the green light is active. Because light-activation happens at specific time points determined by the experimenter, the transition probabilities change over time, i.e. they are non-stationary. The transitions are implemented using ordinary differential equations. One of the assumptions of the model is, that the recorded patch-clamp currents are directly proportional to the proportion of M-state. We don't expect the patch clamp noise level to change during our recording time, and thus we include a constant Gaussian noise term into our model. To summarize: We model the patch clamp currents using a Gaussian where the mean is proportional to the amount of M-state and thus non-stationary over time. The model allows us to intuitively grasp the parameters, interactions and

mechanisms that are needed to model our data. A more formal way to describe this implementation is to describe the model as a non-stationary two-state hidden Markov model. We then estimate the transition probabilities and relevant factors. All scripts and models are documented and publicly available under <http://osf.io/bn6pk>. In this paper we make use of the STAN packages (B Carpenter et al., 2017), in combination with R (R Core Team, 2013). The non-stationarity in our case was implemented by a logistic linear model with time-varying predictors. The model code is shown in Figure 2 B, parallel to the model graph. In the following, square brackets reflect arrays, round brackets reflect functions. In our case, the linear model can be described by:

$$p_{RtoM}[t] = \text{logit}(C_{RM} + L_B[t] \cdot L_{RM})$$

Where c_{RM} is the constant change parameter, $L_B[t]$ defines at which time intervals blue light is active and L_{RM} is the blue light dependent change parameter. The logit function maps values from the domain $-\infty$ to ∞ to the domain of 0 to 1, thus in the domain of probabilities. This formulation as a logistic linear model allows us to connect the estimation of parameters over multiple cells with the idea of hierarchical or mixed models (see section Modular Improvements, hierarchical fit further down). The same formula defines the spontaneous transition probability from M- to R-state. Thus for the size of change of R-state at each point in time, there exist two influences: Some fraction of melanopsin changing their state from M to R and in the same time step some spontaneous change from resting state to active state. The combined probability determines the proportion of R (or M respectively) as captured by using ordinary differential equations. At each simulated time step (with a predefined time-resolution Δt) we update our R-parameter (and M respectively) by a first order integration:

$$\Delta R = p_{MtoR}[t] * M[t - 1] - p_{RtoM}[t] * R[t - 1]$$

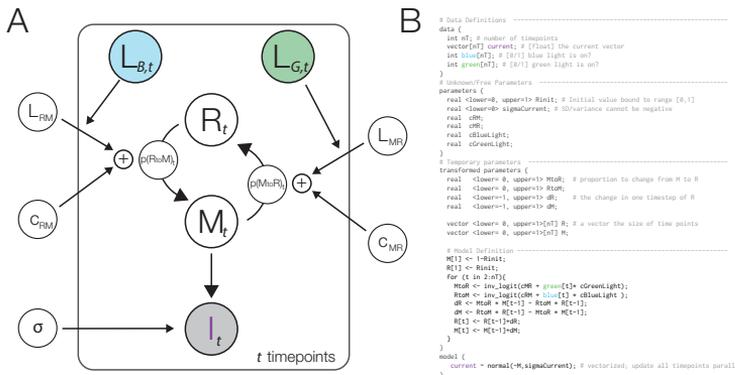


Fig. 6.2 A) A graphical model description of the basic model. Filled parameters depict data that is given. At each point in time a fraction of the R state is changed into the M state with the non-stationary probability $p(RtoM)$. The same process governs the change from M to R. The transition probabilities are influenced by constant (stationary) leakage probabilities and non-stationary, light dependent activations. The active M state is used as a model of the measured current of the patch clamp. These recording are inherently noisy, and we model this noise using a Gaussian function with the non-stationary mean M_t and the standard deviation σ . The parameter for the initial M/R state at $t=1$ was omitted from the graph. B) The graphical model implemented in the STAN programming language.

$$R[t] = R[t - 1] + \Delta R$$

We use a discrete time notation here to parallel the code of the implementation. In the two-state model it is necessary that the amount of M state is equivalent to the inverse of the R state. Thus:

$$M[t] = 1 - R[t]$$

We can make use of this relation and only calculate the change in R state and invert the change in the M state, but if we want to enhance the model to three states, it is more sensible to implement both changes, dR and dM.

The final important relationship to define is the relation to our data and including a noise distribution. In STAN this can be achieved by using:

$$I[t] \propto \text{normal}(M[t], \sigma_{\text{Current}})$$

In STAN the tilde (\propto) means "is sampled from". Thus, the line defines that the measured current I is sampled from a normal distribution with time-varying mean and constant variance σ^2 . Before the model fit we need additional statements about the type and range of parameters, we need to define the initial state, e.g. which could be random. This concludes the implementation of the specified graphical model into STAN.

6.4.2 Bayesian Parameter estimation

In the next step we estimate the posterior parameter distributions. Here we will give a short introduction of Bayesian data analysis and Monte-Carlo sampling methodology. Our goal is to estimate the posterior probability distribution: colloquially, what is the probability that each possible parameter value could underlie our data? According to the Bayesian framework, this consists of

firstly the likelihood of the data given the parameter. In other words how likely is it, that the data are generated from a specific set of parameters. Secondly, from the prior distribution, which states how probable a parameter is before using the data, that means from prior knowledge. In more formal terms, we are interested in the posterior distribution ($p(\Theta|D)$) given the likelihood of the data ($p(D|\Theta)$) and prior parameter probabilities ($p(\Theta)$). Bayes theorem states that these are directly related to each other ($P(\Theta D) \propto p(D|\Theta) * p(\Theta)$). An example: We record a neuron spiking with 10Hz. Our imaginative model assumes that the spiking rate of the cell is sampled from a normal distribution with a mean and a fixed standard deviation at 2 Hz. This model has only a single parameter to be estimated. We can easily calculate the likelihood of the Gaussian: We will get a low likelihood for a set of parameters where the mean is 5 Hz, a higher likelihood for a mean of 12 Hz and an even higher likelihood for a mean of 10 Hz. If we incorporate prior knowledge that these specific neuron types are very rarely observed with a spiking rate of higher than 5Hz, Bayes rule will integrate the information gained from the data and the prior-information and we will find the most likely parameter, given prior and data, at a lower estimate, for example 8 Hz. Whether data or prior dominates the posterior depends on how accurate, or certain, your prior knowledge was specified, and how much uncertainty, or noise, about the parameter the data has. Bayes rule automatically finds the optimal compromise between prior knowledge (what we think is a likely result) and our data (what actually happened).

Calculating the posterior is straight forward for a single parameter: We could randomly try out all parameter values using a grid approach, calculate the likelihoods and priors, and observe the posterior. This would be very ineffective, especially for if we have to estimate multiple parameters concurrently as there is a combinatory explosion. This is where the markov chain monte-

carlo (MCMC) sampling comes into play. Instead of randomly sampling the space, we start at a random initial value and propose to jump to a new value. We evaluate the posterior at this value, if it is higher (thus more likely) than the current value, we will go there. If it is lower, we will go there only with a probability inversely proportional to the difference. From there on we repeat the procedure for many iterations. This simple rule (known as the metropolis algorithm, (Hastings, 1970)) will ensure that we visit areas more often where the posterior is high, but from time to time explore other, less probable areas as well. Moreover our Markov chains fulfill all assumptions of the ergodicity theorem, thus it is guaranteed that the Markov chain will ultimately converge to the true posterior. In the end, our estimate of the posterior consists of how often we visited a certain parameter value. The MCMC sampling algorithm allows us to estimate highly complex models with many parameters.

Over the years more sophisticated algorithms have been developed. In this paper, we use NUTS, the No-U-Turn sampler, it is more efficient than the metropolis samplers in the case of hierarchical linear models with correlated parameters. This algorithm stems from the family of Hamiltonian monte-carlo (HMCs) algorithms. With HMC algorithms we replace the randomly chosen proposal step of metropolis with an algorithm that more effectively samples the posterior. Imagine that the inverse of the posterior has a bowl shape, thus the most likely points are at the valley, and the most unlikely one raise as mountains the further away you go. We randomly start at a point in the posterior and place a marble and send it with a small push in a random direction on its way. We now simulate for a while and the position the marble ends up, is our new proposed value. We compare it again to the current value and proceed as before. The marble has some momentum so it might just be enough to roll through local minima. The NUTS algorithm is based on HMC but in addition makes certain to not allow any

u-turns where the marble rolls uphill (due to gained or initial momentum) and would come down the same way again. The exact algorithm is somewhat more difficult because it needs to make certain that it converges towards the posterior but this is the general idea. For details we refer the interested reader to (Homan and Gelman, 2014). NUTS allows for an effective sampling of the posterior and reduces the risk to get stuck in local minima or passages where the chains could get stuck in the posterior landscape.

6.4.3 Model Fit & Sampling Diagnostic

Next we describe how STAN estimates the posterior distribution. Stan is a sophisticated open source implementation of HMC/NUTS for a multitude of programming languages (R, Matlab, Python, Julia, Stata and a command line tool). It allows to specify models in a comparatively simple way and has many tools to evaluate the results. The model comes with their own programming language which is not difficult to learn if experience in python, R, matlab or c++ is available. The STAN-model is then compiled to c++ code by the STAN interface and sampled by the MCMC algorithm. Sampling consists of two phases, the first is a warmup period where sampling-parameters are calibrated by the NUTS algorithm to effectively sample from the shape of the posterior. This is necessary as new proposed values could be outside the allowed range of the parameter and in that case we would have to reject this location proposal, thus we have an overhead of likelihood calculations. If this happens too often, we sample ineffectively. But at the same time, we do not want redundancies in the sampling resulting from small (but not rejected) step sizes. This tradeoff is automatically calibrated in the warmup period and the following sampling period defines the final outcome of our posterior.

The chains of an MCMC sampler need to be diagnosed for proper convergence. Sometimes we can get stuck in certain parameter value constellations, for example in a bimodal posterior distribution, or the MCMC algorithm makes too small jumps and we do not explore the space appropriately. It is difficult to diagnose those problems when we only look at a single chain with a single starting value. Therefore we use multiple chains which run independently. This allows us to check whether the chains converged in the same posterior distribution, which is necessary (but not sufficient) for successful sampling. There are several features that can indicate proper convergence: We visually inspect the chains (Figure 3 A), compare the variance between chains to the variance in one chain (termed $R\hat{H}at$, and should be close to 1), look at the overlap of the posterior densities of the chains (Figure 3 B) or we check the autocorrelation of a chain (Figure 3 C), how independent two following samples are from each other. A high independency is preferred here. In Stan this is often reported as a single number, the effective number of samples, N_{eff} , which is the number of samples corrected by the autocorrelogram ((Gelman, Carlin, et al., 2013) p. 286). In our first model, we ran 5 chains with 300 warmup iterations and 500 samples. Visually we see that the chains seem converged and the posterior overlap. Similarly the $R\hat{H}at$ is below 1.1 for all parameters. The autocorrelogram shows autocorrelation up to a certain degree, but it does not seem worrying (Figure 3 C, upper panel). In a similar vein, the effective samples are 700 for the upper and 1670 for the lower parameter, representing the 'best' and 'worst' effective sample value in this model. According to all our criterions, the chains of the MCMC seem to have converged.

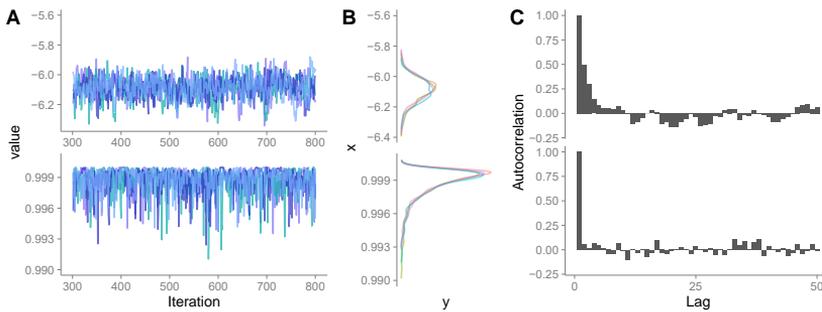


Fig. 6.3 A) four independent MCMC chains with 500 samples each of two parameters, RtoM and Rinit. The chains all converged to the same value range. The variance between chains is similar to the variance within chains. Visually, these chains seem to converge to the same value. B) The posterior density (marginals) of the chains in A. The chains all sample the same region of the posterior, this is an indication for convergence of the chains. These densities can further be simplified by specifying for example the medians and 95% quantiles of the distribution. C) The autocorrelogram of the two parameters. The upper parameter (MtoR) has a higher autocorrelation, thus the effective number of independent samples we drew from the posterior is smaller than for the lower parameter (Rinit).

6.4.4 Posterior Predictive / Model checks

After we have samples of the posterior distribution and preferably before interpreting the results we need to check the adequacy of our model. A powerful tool of generative models is, that they are able to simulate new data from the current estimated parameters. These new data should capture the important dynamics and effects of our original data. Otherwise the model would be inadequate. When we sample new data from the posterior parameter estimates, this process is called posterior predictive model check. In our example, we sample 1000 new traces from our posterior parameter estimate distributions (Figure 4). Because we randomly sample from a distribution of estimates, each trace will be a little bit different. Our original data should be in the 95% credibility interval of the posterior predictive set. Only after we ensured that the model is adequate, we inspect the posterior parameter estimates visually or calculate and interpret summary statistics (often median and

percentiles) of the parameter distributions and interpret the results.

The posterior check reveals two problems with our model. In the initial phase, marked with (1), we observe a mismatch between the observed and the predicted data. Here, the posterior predictive indicate that the current is slowly increasing, whereas the data indicate no such trend. This first model mismatch can be readily explained: In the initial phase, the expressed melanopsin proteins are not activated, they need a first activation by blue light, before they can acquire a new photoactive equilibrium between the R and M state. But the model assumes falsely, that this equilibrium can be acquired from the beginning. By either excluding this portion or adding another initial state for melanopsin, this difference could be modeled. The model mismatch at (2) is currently not well understood. Even though blue light is still activating melanopsin proteins and forcing them to the M state, the current is diminished again. Mechanism that are able to resolve this range from internalization of receptors, to effects of delay due to the G protein-coupled receptors (GPCRs) or due to a refractory period of melanopsin where the stability of the Schiff-base, which connects the retinal with the protein is unstable (Tsukamoto et al., 2015) This is cannot be captured by the current model and thus the posterior predictive show an expected maximum at the end of the blue period. The posterior checks revealed two problems with this model, especially the first one could bias our parameters. To cope with these problems is left open for now, but it is not difficult to resolve them by enhancing the model.

We are now ready to interpret our parameters for this single cell fit. We expected the initial R-state parameter to be around one, due to our baseline correction. This is indeed the case, the average initial state for R is 1 [0.99,1]. The estimated standard deviation (the estimated measurement noise) of our signal is 0.052

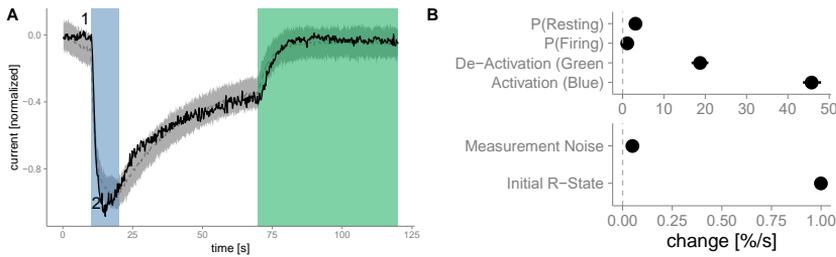


Fig. 6.4 A) The light gray band depicts the 95% credibility interval of 1000 posterior predictives. Posterior predictives are 'new cells' that are simulated from our posterior parameter estimates and reflect the range of possible outcomes of the posterior model fit. The dark gray line depicts the median posterior predictive value. The black curve depicts the original data. The annotation "1" and "2" are discussed in the text. B) Parameter estimates of the single cell shown in A). Median and 95% percentiles are shown.

[0.050,0.054]. We defined four main parameters in our model: The first is c_{RM} , it indicates the constant and spontaneous transition probability from the resting to the active state. The estimate is -6.5 [-6.6, -6.5] on the logit scale. In order to convert this to a more sensible unit, first we take the inverse of the logit function. Then we need to raise the 1-x to the sampling frequency to gain the probability per second:

$$p(x)\left[\frac{\%}{s}\right] = 1 - (1 - \text{invlogit}(x))^{Fs}$$

Thus converted to percent per second, the spontaneous change is on average 1.4 $\frac{\%}{s}$ [1.3 $\frac{\%}{s}$, 1.5 $\frac{\%}{s}$]. The spontaneous transition back to the resting state c_{MR} is the second parameter and for this cell it is a bit higher with on average 3.9 $\frac{\%}{s}$ [3.7 $\frac{\%}{s}$, 4.0 $\frac{\%}{s}$]. We can also construct the equilibrium point from these data, $(1.4)/((1.4+3.9)) \approx 0.26$, thus we expect the equilibrium state to be at around 26% of the maximal theoretical current (the maximal M state). In order to convert the parameters L_{RM} , the activation by blue light, one

needs to take the concurrent constant change into account, the formula changes to:

$$p(x)\left[\frac{\%}{s}\right] = 1 - (1 - (\text{invlogit}(\text{const} + x) - \text{invlogit}(\text{const})))^{Fs}$$

Thus for the activation by blue light we get $38.1 \frac{\%}{s}$ [$36.8 \frac{\%}{s}$, $39.6 \frac{\%}{s}$] and for green light deactivation we see a change of on average $24.7 \frac{\%}{s}$ [$22.7 \frac{\%}{s}$, $26.9 \frac{\%}{s}$]. Keep in mind that this is an estimate for a single cell, thus the posteriors are comparably tight, the uncertainty about the parameters is low. More complex models take the data of multiple cells in account and are introduced in the next chapter. This concludes the bayesian model fit. To go further from here we recommend the introduction book by Kruschke (Kruschke, 2014), the applied problem-centered book by Wagenmaker (MD Lee and EJ Wagenmakers, 2014) and the book by Gelman (Gelman, Carlin, et al., 2013).

6.5 Model Extensions: Multiple cells, priors, different melanopsin variants and latent models

We are now ready to discuss further enhancements to the model. We advocate to start simple, with a basic working model and after thorough checks, add the modules that are needed for your analysis.

6.5.1 Hierarchical Model Fit

We successfully estimated parameters for a single cell. Now we need to check whether these hold for the whole population of cells. A standard procedure is estimating the parameters of each cell individually and then taking the average as the population average. This is a valid and straight forward approach, but has some drawbacks: Cells where parameters are difficult to estimate are weighted the same as cells where parameters are certain. In a similar vein, the single cell parameter estimates are not influenced by the parameters of other cells, even though we can leverage this population knowledge to get better single cell estimates. In recent years mixed linear models (also known as hierarchical models) are becoming more and more popular. In mixed models we fit all cells at the same time and assume that the parameter value of each cell is sampled from a parent population parameter-distribution. In Figure 5 A we see that the single cell estimates (green, line shows mean and distribution shows the estimation precision) are samples from an overarching population distribution of the parameter, in this case a normal distribution with two parameters. If all cell-parameters are sampled from the population-distribution, it is reasonable to expect that single cell parameters that are in the tail

of the population (thus extreme outcomes or outliers) are unlikely. We thus move our single cell estimate closer to the mean of the population-distribution, an effect termed shrinkage. The amount of shrinkage depends on the probable distribution of the single cell mean and the distance of the cell mean to the population mean (the variance of the population needs to be included in the distance). The population distribution parameters are estimated concurrently to the shrunk single cell estimates. Because we estimated parameters from the same cell, similar to a within-subject design, we expect that multiple parameters could be correlated with each other (Figure 5 C). For example, if we estimate the refractory period of a neuron to be short we might also suspect that it shows a higher maximal firing rate. Thus when estimating multiple parameters of a cell, we have to take correlations between the parameters into account. This also allows for shrinkage over the correlation parameter. If there is no correlation in the data nor prior, the estimate will also be close to zero and shrinkage will not take place. Because population distributions are usually normal distributions we can elegantly assume all parameters are based on a multivariate normal distribution with means, variances and a correlation matrix (or equivalently means and a covariance matrix). This part is equivalent to a linear mixed model where all parameters have random slopes and the complete correlation matrix is estimated.

In practical terms we need to introduce some more parameters to be estimated. The model is kept untouched for the critical calculations in each time step, but of course the underlying data and the parameters are different for each cell. We introduce a matrix notation in the code, where the parameters are saved in a matrix termed beta (dimensions n-parameter * n-cells). We also introduce population-vectors with the prefix 'm_' or 's_' for population mean value or population standard deviation value, for example m_{beta} with n-parameters for the mean. Further we need a correlation

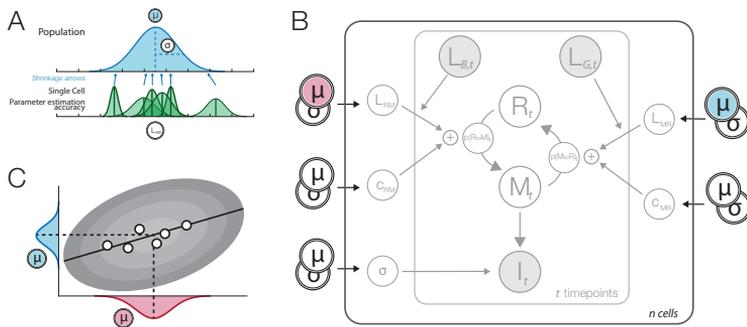


Fig. 6.5 A) Hierarchical parameters. The blue distribution is our population distribution with two parameters, μ and σ . Below the posterior estimates of the single cells are shown in green. They give an estimate of the mean and its uncertainty. When fitting a hierarchical model, the single cell posterior means are shrunk towards the population mean (blue arrows). Shrinkage is strongest for uncertain (broad distribution) parameters and parameters that are furthest away from the population mean. B) Hierarchical model graph. The parameters of a single cell (see Figure 2) are assumed to be sampled from an overarching population distribution. Thus each single-cell parameter is assumed to come from a population with mean and variance as shown in A). C) Dots represent parameter estimates of single cells. Here we observe a correlation between two population parameters. In order to capture this relationship, we need to include the correlation term between all population parameters and model it as one combined multivariate normal distribution

matrix (n-parameter * n-parameter) and a population-vector for the variance (dimensions n-parameter, needs to be positive). In stan we can conveniently calculate the covariance matrix using:

$$covmat < -quadformdiag(corrmat, s_{beta});$$

Finally we need to define the relation of the single cell parameters with the multivariate population:

$$beta[cell_i, dx] \propto multinormal(m_{beta}, covmat);$$

This statement is repeated for each cell through a loop. This states that the beta values (the n-parameter dimension is vectorized, thus hidden) are sampled from a multivariate normal with the given mean and covariance matrix. The initial value of R for each cell has to be between 0 and 1. But if we sample from a normal distribution with mean 0.9 and SD of 0.1, we will sometimes sample values greater than 0. We can simply ignore those values and in those cases resample until we get a value <1. Alternatively we can use a function that is strictly bounded between 0 and 1, for example a beta-distribution:

$$Rinit[cIdx] beta(alpha_{Rinit}, beta_{Rinit})$$

Using pairwise scatter plots of the MCMC values, we noticed that two parameters of the posterior estimates are highly correlated: c_{RM} , the spontaneous change to an active state, and L_{RM} the activation through blue light. The correlation stems from the linear model definition and due to the logit scale. In order to activate the cell by blue light, L_{RM} needs to act against the very large negative number of c_{RM} (a large negative number on the logit scale forces the constant firing probability of the cell to be close to 0). The change at each point in time is:

$$p(RtoM) = invlogit(c_{RM} + L_{RM})$$

Thus L_{RM} needs to counteract c_{RM} . Let's take for example $c_{RM} = -10 \pm 1$ ($\approx 0.0001 \frac{\%}{s}$). When light activates the cell, the total should be around $75 \frac{\%}{s}$:

$$p(RtoM) = \text{invlogit}(1 \pm 0.1)$$

Therefore it is clear that $L_{RM} = 11 \pm 1.1$. Because the value of c_{RM} is expectedly very negative on the logit scale, it will always be a large part of L_{RM} and therefore we get the correlations. This is problematic for MCMC sampling algorithms, they do not converge well with high correlations between parameters. There is a trick to reduce the correlation: reparameterization. Reparameterization changes how parameters are related to each other. It only changes the sampling procedure, but not the outcome or the estimated model because we keep the relation between parameters the same. In this case we change:

$$L_{RM}[cIdx] < -\text{beta}[cIdx, 3]$$

to

$$L_{RM}[cIdx] < -\text{beta}[cIdx, 3] - c_{RM}[cIdx]$$

Because we sample beta and not L_{RM} , we changed the parameter space that is sampled by the MCMC algorithm, to one that does not show the high correlation between parameters, but we don't change the actual parameter value. The reparameterization greatly reduced the time to convergence and in addition improved the effective samples N_{eff} . With some simple addition to the model we are now able to estimate shranked parameter values for all cells concurrently. This model is more complex than the simple model, in order for it to converge we needed to initialize the chains at values in the range of the posterior, we used the same values on both the single cell and the population level and initialized the means but not the variances.

It is now necessary to draw posterior predictives to evaluate whether our model is adequate. In hierarchical models, we can perform posterior predictives in at least two cases: either we take the estimated parameters of each cell and do the same procedure as in the basic model for each cell, or we sample "new cells" from the estimated population multivariate normal distribution. These predicted new cells reflect the range of possible results predicted by our model, prior and parameter estimates. For ease of display, we directly plot the amount of M state without the additional noise term added. The first case, selecting the parameters of the single cell, can be seen in Figure 6A. Here the posterior predictives match the real data (Figure 1 B) very well. In the second case we sample new cells, as expected, this results in a broader distribution (Figure 6 B). The general shape again matches the original data very closely.

After the model posterior predictive tests we look at the results of the model. Similarly to the posterior predictive we can observe results at two different levels. Those two levels, single cell and population, can be seen in figure Figure 7 A,B. In the top posterior estimates of the population distribution, the median distribution and the mean $\pm 95\%$ credibility interval of the mean are shown. In the lower row the single cell uncertainty estimates and their respective means are shown. The population distribution should

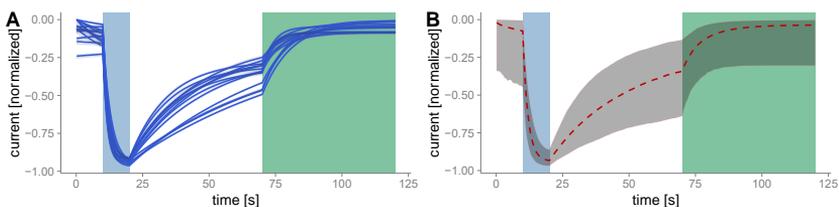


Fig. 6.6 The shaded region depicts the 95% interval of posterior predictives A) Single cell posterior predictive. The parameter estimate of each cell was used to sample new timeseries. Posterior predictives are very similar for single cells, thus the shaded region is nearly invisible. B) Posterior predictive if we sample new cells from the population distribution. Compare with Figure 1 to see the similarity.

match the distribution of the single cells, as is the case in A, for L_{MR} and in B for the beta-distribution of the initial R-state.

We can summarize the values using median and 95% percentiles as in Figure 7 C. The spontaneous firing rate is $0.5 \frac{\%}{s}$ [$0.2 \frac{\%}{s}$, $1.0 \frac{\%}{s}$], while the spontaneous change to the resting state is $2.7 \frac{\%}{s}$ [$2.1 \frac{\%}{s}$, $3.4 \frac{\%}{s}$], thus the equilibrium point of the population is at $17.4 \frac{\%}{s}$ [$8.6 \frac{\%}{s}$, $28.6 \frac{\%}{s}$]. Activation by blue light changes the transition probability by $37.3 \frac{\%}{s}$ [$34 \frac{\%}{s}$, $41.1 \frac{\%}{s}$] whereas green light deactivates with a lesser rate of by $9.1 \frac{\%}{s}$ [$8.1 \frac{\%}{s}$, $12.0 \frac{\%}{s}$]. We can also estimate the probabilities of the cell we fitted in the beginning, which will be affected by the shrinkage factor. Here we see that the single cell estimate of the spontaneous firing rate was $1.4 \frac{\%}{s}$ [$1.3 \frac{\%}{s}$, $1.5 \frac{\%}{s}$] but in the hierarchical model it is $1.94 \frac{\%}{s}$ [$1.64 \frac{\%}{s}$, $2.04 \frac{\%}{s}$]. Thus the shrinkage moved the single cell estimate towards the population-mean of $2.7 \frac{\%}{s}$. This new estimate will be a better prediction of a new measurement of the same cell because it is informed by the estimates of all other cells via shrinkage. In order to fit multiple cells we needed to add hierarchical population distributions and use a reparameterization-trick. From the model we can sample new cells and estimate in what range new cells will be.

6.5.2 Priors

Another strength of Bayesian data analysis is the possibility to add prior knowledge to your data. In STAN this is straight forward, for example if we expect that our estimated noise-level is around 0.02 with a standard deviation of 0.01 we add in the STAN-model block:

$$\sigma_{Current} \propto normal(0.02, 0.01)$$

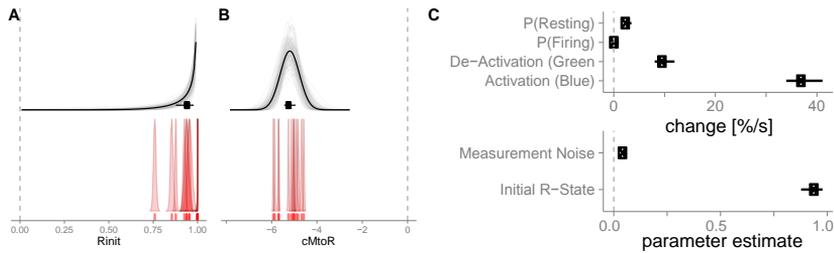


Fig. 6.7 A,B) Population distribution of deactivation with green light (L_{MR}) and the initial R state (R_{init}) respectively with 100 redraws from the posterior chains, the pointrange depicts the mean and 95%-percentile. The lower plots depict single cell posterior estimates and respect mean posterior. L_{MR} is depicted on the logit-scale. C) Results of all parameters. The top plot is in $\frac{\%}{s}$, the lower in natural units for the respective parameters.

The MCMC sampler incorporates this prior in the appropriate way and integrates it with the likelihood of the standard deviation of the data. Importantly, if we would use a uniform-prior, for example 0.01 - 0.03, we restrict the domain of possible parameter values. Thus even if we have strong evidence from the data that the standard deviation should be 0.05, our posterior will not be able to put any weight, because the prior is zero. This cannot happen with the above normal distribution, because the normal has non-zero weight (albeit very small) from minus infinity to infinity. Another more elaborate example could be to include previously measured biological constants into the model. For example Emanuel and Do 2015 (Emanuel and Do, 2015) proposed a numerical three state model for melanopsin based on biochemical data. They made use of photon absorption rates, spectral templates and quantum efficiencies to simulate the wavelength dependencies of the distribution of states. They then qualitatively compared it to their data and concluded that melanopsin can occur in three states. It is very well possible to enhance the model and include these biochemical data as priors in the data fit and estimate the certainty of the posterior. Priors allow to appropriately incorporate scientific knowledge already at the stage of data fitting.

6.5.3 Wavelength Dependencies

So far we activated and deactivated melanopsin using two distinct wavelengths. But we can repeat this process with many other wavelengths as well. In that case we are interested to model an activation and a deactivation function of melanopsin based on the wavelength. Of course this function is a priori unknown. While it is possible to use non-parametric basis-functions (e.g. splines) to estimate a non-linear form of the function, in our case there is reasonable evidence (Emanuel and Do, 2015; Spoida et al., 2016), that the activation function follows a Gaussian tuning function. We incorporate this in our model (Figure 8 A) and decided to use a Gaussian with three unknown parameters: a mean, a variance (those two parameters regulate at which wavelengths the cell get de/activated) and a normalization parameter which regulates the strength of the de/activation (Figure 8 B).

6.5.4 Bi- vs tri-stability

It has recently been suggested, that Melanopsin has not two states but a third one (Emanuel and Do, 2015). In that case parts of the M state transfiguration change not to the R state, but to the E (extramelanopsin) configuration. In analogue to Emanuel & Do who proposed a numerical three state model simulation, we add this third state X. Therefore the model changes as follows:

$$\begin{aligned}dR &= M_{toR} * M[t-1] - R_{toM} * R[t-1]; \\dE &= M_{toE} * M[t-1] - E_{toM} * E[t-1]; \\dM &= E_{toM} * E[t-1] + R_{toM} * R[t-1] - M_{toR} * M[t-1] - M_{toE} * M[t-1];\end{aligned}$$

In this model specification transitions from R to E and vice versa are not allowed, which is grounded in energetic constraints where a direct conversion from R to E state has not been observed (Matsuyama et al., 2012). We could perform bayesian model selection

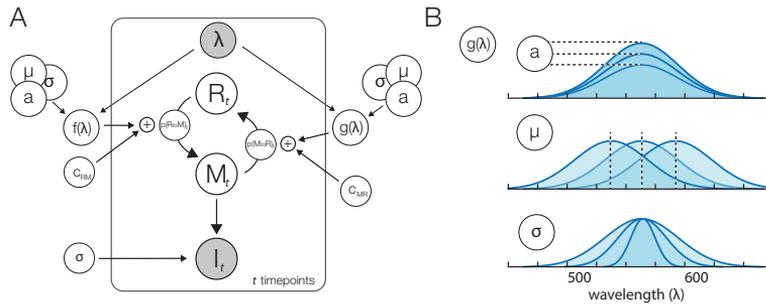


Fig. 6.8 A) Graphical model with wavelength dependency. We replaced the two light sources with a single one that is able to change the wavelength and two functions that translate the wavelength to an activation or deactivation probability. B) The wavelength functions have three parameters. Parameter a regulates how strongly the de/activation is. Parameter μ regulates at which location the maximal de/activation is to be expected and parameter σ regulates on what range the de/activation can occur.

on the two state against the three state model to see which model shows more support from the data. This can generally be done using the bayes-factor or an information criterion for example DIC or WAIC. A discussion of the differences or preferences can be found for example in (Gelman, Carlin, et al., 2013). It is to be expected that we need similar data as (Emanuel and Do, 2015) to be able to show that melanopsin has indeed three states. If our current data is already well explained by two states, adding a third state will not improve the model-fit, if we punished for using the additional number of parameters. Indeed a model with three states of a single cell has a WAIC of -5135, while the two state models has only -3696, where a higher number is better. This does not indicate that melanopsin has two states, only that two states are adequate to describe the very limited data gained from a single cell. This module shows the extension of our basic model to be able to directly test two competing hypothesis in a single coherent framework of data analysis.

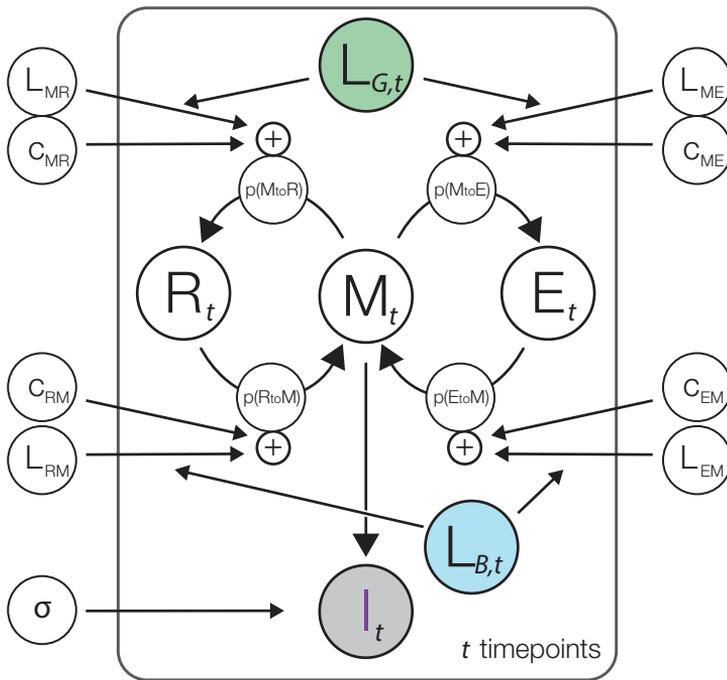


Fig. 6.9 We add a third state to our model. Direct changes from R to E or E to R are not allowed. Therefore we only need to add new changes from M to E (EtoM and MtoE)

6.5.5 Differences between cell types

Melanopsin occurs in different species and has slightly different sequences. Two types can show different activation dynamics and thus different underlying kinetic parameters. We recorded data that allows us to compare a human melanopsin (hOpn4L) to a mouse-origin melanopsin (mOpn4L). We use the basic model with the hierarchical model extension for multiple cells. In our model, we can include this as a factor in the linear model. Thus we adapt two lines in our code:

```

$$p(MtoR)[cell] < -invlogit(c_{MR}[cell] + L_{Green}[t] * L_{MR}[cell] +$$
  

$$i s_{mouse}[cell] * (c_{MR-mouse}[cell] + L_{Green}[t] * L_{MR-mouse}[cell]));$$

```

And the respective $p(RtoM)$ line as well. The idea is to model both cell-types with the same parameters, but allow the parameters to differ if the data of a mOpn4l cell is being fitted. This is the same way one would model this with treatment coding in a classic linear model. We end up with the parameters for a hOpn4L cell (c_{MR} and L_{MR}) and the difference in the parameters to a mOpn4l cell ($c_{MR-mouse}$ and $L_{MR-mouse}$). If we would like get the parameter estimate for mOpn4l directly, we can simply add the two estimates. The results of this model can be seen in the Figure 9 B.

The differences between the two protein variants kinetics' can be explained by various factors. For instance, the active center that connects the photoactive retinal to the protein using a Schiff-base is more stable for mOpn4L than hOpn4L (Tsukamoto et al., 2015). They also differ in phosphorylation of the intracellular loops and the C-termini (Blasic, Lane Brown, et al., 2012; Blasic, Brown, et al., 2012; Blasic, Matos-Cruz, et al., 2014; Fahrenkrug et al., 2014).

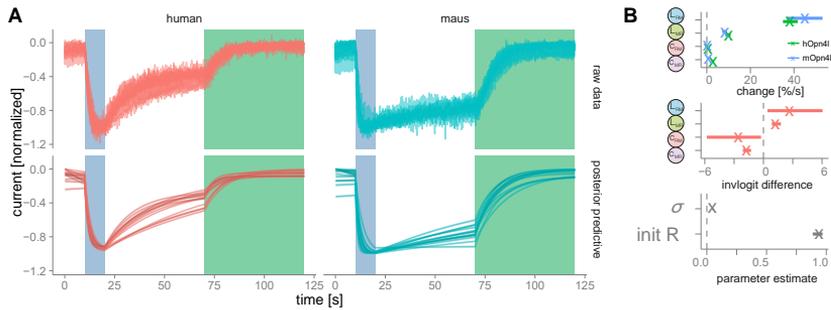


Fig. 6.10 A) The red curves depict 13 cells with human melanopsin (hOpn4L). The blue curves depict 14 mouse melanopsin (mOpn4L). The upper panel depicts the preprocessed raw data, the lower panel the mean posterior predictive model fit. B) The plot depicts the parameter and their differences of a combined model fit of hOpn4L and mOpn4L

But other proteins or factors cannot be excluded to influence the differences as well.

6.6 Chapter conclusion

In this paper we developed a basic generative model for the activation and deactivation kinetics of melanopsin. We inverted the model using bayesian parameter estimation in the STAN framework and show how to interpret the parameters of the model and how to predict future data from the model. Using our generative model we are able to inform new experiments and directly tackle uncertainties of underlying parameters.

General Discussion

7.1 Summary

In this thesis I followed visual information throughout the brain. I discussed how new categories are learned, how internal information is used for decision making and for predicting future input over eye movements. I then showed in a new paradigm how to explore the mechanisms that generate the decisions where and when to look next. Finally I introduced the framework of a Bayesian Generative model based on the example of Melanopsin. In the following, I will give a short summary over the chapters followed up with discussions on the connections between the topics.

In the second chapter, I started with the observation that the human visual system is able to distinguish naturally occurring categories with exceptional speed and accuracy. This hints at a rapid, fixed, one-shot process. Nevertheless, the brain is also capable of learning completely new categories easily. In our study (Kietzmann and König, 2015), we taught subjects new categories to investigate how these category formation develops over the course of learning. In these kind of studies greatest care needs to be employed to account for possible low-level confounds. That is, it could be that the categories to be learned are more similar within themselves than between them by accidental design. Therefore, we adopted a learning paradigm with strictly controlled parametric stimuli. In the study, we made use of Magnetoencephalography (MEG), a method to record brain activity with millisecond resolution. We developed

a new adaptation paradigm, similar to fMRI adaptation, and used a new analysis approach to detect signals of category selectivity. During learning, subjects showed a remarkable shift in cortical activity: After initial training, subjects showed a very high categorization accuracy. We found category selectivity in prefrontal areas starting at 275ms. After twenty-two sessions of training, subjects improved their categorization performance further. Concurrently the category selective activity shifted to 125ms in occipitotemporal cortex. This shift is interesting, as category learning effects are often reported to emerge by multiple different systems (Ashby and Maddox, 2011; EE Smith and Grossman, 2008; Ashby and Ell, 2001) depending on the task-requirements. We build on this idea and interpret our findings accordingly: extended category learning leads to a change through which system categories are extracted. We speculate that category learning starts from a very flexible, slow system in prefrontal areas, and subsequently changes to a fast, feature-based, automatic system in the temporal cortex.

In the third chapter, I was interested in perceptual decision making, especially in the case when percepts are actively generated by the brain. In our study (Ehinger, Häusser, et al., 2017) we made use of the phenomenon of filling-in in the physiological blind spots, the area of the visual field where the optic nerve passes through the retina and no photoreceptors can be found. Even though no direct input from the eyes reaches our brain, the percept at this area is as vivid as in any other region. Therefore, this region gives us the unique opportunity to compare partially inferred to veridical percepts. In the ambiguous condition, when two physically identical stimuli were displayed, subjects showed a bias opposite to the optimal one: They preferred to choose the blind spot stimulus over a veridical one. This is particularly striking as for the blind spot stimulus no relevant information was available to correctly perform the task. We concluded that "Humans treat

filled-in inferred percepts as more real than veridical ones" (Ehinger, Häusser, et al., 2017).

In the fourth chapter I used the filled-in percepts at the blind spots again. In our study (Ehinger, König, et al., 2015), we were interested in violations of peripheral predictions over eye movements. Whenever we perform eye movements, it is reasonable to assume that we produce active predictions on the future incoming signals. We tested this idea by introducing stimulus changes during eye movements. In addition, we modulated the reliability of the peripheral prediction: In half of the trials, the prediction was based on veridical information, in the other half the prediction was based on inferred information from the physiological blind spot. Thus once veridical predictions and once inferred predictions were tested by making eye movements. Concurrently to the eye movements, we measured EEG in order to analyze the occurrence, timing and location of prediction errors. When exchanging the stimulus during the saccade, we observed two separate sets of prediction errors, one early ($<250\text{ms}$) and one late ($>250\text{ms}$) after fixation onset. The early prediction errors showed lateralization, that is a dependence on whether the saccade was performed to the left or the right. This is reasonable, as low-level visual predictions e.g. of the right peripheral field of view, should be represented dominantly in the ipsilateral cortical hemisphere. The late prediction error was a global representation of change. It is interesting to note, that we only observed an interaction of prediction error and the blind spot for the late, but not the early component. This could mean that early prediction errors make use of the inferred prediction *as if* it was veridical, which is not the case for the late prediction errors. Our results also confirm that there are multiple levels of prediction errors in the brain and each processing stage has access to different information on the reliability of the stimulus.

In the fifth chapter I introduced a new paradigm to study where, when and why people look at different parts of a visual scene. We framed eye movements as a constant decision process, deciding whether to exploit the current view and extract information or to move on and further explore the scene. This balance between exploitation and exploration defines the distribution of fixation durations. We developed a new gaze-contingent viewing paradigm that allows us not only to control the location of fixation, but also the timing, that is, the point in time when subjects are allowed to make an eye movement. In other words, we can control the where and when of eye movements. At the same time, our paradigm allows us to utilize reaction times as a proxy for the state of the exploration-exploitation dilemma. In the experiment, subjects observed the underlying stimulus through a 3° (visual degree) aperture. After an experimentally controlled time, a new aperture was visible and subjects continued exploring the image. We observed an exponential decay between the controlled fixation duration and the subsequent reaction time: Short fixation durations elicited longer choice times and vice versa. We interpreted this as a successful interruption of the exploration-exploitation process. Only after a prolonged fixation, all available information was extracted. If we removed the stimulus during the exploitation phase, the subjects' visual system was not yet ready to move on. In addition, in some trials we offered multiple future fixation locations which resulted in an increase of reaction time with increasing number of locations. That is, similar to many decision making tasks (formalized in Hick's law, Hick, 1952), when offered with more alternatives, humans tend to increase their decision time in a logarithmic fashion. We concluded with a description of the merits and limitations of this new exciting paradigm and sketch future applications to the causal study of eye movements.

In the sixth chapter I introduced a tutorial to construct a Bayesian hierarchical generative model. This model type has al-

ready been used in other chapters (Chapter 3 & 5) but was not the focus of these studies. Such models are becoming prominent in many areas, mainly in statistics (Sorensen et al., 2016) and cognitive modeling (MD Lee and EJ Wagenmakers, 2014). Here, I show an application to the kinetics of Melanopsin (Spoida et al., 2016), for instance a hierarchical model to fit multiple cells, a wavelength dependency to investigate Melanopsin reaction to light of different color, and a third state to investigate whether Melanopsin is bi- or tri-stable. I showed how differences between parameters of Melanopsin sourced from two different species can be included into the model. I also discussed how this model can be improved by adding physiological constraints on the parameters in form of priors. Once the model is build, it is easy to interrogate the model on how it would react in hypothetical scenarios. Taken together with the statistical implementations we used in Chapter 3 and 5, this tutorial-chapter shows how powerful and flexible hierarchical Bayesian models are.

7.2 Predictions for eye movements, and decision making in the blind spot

In this section I will discuss the results of two studies relating to the reliability of internal models, specifically internal models of blind spot percepts (Ehinger, Häusser, et al., 2017; Ehinger, König, et al., 2015).

In Chapter 3, I introduced a study on perceptual decision making. We were interested in how the brain relies on filled-in, internally generated percepts. We showed a bias towards the internal model in such decision making paradigm and interpreted this bias in the sense, that the internal model (or equivalently, the

filled-in percept) is attributed a higher reliability than the veridical stimulus representation.

In Chapter 4, I introduce a study that shows that early prediction errors do not differentiate between internal and veridical predictions. In contrast, for late prediction errors, indicated by a P300 electrophysiological correlate, we see a decreased prediction error for the inferred percept. Generally, a decreased P300 amplitude is associated with smaller reliability of the prediction (Polich, 2007; W Sommer et al., 1998; Rosenfeld et al., 2005). This is clearly at odds with the previous study results: Making a prediction only based on the behavioral decision making task, one would expect that the prediction errors should be increased for blind spot predictions compared to veridical ones and not either absent (early prediction errors) or decreased (late prediction errors).

7.2.1 Early Prediction Errors

The missing early prediction errors can have a multitude of explanations. Of course, interpreting compatibility of results with an assumed H_0 of no effect, does not allow one to interpret the results as confirmatory to the H_0 . Thus, first of all we need to discuss sampling errors. Especially with the small number of subjects ($n=15$) in our study, it could be likely that we missed the effect in this sample (nevertheless, each subject had 3000 trials in total). In order to argue for the H_0 , new Bayesian analyses would be needed on this dataset. In general, EEG might not be able to pick up the prediction signal due to physiological constraints of what can be measured with EEG. A slight indication against this explanation is the small effect of filled-in vs. non-filled peripheral stimulus that we demonstrated in the same study (Ehinger, König, et al., 2015).

In order to understand the source of these discrepancies, I propose a new stationary experiment without eye-movements. In this experiment I would show a continuous stimulus monocularly in the periphery while recording EEG. The stimulus should be filled-in or veridical. I would then exchange the stimulus with the perpendicular stimulus as a veridical percept. This experiment could answer whether prediction errors are different based on inferred and filled-in information. Similar to the first control experiment in Chapter 4, one would need to take into account the different density of photoreceptors for the temporal (blind spot) and nasal (veridical) visual field, by additionally testing other areas. I propose this experiment, which is conceptually very similar to our published study, for two reasons: 1) It is a conceptual replication and generalization of the effects found in Ehinger, König, et al. (2015). 2) It removes eye movements from the experiment with the idea to control the fact that predictions for future eye movements are special. Indications that eye movements have a special status comes, for example, from the literature on presaccadic remapping of receptive fields (Rolfs et al., 2011).

7.2.2 Late Prediction Errors

In contrast to the early prediction errors, the late prediction errors showed decreased reliability in the prediction over eye movements. In order to try to explain this discrepancy, I assume that each hierarchical level has access to different information, that is, to different generative models. Support for this assumption comes for example from optical illusions, where we have access to the high-level concept of a stimulus being an illusion, nevertheless we cannot influence our low-level percept. Thus, one explanation for the discrepancy between larger surprise to unreliable stimuli, but behavioral preference in decision making, is that in decision making we do not have access to the full generative model. That

is, the brain puts a higher reliability on the stimulus, because the internally generated representation has a higher signal-to-noise ratio in itself. In the prediction over eye movement context, the brain has additional access to the information that the prediction was produced in the filled-in area and it is able to adjust its P300 response accordingly. This explanation is difficult to test. An experiment is needed that combines decision making and eye-movements. For instance, a filled-in or veridical stimulus could be shown in the blind spot and then checked by performing an eye-movement. One would ask, identical to Chapter 3 / Ehinger, Häusser, et al. (2017), whether the first or second stimulus was continuous, and possibly adjust the difficulty of the task by introducing noise or decreasing the contrast of the inset. Of course recency effects and foveal against peripheral resolution effects would need to be taken into account. This experiment would allow to test whether predictions over eye movements allow the access of special information for decision making.

7.3 Eye movements and Predictions

In Chapter 4 we showed that prediction errors exist over eye movements. We showed large prediction errors if the stimulus is exchanged during the saccade (and thus was unpredicted). We can only speculate if normal saccades also lead to prediction errors due to the mismatch of peripheral and foveal information. I will make this argument and show that there are conceptual indications that this should be the case based on the theory of predictive coding. There are some empirical findings on this. For instance, it is long known, that the amplitude of the P1 (or λ -response in the literature) scales logarithmically with the size of a saccade (Armington and Bloom, 1974; Thickbroom et al., 1991; Dimigen et al., 2011; Dandekar, Privitera, et al., 2012). That is, larger saccades have larger P1 responses. If we take into account that

larger saccades also move the eye into a region that previously was more eccentric in the periphery, this would fit our initial idea: Large eye movements lead to less predictable stimuli, higher uncertainty and larger prediction errors. Note that, that prediction errors can be of equal size when certainty is high but the prediction is slightly wrong, or when uncertainty is low but the prediction is completely wrong. Predictions in the periphery are likely a middle ground where uncertainty exists, but predictions are often accurate on the coarse structures of the visual input (fine detail is not available through peripheral vision).

Combining these ideas with eye movements of unrestricted viewing paradigms, this means that prediction errors measured by ERPs at locations that are well predicted by our internal model should lead to small prediction errors. This allows for predictions in future EEG studies with eye movements behavior. If eye movements are guided to places with high uncertainty, it is likely that initial prediction errors are high and subsequent "double checking" saccades show weaker responses. To get more specific, one possible experiment could be a search paradigm, where the target is only shown after a certain number of saccades. This would allow to compare responses to saccades at locations that were repeatedly revisited. Because less uncertainty on the incoming information exists at revisited locations, I would predict that the P1 response should be smaller. In addition, if a target is gaze-contingently revealed at a revisited fixation location, it should result in a larger prediction error than a previously non-visited location. The guided-viewing bubble paradigm introduced in chapter 5 would be the ideal paradigm, it allows to control (enforce) refixations. This is necessary because during free-viewing, refixations are likely not only elicited by high uncertainty (Wilming, Harst, et al., 2013). Combining the experimental ideas I developed throughout this thesis allows for exiting new ways to study the brain.

7.4 Eye movements and Uncertainty

I proposed to use a search paradigm before. Trying to understand where subjects make eye movements to, will help us to understand what feature the brain regards as important to sample. One influential idea comes from the field of visual search and is based on the minimization of uncertainty. Using Bayesian optimal models as a comparison to human performance, Najemnik and Geisler (2005) showed that humans tend to optimize certainty, that is, minimize uncertainty. They used a visual search paradigm where subjects needed to find a target hidden in noise. Subjects prioritized places where the highest subjective uncertainty exist (and thus the highest possibility that a target is hidden there). This idea has recently been used successfully in several generative models of eye movement behavior (Tatler, Brockmole, et al., 2017; Schütt et al., 2017) (see also K Friston, Adams, et al., 2012). The guided-viewing paradigm presented in Chapter 5 could be used in new search experiments to enhance the validity of these findings. A future step is to analyze the data of our 45 subjects for not only when they made saccades, but also which bubble they selected. One analysis would be to check whether the paradigm also shows optimal choice in the bubble-selection processes (optimal in a sense of maximizing entropy of exploring an image). It is of course also possible to use our paradigm to record a search paradigm and control both the time subjects are exposed to a distractor (or target) as well as where they are allowed to search. I predict results in parallel to our study: A causal dependency of the probability a subjects can recognize a stimulus on the time it is visible. Also longer reaction times to short forced fixation times and longer reaction times when multiple future target locations are shown. In addition, taking Najemnik and Geisler (2005) into account, subjects, when confronted with multiple future target locations, should choose the one that minimizes uncertainty.

7.5 A Statistical Evolution in the Cognitive Sciences

At the beginning of my thesis, a t-test seemed to be state of the art in statistical inference of ERP research (and, often unfortunately, it still is for some ERP literature). In the work presented here, I want to emphasize several statistical advances; two of them (mixed models and Bayesian statistics) are currently taking the field of cognitive neuroscience by storm.

In this section, I want to highlight the role of robust statistical methods. Robust statistical methods are methods that are less severely influenced by outliers and thus represents the main part of the data better. In chapter 4, I make extensive use of robust statistics (Wilcox, 2012). The classical statistical measures of central tendency (e.g. mean) or of spread (e.g. variance) can be heavily influenced by a single measurement point (an outlier). To mitigate this problem, many modern methods are available. For central location, one extreme example is the median, which is not influenced by outliers, but makes certain mathematical procedures unwieldy. Therefore, we make use of methods that take a compromise between both extremes, for example the *winsorized mean* (see Chapter 4.3 for an explanation). By taking a robust approach, one does not loose as much power as by using non-parametric methods, because we still assume the classical assumptions of normality of the sampling distribution. Often only convenience and lack of knowledge are the reasons why the *mean* is used over a *winsorized mean*.

Next, I want to highlight the role of cluster-permutation testing in ERP research. Especially with complex data like EEG, many statistical challenges need to be solved. Today we have the appropriate tools and we should also employ them. Selecting channels

and time windows, *after* the data have been examined, can lead to circular reasoning (J Kilner, 2013; Kriegeskorte, Simmons, PS Bellgowan, et al., 2009). If no a priori information on the selection of channels and time windows are available, it is hard to test all time points at each electrode due to the massive multiple comparison problem. In short, this problem occurs when many tests are performed, and for each the type-I error is controlled only independently from each other. Effective solutions are non-parametric cluster permutation tests (Maris and Oostenveld, 2007), where one does not test each time point x electrode, but connected cluster of significant values over space and time. We used the classical cluster permutation test in Chapter 2, and the successor Threshold-Free Cluster Enhancement (TFCE, SM Smith and Nichols, 2009) in Chapter 4.

Next, I want to highlight the role of mixed / hierarchical models in my and others research. In cognitive psychology, it is very common to record data at two levels. One level is the trial level, the second level is the subject level. A traditional approach for these within-subject data is to aggregate over the trial level to receive a single data point per subject followed by a statistical test like the t-test or, in case of multiple values per subject, a repeated measures ANOVA. The step of aggregating data is suboptimal as soon as unbalanced number of trials or varying variance between subjects exist. This step even becomes impossible if continuous predictors on the trial level are used (e.g. forced fixation time or saccade amplitude, Chapter 5). To solve this problem, most parametric statistical tests can be generalized to the *linear model*. A, conceptually, simple extension is the *mixed* linear model (Gelman and Hill, 2007). The mixed model is becoming the dominant model in cognitive psychology (Bagiella et al., 2000; Quené and Van Den Bergh, 2004; Richter, 2006; Baayen and Milin, 2010; Kliegl et al., 2010) (for exemplar use, see Chapter 3, 5 and 6). At the cost of computational complexity but with increased model flexibility,

the mixed model incorporates unbalanced design with varying variances without any further specifications. Because the linear model is extremely flexible incorporating hierarchical models, multiple dependent variables, non-linear relations and non-normal error distributions, it is currently superseding traditional analyses methods.

Finally, I want to highlight role of Bayesian statistics. In my opinion there are three main reasons why Bayesian statistics are currently en vogue. First, we now have the computational power to fit very flexible models. Second, the current focus is shifting to parameter estimation and not hypothesis testing. And third, Bayesian parameter estimates seem to have a direct interpretation without H_0 "tricks". These three points will now be discussed in more detail. First, I devoted Chapter 6 to show the flexibility of this approach. In Bayesian generative models everything is made explicit, and thus transparent, due to the probabilistic programming code that is used to implement the models¹. Priors of the model are equivalently transparent in this model code. Exchanging basic properties of your model, e.g. assuming your error is normal vs. gamma, is done very transparently by modifying a single line. In the meantime, day-to-day statistics is made very simple because of recent developments in open-source statistical packages like *brms* (Bürkner, 2017), *rstanarm* (Stan Development Team, 2016) or *JASP* (JASP Team, 2018). Second, there is a shift away from p-values and towards confidence intervals and effect size estimates (Kruschke, 2014). Effect size estimation is an instance where many people would like to quantify the uncertainty and the size of the parameter estimates they get from their model. This is exactly what the posterior density of a Bayesian analysis represents. Thus the second major benefit is the direct quantification of uncertainty about a parameter. Third, in statistics many people intuitively assume that the quantity of interest is $p(\theta|data)$. The probability of

¹Except when optimization for speed is relevant

a parameter being in a certain value range, given the data. In order to get this quantity, Bayes theorem has to be applied. Therefore, in the Bayesian world view, the posterior estimates can be directly interpreted as probabilities of the parameter.

These three advantages and a pragmatic view "to use what seems to work" makes it very reasonable arguments to use Bayesian statistic as the default case.

7.5.1 Frequentist and Bayesian statistics

If Bayesian data analysis has these benefits, why is not everyone using them? For this, we have to first understand the frequentist tradition of statistics. Frequentists restrict themselves on the likelihood $p(data|\theta)$, the probability of the data given a specific parameter (RA Fisher, 1922) after specifying a H_0 hypothesis. Fisher, the founding father of frequentist statistics, distanced himself from posterior probabilities (referred to as inverse probabilities):

The theory of inverse probability is founded upon an error, and must be wholly rejected. (Fisher, 1925, p.10)

The trouble for Fisher was, that in order to calculate $p(\theta|data) \propto p(data|\theta)p(\theta)$, the prior $p(\theta)$ has to be "subjectively" specified. Subjectively, because a seemingly uninformative uniform prior depends on the parameterization of the problem, which he finds enough reason to reject the idea of priors (Aldrich et al., 2008; Fienberg et al., 2006). He, in the same way as many frequentists, also argues that hypotheses (replacing θ with e.g. H_0) are either correct or not correct in the real world, but do not have probabilities to be correct. If one follows subjective Bayesians, a correct subjective prior (incorporating everything there is to know about the prob-

lem) allows one to directly interpret $p(\theta|data)$ as the probability that a parameter is in a certain range. Similar, for hypotheses it allows to interpret $p(H|data)$ as the probability that a hypothesis H is true. This Bayesian interpretation of the result of a statistical inference is the intuitive interpretation for many scientists. If one follows the Bayesian interpretation strictly, an additional constraint is needed in order to be allowed to interpret the posterior parameter distribution as a probability of the parameter. The constraint is, that everything needs to be incorporated into the prior, that is a priori known about the parameter. This is usually not possible and objective Bayesian approaches with non- or vaguely-informed priors are used (Berger et al., 2006). This issue of pseudo-Bayes (using more or less ad hoc weakly informed priors) can be relevant in theory (Berger et al., 2006) but in practice often works well (Gelman, Jakulin, et al., 2008).

7.6 An Open Science Evolution of the Cognitive Sciences

Not only statistical practice is changing, but also an evolution of scientific practice can be seen in recent years. This has a large impact on my own way to do science and I will spend a few paragraphs to discuss these changes. The changes are under the heading of *Open Science* and encapsulate two main ideas: *Open Access* and *Open Materials*.

Open Access refers to the idea that the results (publications) of a study paid for by public money should be open to the public. This is currently not the case because most publications are behind pay walls, that is, the public has to pay to read publications even though it has already paid the one producing the publication. The percentage (estimated in 2016) of open access publications up to

the 90s was approximately 20% of articles. Since then, the number has increased to 45% in 2015 (Piwowar et al., 2018). In practice, this number is a bit higher, as many current manuscripts are also available via public preprint repositories, university repositories or private web pages.

Open Materials refers to the idea that the experimental code, the analysis scripts and the data should be made publicly available. This encourages reproducibility (the exact analysis can be reproduced by another scientist on the same data) as well as replicability (the exact protocol can be replicated by another scientist on a new dataset). It has been shown empirically, that enforcing that data is available only upon request, is a poorer option compared to making your data directly available. Stodden et al. (2018) showed that in a sample of 204 articles published in *Science*, 40% of the authors refused or ignored the request, even though *Science* has a data-on-request guideline in place. For my own research, all of my manuscripts are publicly accessible and only one of them was not published as *open access* (Kietzmann, Ehinger, et al., 2016). In addition, most of my data, code and materials are published on the open science framework and in github repositories. I also developed and published several other repositories with visualization and analysis tools during my PhD (<https://github.com/behinger>).

For me, a big benefit to go for open science is that all my analysis choices are transparent. Other researchers can judge, and possibly comment on, those choices and in the best case point out mistakes or unknown limitations of the methods used. In the end, science is supposed to be a self-correcting endeavor on the search for an approximate truth. Making others judge your own work in a transparent and as-impartial-as-possible light is for me an ethical imperative in science.

7.7 Concluding remarks

I will make one final argument in this thesis: The brain is an extremely complex structure. It seems unlikely, that such a complex system can be understood by a single discipline, by studying a single paradigm or a single behavior. Cognitive Science, encompassing many different disciplines, is therefore a very likely candidate to understand the computational and algorithmic principles of the brain. But it is probably not sufficient. More work and understanding crossing scales, from neurons to systems is needed. What is true for Cognitive Science is likely also true for Vision Science. I want to conclude with the plea to study vision using a broad view and multifaceted approaches.

References

- Aarts, AA et al. (2015). “Estimating the reproducibility of psychological science”. In: *Science* 349.6251, pp. 253–267. ISSN: 0036-8075. DOI: 10.1126/science.aac4716.
- Adelson, EH and JR Bergen (1985). “Spatiotemporal energy models for the perception of motion”. In: *Josa a* 2.2, pp. 284–299. DOI: 10.1364/josaa.2.000284.
- Aitchison, L and M Lengyel (2017). “With or without you: predictive coding and Bayesian inference in the brain”. In: *Current opinion in neurobiology* 46, pp. 219–227. DOI: 10.1016/j.conb.2017.08.010.
- Alais, D and D Burr (2004). “The ventriloquist effect results from near-optimal bimodal integration.” In: *Current biology : CB* 14.3, pp. 257–262. ISSN: 0960-9822. DOI: 10.1016/j.cub.2004.01.029.
- Albers, AM, P Kok, I Toni, HC Dijkerman, and FP de Lange (2013). “Shared representations for working memory and mental imagery in early visual cortex”. In: *Current Biology* 23.15, pp. 1427–1431. DOI: 10.1016/j.cub.2013.05.065.
- Aldrich, J et al. (2008). “RA Fisher on Bayes and Bayes’ theorem”. In: *Bayesian Analysis* 3.1, pp. 161–170.
- Alink, A, CM Schwiedrzik, A Kohler, W Singer, and L Muckli (2010). “Stimulus predictability reduces responses in primary visual cortex.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30.8, pp. 2960–6. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3730-10.2010.
- Anderson, AJ, H Yadav, and R Carpenter (2008). “Directional Prediction by the Saccadic System”. In: *Current Biology* 18.8, pp. 614–618. ISSN: 09609822. DOI: 10.1016/j.cub.2008.03.057.

- Anderson, SJ, KT Mullen, and RF Hess (1991). "Human peripheral spatial resolution for achromatic retinal factors". In: *Journal of Physiology* 442, pp. 47–64.
- Antzoulatos, EG and EK Miller (2011). "Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories." In: *Neuron* 71.2, pp. 243–9. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2011.05.040.
- Antzoulatos, EG and EK Miller (2014). "Increases in functional connectivity between prefrontal cortex and striatum during category learning". In: *Neuron* 83, pp. 216–225. ISSN: 10974199. DOI: 10.1016/j.neuron.2014.05.005.
- Arcaro, MJ, PF Schade, JL Vincent, CR Ponce, and MS Livingstone (2017). "Seeing faces is necessary for face-domain formation". In: *Nature neuroscience* 20.10, p. 1404.
- Armington, JC and MB Bloom (1974). "Relations between the amplitudes of spontaneous saccades and visual responses". In: *Journal of the Optical Society of America* 64.9, pp. 1263–1271. ISSN: 0030-3941.
- Ashby, FG and SW Ell (2001). "The neurobiology of human category learning". In: *Trends in cognitive sciences* 5.5, pp. 204–210.
- Ashby, FG and WT Maddox (2011). "Human category learning 2.0". In: *Annals of the New York Academy of Sciences* 1224.1, pp. 147–161.
- Astikainen, P, E Lillstrang, and T Ruusuvirta (2008). "Visual mismatch negativity for changes in orientation—a sensory memory-dependent response." In: *The European journal of neuroscience* 28.11, pp. 2319–24. ISSN: 1460-9568. DOI: 10.1111/j.1460-9568.2008.06510.x.
- Awater, H (2005). "Cortical representation of space around the blind spot". In: *Journal of Neurophysiology* 94.5, pp. 3314–3324. ISSN: 0022-3077. DOI: 10.1152/jn.01330.2004.
- Baayen, RH and P Milin (2010). "Analyzing Reaction Times". In: *International Journal of Psychological Research* 3.2, pp. 12–28. ISSN: 2011-7922. DOI: 10.1287/mksc.12.4.395.
- Bagiella, E, RP Sloan, and DF Heitjan (2000). "Mixed-effects models in psychophysiology." In: *Psychophysiology* 37.1, pp. 13–20. ISSN: 0048-5772. DOI: 10.1017/S0048577200980648.
- Bankó, EM, V Gál, J Körtvélyes, G Kovács, and Z Vidnyánszky (2011). "Dissociating the effect of noise on sensory processing and overall decision difficulty." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31.7, pp. 2663–2674. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2725-10.2011.

- Barr, DJ, R Levy, C Scheepers, and HJ Tily (2013). “Random effects structure for confirmatory hypothesis testing: Keep it maximal”. In: *Journal of Memory and Language* 68.3, pp. 255–278. ISSN: 0749596X. DOI: 10.1016/j.jml.2012.11.001.
- Bastos, AM, WM Usrey, Ra Adams, GR Mangun, P Fries, and KJ Friston (2012). “Canonical microcircuits for predictive coding.” In: *Neuron* 76.4, pp. 695–711. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.10.038.
- Bates, DM (2010). *lme4: Mixed-effects modeling with R*, p. 131. DOI: 10.1177/009286150103500418. URL: <http://lme4.r-forge.r-project.org/lmmwR/lrgprt.pdf>.
- Bates, DM, R Kliegl, S Vasishth, and H Baayen (2015). “Parsimonious mixed models”. In: *arXiv preprint arXiv:1506.04967*, pp. 1–27. DOI: arXiv:1506.04967.
- Bates, DM, M Mächler, B Bolker, and S Walker (2014). “Fitting Linear Mixed-Effects Models using lme4”. In: *Available online at: http://CRAN.R-project.org/package=lme4*. arXiv: 1406.5823. URL: <http://arxiv.org/abs/1406.5823>.
- Baumeister, A and CE Joubert (1969). “Interactive effects on reaction time of preparatory interval length and preparatory interval frequency.” In: *Journal of Experimental Psychology* 82.2, pp. 393–395. ISSN: 0022-1015. DOI: 10.1037/h0028119.
- Becker, W and R Jürgens (1979). “An analysis of the saccadic system by means of double step stimuli”. In: *Vision Research* 19.9, pp. 967–983. ISSN: 00426989. DOI: 10.1016/0042-6989(79)90222-0.
- Beintema, JA, EM van Loon, and AV van den Berg (2005). “Manipulating saccadic decision-rate distributions in visual search”. In: *Journal of Vision* 5.3, p. 1. ISSN: 1534-7362. DOI: 10.1167/5.3.1.
- Bekinschtein, TA, S Dehaene, B Rohaut, F Tadel, L Cohen, and L Naccache (2009). “Neural signature of the conscious processing of auditory regularities.” In: *Proceedings of the National Academy of Sciences of the United States of America* 106.5, pp. 1672–7. ISSN: 1091-6490. DOI: 10.1073/pnas.0809667106.
- Bell, AJ and TJ Sejnowski (1997). “The independent components of natural scenes are edge filters”. In: *Vision research* 37.23, pp. 3327–3338. DOI: 10.1016/s0042-6989(97)00121-1.
- Berg, AV van den and EM van Loon (2005). “An invariant for timing of saccades during visual search”. In: *Vision Research* 45.12, pp. 1543–1555. ISSN: 00426989. DOI: 10.1016/j.visres.2004.12.018.

- Berger, J et al. (2006). “The case for objective Bayesian analysis”. In: *Bayesian analysis* 1.3, pp. 385–402.
- Berger-Tal, O, J Nathan, E Meron, and D Saltz (2014). “The exploration-exploitation dilemma: A multidisciplinary framework”. In: *PLOS ONE* 9.4. ISSN: 19326203. DOI: 10.1371/journal.pone.0095693.
- Blakemore, C and GF Cooper (1970). “Development of the Brain depends on the Visual Environment”. In: *Nature* 227, pp. 680–685. DOI: 10.1038/228477a0.
- Blasic, JR, RL Brown, and PR Robinson (2012). “Phosphorylation of mouse melanopsin by protein kinase A.” In: *PLOS ONE* 7.9, e45387. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0045387.
- Blasic, JR, R Lane Brown, and PR Robinson (2012). “Light-dependent phosphorylation of the carboxy tail of mouse melanopsin.” In: *Cellular and molecular life sciences : CMLS* 69.9, pp. 1551–62. ISSN: 1420-9071. DOI: 10.1007/s00018-011-0891-3.
- Blasic, JR, V Matos-Cruz, D Ujla, EG Cameron, S Hattar, ME Halpern, and PR Robinson (2014). “Identification of critical phosphorylation sites on the carboxy tail of melanopsin.” In: *Biochemistry* 53.16, pp. 2644–9. ISSN: 1520-4995. DOI: 10.1021/bi401724r.
- Boerlin, M and S Denève (2011). “Spike-based population coding and working memory”. In: *PLoS computational biology* 7.2, e1001080.
- Boyden, ES (2015). “Optogenetics and the future of neuroscience”. In: *Nature neuroscience* 18.9, p. 1200.
- Boyden, ES, F Zhang, E Bamberg, G Nagel, and K Deisseroth (2005). “Millisecond-timescale, genetically targeted optical control of neural activity”. In: *Nature neuroscience* 8.9, p. 1263.
- Brainard, DH (1997). “The psychophysics toolbox”. In: *Spatial vision* 10.4, pp. 433–436. ISSN: 0169-1015. DOI: 10.1163/156856897X00357.
- Bredberg, G (1968). “Cellular pattern and nerve supply of the human organ of Corti.” In: *Acta oto-laryngologica*, Suppl 236:1+. ISSN: 0001-6489.
- Bridson, R (2007). “Fast Poisson disk sampling in arbitrary dimensions”. In: *ACM SIGGRAPH 2007 sketches on - SIGGRAPH '07*. New York: ACM Press, 22–es. ISBN: 9781450347266. DOI: 10.1145/1278780.1278807.
- Brunswik, E and L Reiter (1938). “Eindrucks-charaktere schematisierter Gesichter.” In: *Zeitschrift Fuer Psychologie* 142, pp. 67–134.
- Bürkner, PC (2017). “brms: An R Package for Bayesian Multilevel Models Using Stan”. In: *Journal of Statistical Software* 80.1, pp. 1–28. DOI: 10.18637/jss.v080.i01.

- Buswell, GT (1935). *How people look at pictures: A study of the psychology of perception in art*. DOI: citeulike-article-id:3847289.
- Bylinskii, Z, T Judd, A Oliva, A Torralba, and F Durand (2016). “What do different evaluation metrics tell us about saliency models?” In: *arXiv*. DOI: 10.1109/tpami.2018.2815601.
- Caharel, S, O D’Arripe, M Ramon, C Jacques, and B Rossion (2009). “Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component.” In: *Neuropsychologia* 47.3, pp. 639–43. ISSN: 0028-3932. DOI: 10.1016/j.neuropsychologia.2008.11.016.
- Campbell, DT and JC Stanley (1963). “Experimental and quasi-experimental designs for research”. In: *Handbook of research on teaching*. Chicago, IL: Rand McNally.
- Carandini, M, JB Demb, V Mante, DJ Tolhurst, Y Dan, BA Olshausen, JL Gallant, and NC Rust (2005). “Do we know what the early visual system does?” In: *Journal of Neuroscience* 25.46, pp. 10577–10597. DOI: 10.1523/jneurosci.3726-05.2005.
- Carlson, T, H Hogendoorn, and R Kanai (2011). “High temporal resolution decoding of object position and category”. In: *Journal of Vision* 11.10, pp. 1–17. DOI: 10.1167/11.10.9.Introduction.
- Carlson, T, D Tovar, A Alink, and N Kriegeskorte (2013). “Representational dynamics of object vision: The first 1000 ms”. In: *Journal of Vision* 13.10, pp. 1–19. DOI: 10.1167/13.10.1.doi.
- Carpenter, B, A Gelman, M Hoffman, D Lee, B Goodrich, M Betancourt, M Brubaker, J Guo, P Li, and A Riddell (2017). “Stan: a probabilistic programming language”. In: *Journal of Statistical Software* 76.1, pp. 1–32. DOI: 10.18637/jss.v076.i01.
- Carpenter, R and MLL Williams (1995). “Neural computation of log likelihood in control of saccadic eye movements”. In: *Nature* 377.6544, pp. 59–62. ISSN: 0028-0836. DOI: 10.1038/377059a0.
- Carr, D, N Lewin-Koh, and M Maechler (2015). *hexbin: Hexagonal Binning Routines*.
- Cauchoix, M, G Barragan-Jason, T Serre, and EJ Barbeau (2014). “The Neural Dynamics of Face Detection in the Wild Revealed by MVPA.” In: *The Journal of Neuroscience* 34.3, pp. 846–54. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3030-13.2014.

- Chapman, B, MP Stryker, and T Bonhoeffer (1996). "Development of orientation preference maps in ferret primary visual cortex". In: *Journal of Neuroscience* 16.20, pp. 6443–6453.
- Cheadle, S, V Wyart, K Tsetsos, N Myers, V de Gardelle, S Herce Castañón, and C Summerfield (2014). "Adaptive gain control during human perceptual choice." In: *Neuron* 81.6, pp. 1429–41. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2014.01.020.
- Chen, CM, P Lakatos, AS Shah, AD Mehta, SJ Givre, DC Javitt, and CE Schroeder (2007). "Functional anatomy and interaction of fast and slow visual pathways in macaque monkeys." In: *Cerebral cortex (New York, N.Y. : 1991)* 17.7, pp. 1561–9. ISSN: 1047-3211. URL: <http://cercor.oxfordjournals.org/content/17/7/1561.short>.
- Chennu, S, V Noreika, D Gueorguiev, A Blenkmann, S Kochen, A Ibáñez, AM Owen, and Ta Bekinschtein (2013). "Expectation and attention in hierarchical auditory prediction." In: *The Journal of neuroscience* 33.27, pp. 11194–205. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0114-13.2013.
- Churchland, AK, R Kiani, and MN Shadlen (2008). "Decision-making with multiple alternatives". In: *Nature Neuroscience* 11.7, pp. 851–851. ISSN: 1097-6256. DOI: 10.1038/nn0708-851c.
- Cichy, RM, D Pantazis, and A Oliva (2014). "Resolving human object recognition in space and time". In: *Nature Neuroscience* 17, pp. 455–462. ISSN: 1097-6256. DOI: 10.1038/nn.3635.
- Clark, A (2013). "Whatever next? Predictive brains, situated agents, and the future of cognitive science". In: *Behavioral and brain sciences* 36.3, pp. 181–204.
- Cohen, JD, SM McClure, and AJ Yu (2007). "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 362.1481, pp. 933–942. ISSN: 0962-8436. DOI: 10.1098/rstb.2007.2098.
- Cook, TD, DT Campbell, and W Shadish (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Boston. DOI: 10.1198/jasa.2005.s22.
- Cornelissen, FW, EM Peters, and J Palmer (2002). "The eyelink toolbox: eye tracking with MATLAB and the psychophysics toolbox". In: *Behavior Research Methods, Instruments, & Computers* 34.4, pp. 613–617. ISSN: 0743-3808. DOI: 10.3758/BF03195489.

- Cousineau, D (2005). “Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method”. In: *Tutorials in Quantitative Methods for Psychology* 1.1, pp. 42–45. ISSN: 19134126. DOI: noDOIfound.
- Covic, EN and SM Sherman (2011). “Synaptic properties of connections between the primary and secondary auditory cortices in mice”. In: *Cerebral Cortex* 21.11, pp. 2425–2441. DOI: 10.1093/cercor/bhr029.
- Crapse, TB and MA Sommer (2008). “Corollary discharge circuits in the primate brain”. In: *Current opinion in neurobiology* 18.6, pp. 552–557. DOI: 10.1016/j.conb.2008.09.017.
- Crick, F and C Koch (1998). “Constraints on cortical and thalamic projections: the no-strong-loops hypothesis”. In: *Nature* 391.6664, p. 245. DOI: 10.1038/34584.
- Cromer, JA, JE Roy, and EK Miller (2010). “Representation of multiple, independent categories in the primate prefrontal cortex.” In: *Neuron* 66.5, pp. 796–807. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2010.05.005.
- Cronin, B, I Stevenson, M Sur, and KP Körding (2010). “Hierarchical Bayesian modeling and Markov chain Monte Carlo sampling for tuning-curve analysis”. In: *Journal of neurophysiology*. DOI: 10.1152/jn.00379.2009.
- Crossland, MD and PJ Bex (2009). “Spatial alignment over retinal scotomas.” In: *Investigative ophthalmology & visual science* 50.3, pp. 1464–1469. ISSN: 1552-5783. DOI: 10.1167/iovs.08-2690.
- Crouzet, SM and SJ Thorpe (2011). “Low-level cues and ultra-fast face detection.” In: *Frontiers in Psychology* 2, p. 342. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2011.00342.
- Crowe, DA, SJ Goodwin, RK Blackman, S Sakellaridi, SR Sponheim, AW MacDonald, and MV Chafee (2013). “Prefrontal neurons transmit signals to parietal neurons that reflect executive control of cognition.” In: *Nature neuroscience* 16.10, pp. 1484–91. ISSN: 1546-1726. DOI: 10.1038/nn.3509.
- Cudeiro, J and AM Sillito (2006). “Looking back: corticothalamic feedback and early visual processing”. In: *Trends in neurosciences* 29.6, pp. 298–306.
- Curcio, CA, KR Sloan, RE Kalina, and AE Hendrickson (1990). “Human photoreceptor topography.” In: *The Journal of comparative neurology* 292.4, pp. 497–523. ISSN: 0021-9967. DOI: 10.1002/cne.902920402.
- Dale, AM, BR Fischl, and MI Sereno (1999). “Cortical surface-based analysis. I. Segmentation and Surface Reconstruction”. In: *Neuroimage* 9.2, pp. 179–194. ISSN: 1053-8119. DOI: 10.1006/nimg.1998.0396.

- Dale, AM, AK Liu, BR Fischl, RL Buckner, JW Belliveau, JD Lewine, and E Halgren (2000). “Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity.” In: *Neuron* 26.1, pp. 55–67. ISSN: 0896-6273.
- Dan, Y, JJ Atick, and RC Reid (1996). “Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory”. In: *Journal of Neuroscience* 16.10, pp. 3351–3362.
- Dandekar, S, C Privitera, T Carney, and SA Klein (2012). “Neural saccadic response estimation during natural viewing”. In: *Journal of Neurophysiology* 107.6, pp. 1776–1790. ISSN: 0022-3077. DOI: 10.1152/jn.00237.2011.
- Dandekar, S, J Ding, C Privitera, T Carney, and Sa Klein (2012). “The fixation and saccade P3.” In: *PLOS ONE* 7.11, e48761. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0048761.
- Darlington, TR, S Tokiyama, and SG Lisberger (2017). “Control of the strength of visual-motor transmission as the mechanism of rapid adaptation of priors for Bayesian inference in smooth pursuit eye movements”. In: *Journal of neurophysiology* 118.2, pp. 1173–1189. DOI: 10.1152/jn.00282.2017.
- Daw, ND, JP O’Doherty, P Dayan, RJ Dolan, and B Seymour (2006). “Cortical substrates for exploratory decisions in humans.” In: *Nature* 441.7095, pp. 876–9. ISSN: 1476-4687. DOI: 10.1038/nature04766.
- De Pasquale, R and SM Sherman (2011). “Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse”. In: *Journal of Neuroscience* 31.46, pp. 16494–16506. DOI: 10.1523/jneurosci.3664-11.2011.
- De Vries, J, R Azadi, and M Harwood (2016). “The saccadic size-latency phenomenon explored: Proximal target size is a determining factor in the saccade latency”. In: *Vision Research* 129, pp. 87–97. ISSN: 00426989. DOI: 10.1016/j.visres.2016.09.006.
- DeAngelis, GC, I Ohzawa, and RD Freeman (1995). “Receptive-field dynamics in the central visual pathways”. In: *Trends in neurosciences* 18.10, pp. 451–458. DOI: 10.1016/0166-2236(95)94496-r.
- Delorme, A and S Makeig (2004). “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis.” In: *Journal of neuroscience methods* 134.1, pp. 9–21. ISSN: 0165-0270. DOI: 10.1016/j.jneumeth.2003.10.009.
- Di Russo, F, A Martínez, MI Sereno, S Pitzalis, and SA Hillyard (2002). “Cortical sources of the early components of the visual evoked potential.” In: *Human*

- brain mapping* 15.2, pp. 95–111. ISSN: 1065-9471. DOI: 10.1002/hbm.10010.
- DiCarlo, JJ, D Zoccolan, and NC Rust (2012). “How does the brain solve visual object recognition?” In: *Neuron* 73.3, pp. 415–34. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.01.010.
- Dimigen, O, W Sommer, A Hohlfeld, AM Jacobs, and R Kliegl (2011). “Coregistration of eye movements and EEG in natural reading: Analyses and review”. In: *Journal of Experimental Psychology: General* 140.4, pp. 552–572. ISSN: 00963445. DOI: 10.1037/a0023885.
- Do, MTH and KW Yau (2010). “Intrinsically photosensitive retinal ganglion cells.” In: *Physiological reviews* 90.4, pp. 1547–81. ISSN: 1522-1210. DOI: 10.1152/physrev.00013.2010.
- Donchin, E and MGH Coles (2010). “Is the P300 component a manifestation of context updating?” English. In: *Behavioral and Brain Sciences* 11.03, p. 357. ISSN: 0140-525X.
- Donner, TH, M Siegel, P Fries, and AK Engel (2009). “Buildup of choice-predictive activity in human motor cortex during perceptual decision making.” In: *Current Biology* 19.18, pp. 1581–5. ISSN: 1879-0445. DOI: 10.1016/j.cub.2009.07.066.
- Dorris, MC, TL Taylor, RM Klein, and DP Munoz (1999). “Influence of previous visual stimulus or saccade on saccadic reaction times in monkey.” In: *Journal of neurophysiology* 81.5, pp. 2429–2436. ISSN: 0022-3077. DOI: 10.1152/jn.1999.81.5.2429.
- Egner, T, JM Monti, and C Summerfield (2010). “Expectation and surprise determine neural population responses in the ventral visual stream.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30.49, pp. 16601–8. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2770-10.2010.
- Ehinger, BV, D Eickelbeck, K Spoida, S Herlitze, and P König (2016). “Understanding melanopsin using bayesian generative models- an Introduction”. In: *bioRxiv*. DOI: 10.1101/043273.
- Ehinger, BV, P Fischer, AL Gert, L Kaufhold, F Weber, G Pipa, and P König (2014). “Kinesthetic and vestibular information modulate alpha activity during spatial navigation: a mobile EEG study”. In: *Frontiers in human neuroscience* 8, p. 71.
- Ehinger, BV, K Häusser, JP Ossandón, and P König (2017). “Humans treat unreliable filled-in percepts as more real than veridical ones”. In: *eLife* 6. DOI: 10.7554/eLife.21761.

- Ehinger, BV, L Kaufhold, and P König (2018). “Probing the temporal dynamics of the exploration–exploitation dilemma of eye movements”. In: *Journal of Vision* 18.3, pp. 6–6. DOI: 10.1167/18.3.6.
- Ehinger, BV, P König, and JP Ossandón (2015). “Predictions of visual content across eye movements and their modulation by inferred information”. In: *Journal of Neuroscience* 35.19, pp. 7403–7413. DOI: 10.1523/jneurosci.5114-14.2015.
- Einhäuser, W and A Nuthmann (2016). “Salient in space, salient in time: Fixation probability predicts fixation duration during natural scene viewing”. In: *Journal of Vision* 16.11, p. 13. ISSN: 1534-7362. DOI: 10.1167/16.11.13.
- Einhäuser, W, M Spain, and P Perona (2008). “Objects predict fixations better than early saliency”. In: *Journal of Vision* 8.14, pp. 18–18. ISSN: 1534-7362. DOI: 10.1167/8.14.18.
- Emanuel, AJ and MTH Do (2015). “Melanopsin Tristability for Sustained and Broadband Article Melanopsin Tristability for Sustained and Broadband Phototransduction”. In: *Neuron* 85.5, pp. 1043–1055. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2015.02.011.
- Engbert, R, A Nuthmann, EM Richter, and R Kliegl (2005). “SWIFT: A Dynamical Model of Saccade Generation During Reading.” In: *Psychological Review* 112.4, pp. 777–813. ISSN: 1939-1471. DOI: 10.1037/0033-295X.112.4.777.
- Engel, AK, A Maye, M Kurthen, and P König (2013). “Where’s the action? The pragmatic turn in cognitive science”. In: *Trends in cognitive sciences* 17.5, pp. 202–209.
- Ernst, MO and MS Banks (2002). “Humans integrate visual and haptic information in a statistically optimal fashion.” In: *Nature* 415.6870, pp. 429–33. ISSN: 0028-0836. DOI: 10.1038/415429a.
- Fahle, M and M Schmid (1988). “Naso-temporal asymmetry of visual perception and of the visual cortex.” In: *Vision research* 28.2, pp. 293–300. ISSN: 0042-6989. DOI: 10.1016/0042-6989(88)90157-5.
- Fahrenkrug, J, B Falktoft, B Georg, J Hannibal, SB Kristiansen, and TK Klausen (2014). “Phosphorylation of rat melanopsin at Ser-381 and Ser-398 by light/dark and its importance for intrinsically photosensitive ganglion cells (ipRGCs) cellular Ca²⁺ signaling.” In: *The Journal of biological chemistry* 289.51, pp. 35482–93. ISSN: 1083-351X. DOI: 10.1074/jbc.M114.586529. URL: <http://www.jbc.org/content/289/51/35482.short>.

- Feldman, H and KJ Friston (2010). "Attention, uncertainty, and free-energy." In: *Frontiers in human neuroscience* 4.December, p. 215. ISSN: 1662-5161. DOI: 10.3389/fnhum.2010.00215.
- Felleman, DJ and DE Van (1991). "Distributed hierarchical processing in the primate cerebral cortex." In: *Cerebral cortex (New York, NY: 1991)* 1.1, pp. 1-47. DOI: 10.1093/cercor/1.1.1.
- Fenko, L, O Yizhar, and K Deisseroth (2011). "The development and application of optogenetics". In: *Annual review of neuroscience* 34.
- Fetsch, CR, A Pouget, GC DeAngelis, and DE Angelaki (2012). "Neural correlates of reliability-based cue weighting during multisensory integration." In: *Nature neuroscience* 15.1, pp. 146-154. ISSN: 1546-1726. DOI: 10.1038/nn.2983.
- Fienberg, SE et al. (2006). "When did Bayesian inference become Bayesian?" In: *Bayesian analysis* 1.1, pp. 1-40.
- Findlay, JM and LR Harris (1984). "Small Saccades to Double-Stepped Targets Moving in Two Dimensions". In: *Advances in Psychology* 22.C, pp. 71-78. ISSN: 01664115. DOI: 10.1016/S0166-4115(08)61820-8.
- Findlay, JM and R Walker (1999). "A model of saccade generation based on parallel processing and competitive inhibition". In: *Behavioral and Brain Sciences* 22, pp. 661-721. ISSN: 0140-525X. DOI: 10.1017/S0140525X99002150.
- Fiorani Júnior, M, MG Rosa, R Gattass, and CE Rocha-Miranda (1992). "Dynamic surrounds of receptive fields in primate striate cortex: a physiological basis for perceptual completion?" In: *Proceedings of the National Academy of Sciences of the United States of America* 89.18, pp. 8547-51. ISSN: 0027-8424. DOI: 10.1073/pnas.89.18.8547.
- Fischl, B, MI Sereno, RB Tootell, and AM Dale (1999). "High-resolution intersubject averaging and a coordinate system for the cortical surface." In: *Human Brain Mapping* 8.4, pp. 272-84. ISSN: 1065-9471. DOI: 10.1002/(sici)1097-0193(1999)8:4<272::aid-hbm10>3.0.co;2-4.
- Fiser, J, P Berkes, G Orbán, and M Lengyel (2010). "Statistically optimal perception and learning: from behavior to neural representations." In: *Trends in cognitive sciences* 14.3, pp. 119-30. ISSN: 1879-307X. DOI: 10.1016/j.tics.2010.01.003.
- Fisher, RA (1925). *Statistical methods for research workers*. Genesis Publishing Pvt Ltd.
- Fitzgerald, JK, SK Swaminathan, and DJ Freedman (2012). "Visual categorization and the parietal cortex." In: *Frontiers in integrative neuroscience* 6.May, p. 18. ISSN: 1662-5145. DOI: 10.3389/fnint.2012.00018.

- Folstein, JR, TJ Palmeri, AE Van Gulick, and I Gauthier (2015). "Category Learning Stretches Neural Representations in Visual Cortex". In: *Current Directions in Psychological Science* 24, pp. 17–23. ISSN: 0963-7214. DOI: 10.1177/0963721414550707.
- Folstein, JR, TJ Palmeri, and I Gauthier (2012). "Category Learning Increases Discriminability of Relevant Object Dimensions in Visual Cortex." In: *Cerebral Cortex*. ISSN: 1460-2199. DOI: 10.1093/cercor/bhs067.
- Folstein, J, I Gauthier, and T Palmeri (2012). "How category learning affects object discrimination: Not all morphspaces stretch alike". In: *Journal of Experimental Psychology-Learning Memory and Cognition* 38.4.807.
- Foulsham, T and G Underwood (2008). "What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition." In: *Journal of vision* 8.2, pp. 1–17. ISSN: 1534-7362. DOI: 10.1167/8.2.6.
- Foulsham, T and G Underwood (2009). "Does conspicuity enhance distraction? Saliency and eye landing position when searching for objects." In: *Quarterly journal of experimental psychology (2006)* 62.6, pp. 1088–1098. ISSN: 1747-0218. DOI: 10.1080/17470210802602433.
- Freedman, DJ, M Riesenhuber, T Poggio, and EK Miller (2003). "A comparison of primate prefrontal and inferior temporal cortices during visual categorization." In: *The Journal of Neuroscience* 23.12, pp. 5235–46. ISSN: 1529-2401.
- Freedman, DJ, M Riesenhuber, T Poggio, and EK Miller (2006). "Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex." In: *Cerebral Cortex* 16.11, pp. 1631–44. ISSN: 1047-3211. DOI: 10.1093/cercor/bhj100.
- Friedman, D, YM Cycowicz, and H Gaeta (2001). "The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty". In: *Neuroscience & Biobehavioral Reviews* 25.4, pp. 355–373. ISSN: 01497634. DOI: 10.1016/S0149-7634(01)00019-7.
- Friston, K (2005). "A theory of cortical responses." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360.1456, pp. 815–36. ISSN: 0962-8436. DOI: 10.1098/rstb.2005.1622.
- Friston, K (2010). "The free-energy principle: a unified brain theory?" In: *Nature reviews. Neuroscience* 11.2, pp. 127–38. ISSN: 1471-0048. DOI: 10.1038/nrn2787.
- Friston, K, Ra Adams, L Perrinet, and M Breakspear (2012). "Perceptions as hypotheses: saccades as experiments." In: *Frontiers in psychology* 3.May, p. 151. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2012.00151.

- Friston, K, J Kilner, and L Harrison (2006). "A free energy principle for the brain." In: *Journal of physiology, Paris* 100.1-3, pp. 70–87. ISSN: 0928-4257. DOI: 10.1016/j.jphysparis.2006.10.001.
- Gameiro, RR, K Kaspar, SU König, S Nordholt, and P König (2017). "Exploration and Exploitation in Natural Viewing Behavior". In: *Scientific Reports* 7.1, p. 2311. ISSN: 2045-2322. DOI: 10.1038/s41598-017-02526-1.
- Garrido, MI, JM Kilner, SJ Kiebel, and KJ Friston (2007). "Evoked brain responses are generated by feedback loops." In: *Proceedings of the National Academy of Sciences of the United States of America* 104.52, pp. 20961–6. ISSN: 1091-6490. DOI: 10.1073/pnas.0706274105.
- Garrido, MI, JM Kilner, KE Stephan, and KJ Friston (2009). "The mismatch negativity: a review of underlying mechanisms." In: *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 120.3, pp. 453–63. ISSN: 1872-8952. DOI: 10.1016/j.clinph.2008.11.029.
- Gauthier, I, P Skudlarski, JC Gore, and aW Anderson (2000). "Expertise for cars and birds recruits brain areas involved in face recognition." In: *Nature Neuroscience* 3.2, pp. 191–7. ISSN: 1097-6256. DOI: 10.1038/72140.
- Gauthier, I and CA Nelson (2001). "The development of face expertise". In: *Current opinion in neurobiology* 11.2, pp. 219–224.
- Gauthier, I, P Skudlarski, JC Gore, and AW Anderson (2000). "Expertise for cars and birds recruits brain areas involved in face recognition". In: *Nature neuroscience* 3.2, p. 191.
- Gelman, A, JB Carlin, HS Stern, DB Dunson, A Vehtari, and DB Rubin (2013). *Bayesian Data Analysis, Third Edition*. ISBN: 1439840954.
- Gelman, A and J Hill (2007). "Data analysis using regression and multilevel/hierarchical models". In: *Policy Analysis*, pp. 1–651. ISSN: 0012-9658. DOI: 10.2277/0521867061.
- Gelman, A, J Hwang, and A Vehtari (2014). "Understanding predictive information criteria for Bayesian models". In: *Statistics and Computing* 24.6, pp. 997–1016. DOI: 10.1007/s11222-013-9416-2.
- Gelman, A, A Jakulin, MG Pittau, YS Su, et al. (2008). "A weakly informative default prior distribution for logistic and other regression models". In: *The Annals of Applied Statistics* 2.4, pp. 1360–1383.
- Ghasemi, O, ML Lindsey, T Yang, N Nguyen, Y Huang, and YF Jin (2011). "Bayesian parameter estimation for nonlinear modelling of biological pathways." En. In: *BMC systems biology* 5 Suppl 3.3, S9. ISSN: 1752-0509. DOI: 10.1186/1752-0509-5-S3-S9.

- Gillebert, CR, HP Op De Beeck, S Panis, and J Wagemans (2009). "Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences." In: *Journal of Cognitive Neuroscience* 21.6, pp. 1054–64. ISSN: 0898-929X. DOI: 10.1162/jocn.2009.21089.
- Glaholt, M, K Rayner, and E Reingold (2012). "The mask-onset delay paradigm and the availability of central and peripheral visual information during scene viewing". In: *Journal of vision* 12.1, pp. 9–9. ISSN: 1534-7362. DOI: 10.1167/12.1.9.Introduction.
- Gnadt, JW and RA Andersen (1988). "Memory related motor planning activity in posterior parietal cortex of macaque". In: *Experimental brain research* 70.1, pp. 216–220.
- Goeke, CM, S Planera, H Finger, and P König (2016). "Bayesian alternation during tactile augmentation". In: *Frontiers in behavioral neuroscience* 10, p. 187. DOI: 10.3389/fnbeh.2016.00187.
- Gold, JI and MN Shadlen (2007). "The Neural Basis of Decision Making". In: *Annual Review of Neuroscience* 30.1, pp. 535–574. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.29.051605.113038.
- Goldberg, ME, JW Bisley, KD Powell, and J Gottlieb (2006). "Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior". In: *Progress in brain research* 155, pp. 157–175.
- Goldstone, RL, Y Lipka, and RM Shiffrin (2001). "Altering object representations through category learning." In: *Cognition* 78.1, pp. 27–43. ISSN: 0010-0277. DOI: 10.1016/s0010-0277(00)00099-8.
- Goldstone, RL, M Steyvers, and K Larimer (1996). "Categorical perception of novel dimensions". In: *Proceedings of the eighteenth annual conference of the Cognitive Science Society*, pp. 243–248. DOI: 10.1080/713756735.
- Goodwin, S and R Blackman (2012). "Executive control over cognition: stronger and earlier rule-based modulation of spatial category signals in prefrontal cortex relative to parietal cortex". In: *The Journal of ...* 32.10, pp. 3499–515. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3585-11.2012.
- Gosselin, F and PG Schyns (2001). "Bubbles: A technique to reveal the use of information in recognition tasks". In: *Vision Research* 41.17, pp. 2261–2271. ISSN: 00426989. DOI: 10.1016/S0042-6989(01)00097-9.
- Gould, JD (1973). "Eye movements during visual search and memory search." In: *Journal of experimental psychology* 98.1, pp. 184–95. ISSN: 0022-1015. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3997357%7B%5C%7D&tool=pmcentrez%7B%5C%7Drendertype=abstract>.

- Greene, HH, JM Brown, and B Dauphin (2014). “When do you look where you look? A visual field asymmetry.” In: *Vision research* 102, pp. 33–40. ISSN: 1878-5646. DOI: 10.1016/j.visres.2014.07.012. URL: <http://www.sciencedirect.com/science/article/pii/S0042698914001710>.
- Grill-Spector, K, R Henson, and A Martin (2006). “Repetition and the brain: neural models of stimulus-specific effects”. In: *Trends in Cognitive Sciences* 10.1, pp. 14–23. ISSN: 13646613. DOI: 10.1016/j.tics.2005.11.006.
- Grill-Spector, K and R Malach (2001). “fMR-adaptation: a tool for studying the functional properties of human cortical neurons.” In: *Acta psychologica* 107.1-3, pp. 293–321. ISSN: 0001-6918. DOI: 10.1016/s0001-6918(01)00019-1.
- Grimes, J (1996). “On the Failure to Detect Changes in Scenes across Saccades : Perception - oi”. In: *Perception*.
- Gu, Y, DE Angelaki, and GC Deangelis (2008). “Neural correlates of multisensory cue integration in macaque MSTd.” In: *Nature neuroscience* 11.10, pp. 1201–10. ISSN: 1546-1726. DOI: 10.1038/nn.2191.
- Gureckis, TM and RL Goldstone (2008). “The effect of the internal structure of categories on perception”. In: *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*, pp. 1876–1881.
- Hagihara, KM, T Murakami, T Yoshida, Y Tagawa, and K Ohki (2015). “Neuronal activity is not required for the initial formation and maturation of visual selectivity”. In: *Nature neuroscience* 18.12, p. 1780.
- Hankins, M, S Peirson, and R Foster (2008). “Melanopsin: an exciting photopigment”. In: *Trends in neurosciences*. DOI: 10.1016/j.tins.2007.11.002.
- Harris, A and K Nakayama (2007). “Rapid Face-Selective Adaptation of an Early Extrastriate Component in MEG”. In: *Cerebral Cortex* 17.1, pp. 63–70. DOI: 10.1093/cercor/bhj124.
- Harrison, SA and F Tong (2009). “Decoding reveals the contents of visual working memory in early visual areas”. In: *Nature* 458.7238, p. 632. DOI: 10.1038/nature07832.
- Haslinger, R, G Pipa, B Lima, W Singer, EN Brown, and S Neuenschwander (2012). “Context Matters: The Illusive Simplicity of Macaque V1 Receptive Fields”. In: *PLOS ONE* 7.7, pp. 1–17. DOI: 10.1371/journal.pone.0039699.
- Hastings, WK (1970). “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57.1, pp. 97–109. ISSN: 0006-3444. DOI: 10.1093/biomet/57.1.97. URL: <http://biomet.oxfordjournals.org/content/57/1/97>.

- Hayhoe, M and D Ballard (2005). "Eye movements in natural behavior". In: *Trends in Cognitive Sciences* 9.4, pp. 188–194. DOI: 10.1016/j.tics.2005.02.009.
- He, T, M Fritsche, and FP de Lange (2018). "Predictive remapping of visual features beyond saccadic targets". In: *bioRxiv*. DOI: 10.1101/297481.
- Heekeren, HR, S Marrett, PA Bandettini, and LG Ungerleider (2004). "A general mechanism for perceptual decision-making in the human brain." In: *Nature* 431.7010, pp. 859–862. ISSN: 1476-4687. DOI: 10.1038/nature02966.
- Heekeren, HR, S Marrett, and LG Ungerleider (2008). "The neural systems that mediate human perceptual decision making". In: *Nature Reviews Neuroscience* 9.6, pp. 467–479. ISSN: 1471-003X. DOI: 10.1038/nrn2374.
- Helie, S and FG Ashby (2012). "Learning and transfer of category knowledge in an indirect categorization task". In: *Psychological Research* 76, pp. 292–303. ISSN: 03400727. DOI: 10.1007/s00426-011-0348-1.
- Helie, S, JL Roeder, and FG Ashby (2010). "Evidence for cortical automaticity in rule-based categorization." In: *The Journal of neuroscience* 30.42, pp. 14225–14234. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2393-10.2010.
- Henderson, JM and A Hollingworth (2013). "Global change blindness during scene perception". In: *Psychological Science* 14.5, pp. 493–497.
- Henderson, JM, J Olejarczyk, SG Luke, and J Schmidt (2014). "Eye movement control during scene viewing: Immediate degradation and enhancement effects of spatial frequency filtering". In: *Visual Cognition* 22.3-4, pp. 486–502. ISSN: 1350-6285. DOI: 10.1080/13506285.2014.897662.
- Henderson, JM and GL Pierce (2008). "Eye movements during scene viewing: evidence for mixed control of fixation durations." In: *Psychonomic Bulletin & Review* 15.3, pp. 566–573. ISSN: 1069-9384. DOI: 10.3758/PBR.15.3.566.
- Henderson, JM and TJ Smith (2009). "How are eye fixation durations controlled during scene viewing? Further evidence from a scene onset delay paradigm". In: *Visual Cognition* 17. February 2015, pp. 1055–1082. ISSN: 1350-6285. DOI: 10.1080/13506280802685552.
- Heywood, S and J Churcher (1980). "Structure of the visual array and saccadic latency: implications for oculomotor control." In: *The Quarterly journal of experimental psychology* 32.2, pp. 335–341. ISSN: 0033-555X. DOI: 10.1080/14640748008401169.
- Hick, W (1952). "On the rate of gain of information". In: *Quarterly Journal of Experimental Psychology* 4.1, pp. 11–26. DOI: 10.1080/17470215208416600.

- Hodges, JS (2014). *Richly Parameterized Linear Models: Additive, Time Series, and Spatial Models Using Random Effects*. Florida: Chapman and Hall. ISBN: 978-1-4398-6683-2.
- Hodgkin, AL and AF Huxley (1952). “A quantitative description of membrane current and its application to conduction and excitation in nerve.” In: *The Journal of physiology* 117.4, pp. 500–44. ISSN: 0022-3751.
- Hollensteiner, KJ, F Pieper, G Engler, P König, and AK Engel (2015). “Crossmodal integration improves sensory detection thresholds in the ferret.” In: *PLOS ONE* 10.5, e0124952. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0124952.
- Homan, M and A Gelman (2014). “The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo”. In: *The Journal of Machine Learning Research*.
- Hooge, IT and CJ Erkelens (1999). “Peripheral vision and oculomotor control during visual search.” In: *Vision research* 39.8, pp. 1567–1575. ISSN: 0042-6989. DOI: 10.1016/S0042-6989(98)00213-2.
- Hooge, IT, EAB Over, RJA Van Wezel, and MA Frens (2005). “Inhibition of return is not a foraging facilitator in saccadic search and free viewing”. In: *Vision Research* 45.14, pp. 1901–1908. ISSN: 00426989. DOI: 10.1016/j.visres.2005.01.030.
- Huang, Y and RP Rao (2011). “Predictive coding”. In: *Wiley Interdisciplinary Reviews: Cognitive Science* 2.5, pp. 580–593. DOI: 10.1002/wcs.142.
- Hubel, DH and TN Wiesel (1968). “Receptive fields and functional architecture of monkey striate cortex”. In: *The Journal of physiology* 195.1, pp. 215–243. DOI: 10.1113/jphysiol.1968.sp008455.
- Huberle, E and W Lutzenberger (2013). “Temporal properties of shape processing by event-related MEG adaptation”. In: *Neuroimage* 67, pp. 119–126. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2012.10.070.
- Hung, CP, G Kreiman, T Poggio, and JJ DiCarlo (2005). “Fast readout of object identity from macaque inferior temporal cortex.” In: *Science* 310.5749, pp. 863–6. ISSN: 1095-9203. DOI: 10.1126/science.1117593.
- Hupé, Jm, AC James, P Girard, SG Lomber, BR Payne, and J Bullier (2000). “Feedback connections act on the early part of the responses in monkey visual cortex”. In: *Journal of neurophysiology* 85.1, pp. 134–145. DOI: 10.1152/jn.2001.85.1.134.
- Hutton, SB and U Ettinger (2006). “The antisaccade task as a research tool in psychopathology: a critical review”. In: *Psychophysiology* 43.3, pp. 302–313. DOI: 10.1111/j.1469-8986.2006.00403.x.

- Itti, L, C Koch, and E Niebur (1998). “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis”. In: *IEEE transactions on pattern analysis and machine intelligence* 20.11, pp. 1254–1259. DOI: 10.1109/34.730558.
- James, TW and I Gauthier (2006). “Repetition-induced changes in BOLD response reflect accumulation of neural activity.” In: *Human brain mapping* 27.1, pp. 37–46. ISSN: 1065-9471. DOI: 10.1002/hbm.20165.
- JASP Team (2018). *JASP (Version 0.8.6)[Computer software]*. URL: <https://jasp-stats.org/>.
- Jiang, X, E Bradley, RA Rini, T Zeffiro, J VanMeter, and M Riesenhuber (2007). “Categorization training results in shape-and category-selective human neural plasticity”. In: *Neuron* 53.6, pp. 891–903.
- Jonas, JB, U Schneider, and GO Naumann (1992). “Count and density of human retinal photoreceptors”. In: *Graefe’s Archive for Clinical and Experimental Ophthalmology* 230.6, pp. 505–510. ISSN: 0721832X. DOI: 10.1007/BF00181769.
- Kaas, JH (2008). “The evolution of the complex sensory and motor systems of the human brain”. In: *Brain research bulletin* 75.2-4, pp. 384–390. DOI: 10.1016/j.brainresbull.2007.10.009.
- Kamienkowski, JE, MJ Ison, RQ Quiroga, and M Sigman (2012). “Fixation-related potentials in visual search : a combined EEG and eye tracking study”. In: *Journal of Vision* 12, pp. 1–20. DOI: 10.1167/12.7.4.Introduction.
- Kaspar, K and P König (2011a). “Viewing behavior and the impact of low-level image properties across repeated presentations of complex scenes”. In: *Journal of Vision* 11.13, pp. 26–26. ISSN: 1932-6203. DOI: 10.1167/11.13.26.
- Kaspar, K and P König (2011b). “Overt Attention and Context Factors: The Impact of Repeated Presentations, Image Type, and Individual Motivation”. In: *PLOS ONE* 6.7. Ed. by JZ Tsien, e21719. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0021719.
- Kaunitz, LN, JE Kamienkowski, A Varatharajah, M Sigman, RQ Quiroga, and MJ Ison (2014). “Looking for a face in the crowd: fixation-related potentials in an eye-movement visual search task.” In: *Neuroimage* 89, pp. 297–305. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2013.12.006.
- Keliris, GA, NK Logothetis, and AS Tolias (2010). “The role of the primary visual cortex in perceptual suppression of salient visual stimuli”. In: *Journal of Neuroscience* 30.37, pp. 12353–12365.
- Kelly, SP and RG O’Connell (2013). “Internal and external influences on the rate of sensory evidence accumulation in the human brain.” In: *The Journal*

- of neuroscience : the official journal of the Society for Neuroscience* 33.50, pp. 19434–41. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3355-13.2013.
- Kietzmann, TC and P König (2015). “Effects of contextual information and stimulus ambiguity on overt visual sampling behavior.” In: *Vision research* 110.Pt A, pp. 76–86. ISSN: 1878-5646. DOI: 10.1016/j.visres.2015.02.023.
- Kietzmann, TC, BV Ehinger, D Porada, AK Engel, and P König (2016). “Extensive training leads to temporal and spatial shifts of cortical activity underlying visual category selectivity”. In: *NeuroImage* 134, pp. 22–34. DOI: 10.1016/j.neuroimage.2016.03.066.
- Kietzmann, TC and P König (2010). “Perceptual learning of parametric face categories leads to the integration of high-level class-based information but not to high-level pop-out”. In: *Journal of vision* 10.13, pp. 20–20.
- Kilner, J (2013). “Bias in a common EEG and MEG statistical analysis and how to avoid it”. In: *Clinical Neurophysiology* 124.10, pp. 2062–2063.
- Kirchner, H and SJ Thorpe (2006). “Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited.” In: *Vision Research* 46.11, pp. 1762–76. ISSN: 0042-6989. DOI: 10.1016/j.visres.2005.10.002.
- Kleiner, M, D Brainard, and DG Pelli (2007). “What’s new in Psychtoolbox-3?” In: *Perception 36 ECVF Abstract Supplement*.
- Kliegl, R, MEJ Masson, and EM Richter (2010). “A linear mixed model analysis of masked repetition priming”. In: *Visual Cognition* 18.5, pp. 655–681. ISSN: 1350-6285. DOI: 10.1080/13506280902986058.
- Koch, C and S Ullman (1985). “Shifts in selective visual attention: towards the underlying neural circuitry.” In: *Human neurobiology* 4.4, pp. 219–27. ISSN: 0721-9075. DOI: 10.1007/978-94-009-3833-5_5.
- Koehler, K, F Guo, S Zhang, and MP Eckstein (2014). “What do saliency models predict?” In: *Journal of Vision* 14.3, p. 14. DOI: 10.1167/14.3.14.
- Kok, P, JF Jehee, and FP De Lange (2012). “Less is more: expectation sharpens representations in the primary visual cortex”. In: *Neuron* 75.2, pp. 265–270. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.04.034.
- Kollmorgen, S, N Nortmann, S Schröder, and P König (2010). “Influence of low-level stimulus features, task dependent factors, and spatial biases on overt visual attention.” In: *PLoS computational biology* 6.5, e1000791. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1000791.
- Kolossa, A, T Fingscheidt, K Wessel, and B Kopp (2013). “A model-based approach to trial-by-trial p300 amplitude fluctuations.” In: *Frontiers in human neu-*

- rosience* 6.February, p. 359. ISSN: 1662-5161. DOI: 10.3389/fnhum.2012.00359.
- Komatsu, H, M Kinoshita, and I Murakami (2000). "Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 20.24, pp. 9310–9. ISSN: 1529-2401.
- Komatsu, H (2006). "The neural mechanisms of perceptual filling-in". In: *Nature Reviews Neuroscience* 7.3, pp. 220–231. ISSN: 1471-003X. DOI: 10.1038/nrn1869.
- König, P, N Wilming, TC Kietzmann, JP Ossandòn, S Onat, BV Ehinger, RR Gameiro, and K Kaspar (2016). "Eye movements as a window to cognitive processes". In: *Journal of Eye Movement Research* 9, pp. 1–16. DOI: 10.16910/jemr.9.5.3.
- Körding, KP, U Beierholm, WJ Ma, S Quartz, JB Tenenbaum, and L Shams (2007). "Causal inference in multisensory perception". In: *PLOS ONE* 2.9. ISSN: 19326203. DOI: 10.1371/journal.pone.0000943.
- Körding, KP and DM Wolpert (2004). "Bayesian integration in sensorimotor learning." In: *Nature* 427.6971, pp. 244–7. ISSN: 1476-4687. DOI: 10.1038/nature02169.
- Krekelberg, B, GM Boynton, and RJa van Wezel (2006). "Adaptation: from single cells to BOLD signals." In: *Trends in neurosciences* 29.5, pp. 250–6. ISSN: 0166-2236. DOI: 10.1016/j.tins.2006.02.008.
- Kriegeskorte, N, M Mur, DA Ruff, R Kiani, J Bodurka, H Esteky, K Tanaka, and PA Bandettini (2008). "Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey". In: *Neuron* 60, pp. 1–16. DOI: 10.1016/j.neuron.2008.10.043.
- Kriegeskorte, N, WK Simmons, PSF Bellgowan, and CI Baker (2009). "Circular analysis in systems neuroscience: the dangers of double dipping." In: *Nature neuroscience* 12.5, pp. 535–540. ISSN: 1097-6256. DOI: 10.1167/8.6.88.
- Kriegeskorte, N, WK Simmons, PS Bellgowan, and CI Baker (2009). "Circular analysis in systems neuroscience: the dangers of double dipping". In: *Nature neuroscience* 12.5, p. 535.
- Kruschke, J (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. ISBN: 0124059163.
- Kümmerer, M, TSA Wallis, and M Bethge (2015). "Information-theoretic model comparison unifies saliency metrics." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.52, pp. 16054–9. ISSN: 1091-6490. DOI: 10.1073/pnas.1510393112.

- Kwisthout, J and I van Rooij (2013). "Predictive coding and the Bayesian brain: Intractability hurdles that are yet to be overcome." In: *CogSci*.
- Lamme, VA (1995). "The neurophysiology of figure-ground segregation in primary visual cortex". In: *Journal of Neuroscience* 15.2, pp. 1605–1615.
- Lawrence, BM (2010). "An anti-Hick's effect for exogenous, but not endogenous, saccadic eye movements." In: *Experimental Brain Research* 204, pp. 115–118. ISSN: 1432-1106. DOI: 10.1007/s00221-010-2301-8. URL: <http://www.ncbi.nlm.nih.gov/pubmed/20521032>.
- Lawrence, BM, A St John, Ra Abrams, and LH Snyder (2008). "An anti-Hick's effect in monkey and human saccade reaction times." In: *Journal of vision* 8.3, pp. 26.1–7. ISSN: 1534-7362. DOI: 10.1167/8.3.26.
- Leach, J and R Carpenter (2001). "Saccadic choice with asynchronous targets: evidence for independent randomisation". In: *Vision Research* 41.2526, pp. 3437–3445. ISSN: 0042-6989. DOI: 10.1016/S0042-6989(01)00059-1.
- Lee, MD and EJ Wagenmakers (2014). *Bayesian Cognitive Modeling: A Practical Course*. ISBN: 1107653916.
- Lee, TS, D Mumford, R Romero, and VA Lamme (1998). "The role of the primary visual cortex in higher level vision". In: *Vision research* 38.15-16, pp. 2429–2454.
- Levine, MW, BG Cleland, and RP Zimmerman (1992). "Variability of responses of cat retinal ganglion cells." In: *Vis Neurosci* 8.3, pp. 277–279.
- Leys, C, C Ley, O Klein, P Bernard, and L Licata (2013). "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median". In: *Journal of Experimental Social Psychology* 49.4, pp. 764–766. ISSN: 00221031. DOI: 10.1016/j.jesp.2013.03.013.
- Li, S, D Ostwald, M Giese, and Z Kourtzi (2007). "Flexible coding for categorical decisions in the human brain." In: *The Journal of Neuroscience* 27.45, pp. 12321–30. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3795-07.2007.
- Li, X, DV Gutierrez, MG Hanson, J Han, MD Mark, H Chiel, P Hegemann, LT Landmesser, and S Herlitze (2005). "Fast noninvasive activation and inhibition of neural and network activity by vertebrate rhodopsin and green algae channelrhodopsin". In: *Proceedings of the National Academy of Sciences of the United States of America* 102.49, pp. 17816–17821.
- Linden, M van der, M van Turennout, and P Indefrey (2010). "Formation of category representations in superior temporal sulcus." In: *Journal of Cognitive Neuroscience* 22.6, pp. 1270–82. ISSN: 1530-8898. DOI: 10.1162/jocn.2009.21270.

- Linden, M van der, J Wegman, and G Fernández (2013). “Task- and Experience-dependent Cortical Selectivity to Features Informative for Categorization”. In: *Journal of Cognitive Neuroscience*, pp. 1–15. DOI: 10.1162/jocn.
- Litvak, V et al. (2011). “EEG and MEG data analysis in SPM8.” In: *Computational intelligence and neuroscience 2011*, p. 852961. ISSN: 1687-5273. DOI: 10.1155/2011/852961.
- Liu, H, Y Agam, JR Madsen, and G Kreiman (2009). “Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex.” In: *Neuron* 62.2, pp. 281–90. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2009.02.025.
- Liu, J, A Harris, and N Kanwisher (2002). “Stages of processing in face perception: an MEG study.” In: *Nature Neuroscience* 5.9, pp. 910–6. ISSN: 1097-6256. DOI: 10.1038/nn909.
- Liu, N, N Kriegeskorte, M Mur, F Hadj-Bouziane, WM Luh, RBH Tootell, and LG Ungerleider (2013). “Intrinsic structure of visual exemplar and category representations in macaque brain.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 33.28, pp. 11346–60. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.4180-12.2013.
- Loschky, L, G McConkie, J Yang, and M Miller (2005). “The limits of visual resolution in natural scene viewing”. In: *Visual Cognition* 12.6, pp. 1057–1092. ISSN: 1350-6285. DOI: 10.1080/13506280444000652.
- Ludwig, CJH (2009). “Temporal integration of sensory evidence for saccade target selection”. In: *Vision Research* 49.23, pp. 2764–2773. ISSN: 00426989. DOI: 10.1016/j.visres.2009.08.012.
- Ludwig, CJH, JR Davies, and MP Eckstein (2014). “Foveal analysis and peripheral selection during active visual sampling.” In: *Proceedings of the National Academy of Sciences of the United States of America* 111.2. ISSN: 1091-6490. DOI: 10.1073/pnas.1313553111.
- Makeig, S, M Westerfield, TP Jung, S Enghoff, J Townsend, E Courchesne, and TJ Sejnowski (2002). “Dynamic brain sources of visual evoked responses.” In: *Science* 295.5555, pp. 690–4. ISSN: 1095-9203. DOI: 10.1126/science.1066168.
- Marinkovic, K, RP Dhond, AM Dale, M Glessner, V Carr, and E Halgren (2003). “Spatiotemporal dynamics of modality-specific and supramodal word processing.” In: *Neuron* 38.3, pp. 487–97. ISSN: 0896-6273. DOI: 10.1016/s0896-6273(03)00197-1.
- Maris, E and R Oostenveld (2007). “Nonparametric statistical testing of EEG-and MEG-data”. In: *Journal of neuroscience methods* 164.1, pp. 177–190.

- Markov, NT, M Ercsey-Ravasz, C Lamy, ARR Gomes, L Magrou, P Misery, P Giroud, P Barone, C Dehay, Z Toroczkai, et al. (2013). “The role of long-range connections on the specificity of the macaque interareal cortical network”. In: *Proceedings of the National Academy of Sciences* 110.13, pp. 5187–5192.
- Markov, NT, J Vezoli, P Chameau, A Falchier, R Quilodran, C Huissoud, C Lamy, P Misery, P Giroud, S Ullman, et al. (2014). “Anatomy of hierarchy: feed-forward and feedback pathways in macaque visual cortex”. In: *Journal of Comparative Neurology* 522.1, pp. 225–259.
- Marr, D (1982). “Vison”. In: *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*.
- Mars, RB, S Debener, TE Gladwin, LM Harrison, P Haggard, JC Rothwell, and S Bestmann (2008). “Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28.47, pp. 12539–45. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2925-08.2008.
- Martikainen, MH, Ki Kaneko, and R Hari (2005). “Suppressed responses to self-triggered sounds in the human auditory cortex.” In: *Cerebral cortex (New York, N.Y. : 1991)* 15.3, pp. 299–302. ISSN: 1047-3211. DOI: 10.1093/cercor/bhh131.
- Matsumoto, M and H Komatsu (2005). “Neural responses in the macaque v1 to bar stimuli with various lengths presented on the blind spot.” In: *Journal of neurophysiology* 93.5, pp. 2374–87. ISSN: 0022-3077. DOI: 10.1152/jn.00811.2004.
- Matsuyama, T, T Yamashita, Y Imamoto, and Y Shichida (2012). “Photochemical properties of mammalian melanopsin.” In: *Biochemistry* 51.27, pp. 5454–62. ISSN: 1520-4995. DOI: 10.1021/bi3004999.
- Matuschek, H, R Kliegl, S Vasishth, H Baayen, and DM Bates (2015). “Balancing Type I Error and Power in Linear Mixed Models”. In: *arXiv* 2013, pp. 1–14. ISSN: 0749596X. DOI: 10.1016/j.jml.2017.01.001.
- Maunsell, J and DC van Essen (1983). “The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey”. In: *Journal of Neuroscience* 3.12, pp. 2563–2586.
- McGregor, S, M Baltieri, and CL Buckley (2015). “A Minimal Active Inference Agent”. In: *arXiv preprint arXiv:1503.04187*.
- Mckee, XJL, M Riesenhuber, EK Miller, and DJ Freedman (2014). “Task Dependence of Visual and Category Representations in Prefrontal and Inferior

- Temporal Cortices". In: *Journal of Neuroscience* 34.48, pp. 16065–16075. DOI: 10.1523/JNEUROSCI.1660-14.2014.
- Melcher, D (2008). "Dynamic , object-based remapping of visual features in trans-saccadic perception". In: *Journal of Vision* 8.14, pp. 1–17. DOI: 10.1167/8.14.2. Introduction.
- Merriam, EP, CR Genovese, and CL Colby (2007). "Remapping in human visual cortex." In: *Journal of neurophysiology* 97.2, pp. 1738–55. ISSN: 0022-3077. DOI: 10.1152/jn.00189.2006.
- Midgley, CA (1998). "Binocular interactions in human vision Binocular Interactions in Human Vision". PhD thesis.
- Mignard, M and JG Malpeli (1991). "Paths of information flow through visual cortex". In: *Science* 251.4998, pp. 1249–1251. DOI: 10.1126/science.1848727.
- Minamimoto, T, RC Saunders, and BJ Richmond (2010). "Monkeys Quickly Learn and Generalize Visual Categories without Lateral Prefrontal Cortex". In: *Neuron* 66.4, pp. 501–507. ISSN: 08966273. DOI: 10.1016/j.neuron.2010.04.010.
- Mishkin, M and LG Ungerleider (1982). "Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys". In: *Behavioural brain research* 6.1, pp. 57–77.
- Motter, BC and EJ Belky (1998). "The guidance of eye movements during active visual search". In: *Vision Research* 38.12, pp. 1805–1815. ISSN: 00426989. DOI: 10.1016/S0042-6989(97)00349-0.
- Mowbray, GH (1964). "Subjective expectancy and choice reaction times." In: *Quarterly Journal of Experimental Psychology*, 16, 216-223 . DOI: 10.1080/17470216408416371.
- Muhammad, R, JD Wallis, and EK Miller (2006). "A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum." In: *Journal of cognitive neuroscience* 18, pp. 974–989. ISSN: 0898-929X. DOI: 10.1162/jocn.2006.18.6.974.
- Mumford, D (1992). "On the computational architecture of the neocortex". In: *Biological cybernetics* 251, pp. 241–251.
- Mumford, D, S Kosslyn, L Hillger, and R Herrnstein (1987). "Discriminating figure from ground: the role of edge detection and region growing". In: *Proceedings of the National Academy of Sciences* 84.20, pp. 7354–7358. DOI: 10.1073/pnas.84.20.7354.

- Mur, M, DA Ruff, J Bodurka, P De Weerd, PA Bandettini, and N Kriegeskorte (2012). "Categorical, yet graded - single-image activation profiles of human category-selective cortical regions." In: *The Journal of neuroscience* 32.25, pp. 8649–62. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2334-11.2012.
- Murray, SO, D Kersten, Ba Olshausen, P Schrater, and DL Woods (2002). "Shape perception reduces activity in human primary visual cortex." In: *Proceedings of the National Academy of Sciences of the United States of America* 99.23, pp. 15164–9. ISSN: 0027-8424. DOI: 10.1073/pnas.192579399.
- Nadol, JB (1988). "Comparative anatomy of the cochlea and auditory nerve in mammals". In: *Hearing Research* 34.3, pp. 253–266. ISSN: 03785955. DOI: 10.1016/0378-5955(88)90006-8.
- Najemnik, J and WS Geisler (2005). "Optimal eye movement strategies in visual search." In: *Nature* 434.7031, pp. 387–391. ISSN: 0028-0836. DOI: 10.1126/science.1116758.778.
- Nakamura, K and CL Colby (2002). "Updating of the visual representation in monkey striate and extrastriate cortex during saccades." In: *Proceedings of the National Academy of Sciences of the United States of America* 99.6, pp. 4026–31. ISSN: 0027-8424. DOI: 10.1073/pnas.052379899.
- Nienborg, H and BG Cumming (2009). "Decision-related activity in sensory neurons reflects more than a neurons causal effect". In: *Nature* 459.7243, p. 89.
- Noorani, I and RH Carpenter (2016). *The LATER model of reaction time and decision*. DOI: 10.1016/j.neubiorev.2016.02.018.
- Nosofsky, RM (1991). "Tests of an exemplar model for relating perceptual classification and recognition memory." In: *Journal of experimental psychology. Human perception and performance* 17.1, pp. 3–27. ISSN: 0096-1523. DOI: 10.1037/0096-1523.17.1.3.
- Nuthmann, A (2014). "How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements." In: *Journal of Experimental Psychology: Human Perception and Performance* 40.1, pp. 342–360. ISSN: 1939-1277. DOI: 10.1037/a0033854.
- Nuthmann, A (2017). "Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task". In: *Psychonomic Bulletin & Review* 24, pp. 370–392. ISSN: 1069-9384. DOI: 10.3758/s13423-016-1124-4.
- Nuthmann, A and JM Henderson (2012). "Using CRISP to model global characteristics of fixation durations in scene viewing and reading with a common

- mechanism". In: *Visual Cognition* 20.4-5, pp. 457–494. ISSN: 1350-6285. DOI: 10.1080/13506285.2012.670142.
- Nuthmann, A, TJ Smith, R Engbert, and JM Henderson (2010). "CRISP: A computational model of fixation durations in scene viewing." In: *Psychological Review* 117.2, pp. 382–405. ISSN: 1939-1471. DOI: 10.1037/a0018924. URL: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0018924>.
- O'Connell, RG, PM Dockree, and SP Kelly (2012). "A supramodal accumulation-to-bound signal that determines perceptual decisions in humans." In: *Nature neuroscience* 15.12, pp. 1729–35. ISSN: 1546-1726. DOI: 10.1038/nn.3248.
- Olshausen, BA and DJ Field (1996). "Emergence of simple-cell receptive field properties by learning a sparse code for natural images". In: *Nature* 381.6583, p. 607. DOI: 10.1038/381607a0.
- Oostenveld, R, P Fries, E Maris, and JM Schoffelen (2011). "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data." In: *Computational intelligence and neuroscience* 2011, pp. 1–10. ISSN: 1687-5273. DOI: 10.1155/2011/156869.
- O'Regan, JK and A Noë (2001). "A sensorimotor account of vision and visual consciousness". In: *Behavioral and brain sciences* 24.5, pp. 939–973.
- Ossandón, JP, S Onat, and P König (2014). "Spatial biases in viewing behavior". In: *Journal of Vision* 14.2, pp. 20–20. ISSN: 1534-7362. DOI: 10.1167/14.2.20.
- Oswal, A, M Ogden, and RHS Carpenter (2007). "The Time Course of Stimulus Expectation in a Saccadic Decision Task". In: *Journal of Neurophysiology* 97.4, pp. 2722–2730. ISSN: 0022-3077. DOI: 10.1152/jn.01238.2006.
- Palmer, SE (1999). *Vision science: Photons to phenomenology*. MIT press.
- Pama, E, LS Colzato, and B Hommel (2013). "Optogenetics as a neuromodulation tool in cognitive neuroscience". In: *Frontiers in psychology* 4, p. 610.
- Pannasch, S, J Schulz, and BM Velichkovsky (2011). "On the control of visual fixation durations in free viewing of complex images." In: *Attention, perception & psychophysics* 73.4, pp. 1120–32. ISSN: 1943-393X. DOI: 10.3758/s13414-011-0090-1.
- Paradiso, MA and T Carney (1988). "Orientation discrimination as a function of stimulus eccentricity and size: nasal/temporal retinal asymmetry". In: *Vision Research* 28.8, pp. 867–874. ISSN: 00426989. DOI: 10.1016/0042-6989(88)90096-X.
- Pascual-Marqui, R (2002). "Standardized low resolution brain electromagnetic tomography (sLORETA): technical details". In: *Methods & Findings in Experimental & Clinical Pharmacology* 24.D, pp. 5–12.

- Pashler, H and E Wagenmakers (2012). "Editors' introduction to the special section on replicability in psychological science: A crisis of confidence?" In: *Perspectives on Psychological Science* 7.6, pp. 528–530. ISSN: 1745-6916. DOI: 10.1177/1745691612465253.
- Pearl, J (2009). *Causality*. Cambridge university press.
- Pernet, C, M Latinus, T Nichols, and G Rousselet (2015). "Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study". In: *Journal of neuroscience methods* 250, pp. 85–93.
- Pernet, CR, N Chauveau, C Gaspar, and GA Rousselet (2011). "LIMO EEG: a toolbox for hierarchical Linear MOdeling of ElectroEncephaloGraphic data." In: *Computational intelligence and neuroscience* 2011, p. 831409. ISSN: 1687-5273. DOI: 10.1155/2011/831409.
- Petit, JP, KJ Midgley, PJ Holcomb, and J Grainger (2006). "On the time course of letter perception: a masked priming ERP investigation." In: *Psychonomic bulletin & review* 13.4, pp. 674–81. ISSN: 1069-9384. DOI: 10.3758/bf03193980.
- Philiastides, MG and P Sajda (2006). "Temporal characterization of the neural correlates of perceptual decision making in the human brain." In: *Cerebral Cortex* 16.4, pp. 509–18. ISSN: 1047-3211. DOI: 10.1093/cercor/bhi130.
- Pinheiro, J, D Bates, S DebRoy, D Sarkar, and R-Core-Team (2016). *{nlme}: Linear and Nonlinear Mixed Effects Models*.
- Piwowar, H, J Priem, V Larivière, JP Alperin, L Matthias, B Norlander, A Farley, J West, and S Haustein (2018). "The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles". In: *PeerJ* 6, e4375.
- Platt, ML and PW Glimcher (1999). "Neural correlates of decision variables in parietal cortex". In: *Nature* 400.6741, p. 233.
- Plöchl, M, JP Ossandón, and P König (2012). "Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data." In: *Frontiers in human neuroscience* 6. October, p. 278. ISSN: 1662-5161. DOI: 10.3389/fnhum.2012.00278.
- Pockett, S, S Whalen, AVH McPhail, and WJ Freeman (2007). "Topography, independent component analysis and dipole source analysis of movement related potentials." In: *Cognitive neurodynamics* 1.4, pp. 327–40. ISSN: 1871-4080. DOI: 10.1007/s11571-007-9024-y.
- Polich, J (2007). "Updating P300: an integrative theory of P3a and P3b." In: *Clinical neurophysiology : official journal of the International Federation*

- of Clinical Neurophysiology* 118.10, pp. 2128–48. ISSN: 1388-2457. DOI: 10.1016/j.clinph.2007.04.019.
- Pöppel, E, R Held, and D Frost (1973). “Residual visual function after brain wounds involving the central visual pathways in man”. In: *Nature* 243.5405, pp. 295–296. ISSN: 0028-0836. DOI: 10.1038/243295a0.
- Posner, MI and Y Cohen (1980). “Attention and the control of movements.” In: *Tutorials in motor behavior*, pp. 243–258. DOI: 10.1016/s0166-4115(08)61949-4.
- Potter, MC (1976). “Short-Term Conceptual Memory for Pictures”. In: *Journal of Experimental Psychology: Human Learning and Memory* 2.5, pp. 509–522. ISSN: 0096-1515. DOI: 10.1037/0278-7393.2.5.509.
- Proctor, RW and DW Schneider (2017). “Hick’s Law for Choice Reaction Time: A Review”. In: *The Quarterly Journal of Experimental Psychology* 0218.April, pp. 1–56. ISSN: 1747-0218. DOI: 10.1080/17470218.2017.1322622.
- Purves, D, RB Lotto, SM Williams, S Nundy, and Z Yang (2001). “Why we see things the way we do: evidence for a wholly empirical strategy of vision”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 356.1407, pp. 285–297. DOI: 10.1098/rstb.2000.0772.
- Quené, H and H Van Den Bergh (2004). “On multi-level modeling of data from repeated measures designs: A tutorial”. In: *Speech Communication* 43.1-2, pp. 103–121. ISSN: 01676393. DOI: 10.1016/j.specom.2004.02.004.
- Quigley, C, S Onat, S Harding, M Cooke, and P König (2008). “Audio-visual integration during overt visual attention”. In: 1.2, p. 4. ISSN: 1995-8692.
- Quiroga, RQ, L Reddy, G Kreiman, C Koch, and I Fried (2005). “Invariant visual representation by single neurons in the human brain”. In: *Nature* 435.7045, p. 1102. DOI: 10.1038/nature03687.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <http://www.r-project.org/>.
- RA Fisher, M (1922). “On the mathematical foundations of theoretical statistics”. In: *Phil. Trans. R. Soc. Lond. A* 222.594-604, pp. 309–368.
- Rahnev, D (2017). “The case against full probability distributions in perceptual decision making”. In: *bioRxiv*, p. 108944. DOI: 10.1101/108944.
- Rahnev, D and RN Denison (2018). “Suboptimality in Perceptual Decision Making.” In: *The Behavioral and brain sciences*, pp. 1–107. DOI: 10.1101/060194.

- Rao, RP and DH Ballard (1999). "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects." In: *Nature neuroscience* 2.1, pp. 79–87. ISSN: 1097-6256. DOI: 10.1038/4580.
- Ratcliff, R (1993). *Methods for dealing with reaction time outliers*. DOI: 10.1037/0033-2909.114.3.510. arXiv: 93/3.00 [0033-2909].
- Ratcliff, R (2001). "Diffusion and random walk processes". In: *International encyclopedia of the social and behavioral sciences* 6.1977, pp. 3668–3673.
- Ravden, D and J Polich (1998). "Habituation of P300 from visual stimuli." In: *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 30.3, pp. 359–65. ISSN: 0167-8760. DOI: 10.1016/S0167-8760(98)00039-7.
- Ravden, D and J Polich (1999). "On P300 measurement stability: habituation, intra-trial block variation, and ultradian rhythms." In: *Biological psychology* 51.1, pp. 59–76. ISSN: 0301-0511. DOI: 10.1016/S0301-0511(99)00015-0.
- Rayner, K (1978). "Eye movements in reading and information processing." In: *Psychological Bulletin* 85.3, pp. 618–660. ISSN: 0033-2909. DOI: 10.1037/0033-2909.85.3.618.
- Rayner, K (1998). "Eye Movements in Reading and Information Processing : 20 Years of Research". In: 124.3, pp. 372–422. DOI: 10.1037/0033-2909.124.3.372.
- Rayner, K (2009). "Eye movements and attention in reading, scene perception, and visual search". In: *The Quarterly Journal of Experimental Psychology* 62.8, pp. 1457–1506. ISSN: 1747-0218. DOI: 10.1080/17470210902816461.
- Rayner, K, TJ Smith, GL Malcolm, and JM Henderson (2009). "Eye movements and visual encoding during scene perception". In: *Psychological Science* 20.1, pp. 6–10. ISSN: 09567976. DOI: 10.1111/j.1467-9280.2008.02243.x.
- Rayner, K, a W Inhoff, RE Morrison, ML Slowiaczek, and JH Bertera (1981). "Masking of foveal and parafoveal vision during eye fixations in reading." In: *Journal of experimental psychology. Human perception and performance* 7.1, pp. 167–179. ISSN: 0096-1523. DOI: 10.1037/0096-1523.7.1.167.
- Reed, SK and MP Friedman (1973). "Perceptual vs conceptual categorization". In: *Memory & Cognition* 1.2, pp. 157–163. ISSN: 0090-502X. DOI: 10.3758/BF03198087.
- Reichle, ED, A Pollatsek, DL Fisher, and K Rayner (1998). "Toward a model of eye movement control in reading." In: *Psychological Review* 105.1, pp. 125–157. ISSN: 0033-295X. DOI: 10.1037/0033-295X.105.1.125.

- Reingold, EM and DM Stampe (2002). "Saccadic inhibition in voluntary and reflexive saccades." In: *Journal of cognitive neuroscience* 14.3, pp. 371–388. ISSN: 0898-929X. DOI: 10.1162/089892902317361903.
- Richter, T (2006). "What Is Wrong With ANOVA and Multiple Regression? Analyzing Sentence Reading Times With Hierarchical Linear Models". In: *Discourse Processes* 41.3, pp. 221–250. ISSN: 0163-853X. DOI: 10.1207/s15326950dp4103_1.
- Riesenhuber, M and T Poggio (2002). "How Visual Cortex Recognizes Objects: The Tale of the Standard Model". In: *The visual neurosciences*. Ed. by JSW (L. M. Chalupa. Cambridge, MA: MIT Press, pp. 1640–1653.
- Rock, I and J Victor (1964). "Vision and Touch: An Experimentally Created Conflict between the Two Senses". In: *Science* 143.3606, pp. 594–596. ISSN: 0036-8075. DOI: 10.1126/science.143.3606.594.
- Roelfsema, PR, VA Lamme, H Spekreijse, and H Bosch (2002). "Figureground segregation in a recurrent network architecture". In: *Journal of cognitive neuroscience* 14.4, pp. 525–537. DOI: 10.1162/08989290260045756.
- Roelfsema, PR and FP de Lange (2016). "Early visual cortex as a multiscale cognitive blackboard". In: *Annual review of vision science* 2, pp. 131–151.
- Roelfsema, PR and H Spekreijse (2001). "The representation of erroneously perceived stimuli in the primary visual cortex". In: *Neuron* 31.5, pp. 853–863.
- Rolfs, M, D Jonikaitis, H Deubel, and P Cavanagh (2011). "Predictive remapping of attention across eye movements." In: *Nature neuroscience* 14.2, pp. 252–6. ISSN: 1546-1726. DOI: 10.1038/nn.2711.
- Rosenfeld, JP, JR Biroshak, MJ Kleschen, and KM Smith (2005). "Subjective and objective probability effects on P300 amplitude revisited". In: *Psychophysiology* 42.3, pp. 356–359.
- Rossion, B and S Caharel (2011). "ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception." In: *Vision Research* 51.12, pp. 1297–311. ISSN: 1878-5646. DOI: 10.1016/j.visres.2011.04.003.
- Rossion, B and C Jacques (2008). "Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170." In: *Neuroimage* 39.4, pp. 1959–79. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2007.10.011.
- Rothkopf, CA, DH Ballard, and MM Hayhoe (2007). "Task and context determine where you look". In: *Journal of Vision* 7.14, pp. 16–20. DOI: 10.1167/7.14.16.

- Rousselet, GA, CM Gaspar, KP Wierzchorek, and CR Pernet (2011). “Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects).” In: *Frontiers in psychology* 2, p. 137. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2011.00137.
- Rovamo, J, V Virsu, P Laurinen, and L Hyvärinen (1982). “Resolution of gratings oriented along and across meridians in peripheral vision.” In: *Investigative Ophthalmology & Visual Science* 23.5, pp. 666–670. ISSN: 1552-5783.
- Roy, JE, M Riesenhuber, T Poggio, and EK Miller (2010). “Prefrontal cortex activity during flexible categorization.” In: *The Journal of Neuroscience* 30.25, pp. 8519–28. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.4837-09.2010.
- Salthouse, TA and CL Ellis (1980). *Determinants of eye-fixation duration*. DOI: 10.2307/1422228.
- Schafer, EW and MM Marcus (1973). “Self-stimulation sensory responses”. In: *Science* 181.4095, pp. 175–177.
- Schall, JD and KG Thompson (1999). “neural selection and control of visually guided eye movements”. In: *Annual Review of Neuroscience* 22.1, pp. 241–259. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.22.1.241.
- Schall, S, C Quigley, S Onat, and P König (2009). “Visual stimulus locking of EEG is modulated by temporal congruency of auditory stimuli.” In: *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale* 198.2-3, pp. 137–51. ISSN: 1432-1106. DOI: 10.1007/s00221-009-1867-5.
- Schielzeth, H and W Forstmeier (2009). “Conclusions beyond support: overconfident estimates in mixed models.” In: *Behavioral ecology : official journal of the International Society for Behavioral Ecology* 20.2, pp. 416–420. ISSN: 1045-2249. DOI: 10.1093/beheco/arn145.
- Schmidt, TM and P Kofuji (2009). “Functional and morphological differences among intrinsically photosensitive retinal ganglion cells.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29.2, pp. 476–82. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.4117-08.2009.
- Scholl, C, X Jiang, J Martin, and M Riesenhuber (2014). “Time Course of Shape and Category Selectivity Revealed by EEG Rapid Adaptation”. In: *Journal of Cognitive Neuroscience* 26.2, pp. 408–421. DOI: 10.1162/jocn.
- Schurger, A, JD Sitt, and S Dehaene (2012). “An accumulator model for spontaneous neural activity prior to self-initiated movement.” In: *Proceedings of the National Academy of Sciences of the United States of America* 109.42, E2904–13. ISSN: 1091-6490. DOI: 10.1073/pnas.1210467109.
- Schütt, HH, LOM Rothkegel, HA Trukenbrod, S Reich, FA Wichmann, and R Engbert (2017). “Likelihood-based parameter estimation and comparison of

- dynamical cognitive models.” In: *Psychological Review* 124.4, pp. 505–524. ISSN: 1939-1471. DOI: 10.1037/rev0000068.
- Segaert, K, K Weber, FP de Lange, KM Petersson, and P Hagoort (2013). *The suppression of repetition enhancement: A review of fMRI studies*. DOI: 10.1016/j.neuropsychologia.2012.11.006.
- Seger, CA and EK Miller (2010). “Category learning in the brain.” In: *Annual review of neuroscience* 33, pp. 203–19. ISSN: 1545-4126. DOI: 10.1146/annurev.neuro.051508.135546.
- Serre, T, A Oliva, and T Poggio (2007). “A feedforward architecture accounts for rapid categorization.” In: *Proceedings of the National Academy of Sciences of the United States of America* 104.15, pp. 6424–9. ISSN: 0027-8424. DOI: 10.1073/pnas.0700622104.
- Seth, AK (2014). “A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia”. In: *Cognitive neuroscience* 5.2, pp. 97–118.
- Shadlen, MN, K Britten, W Newsome, and J Movshon (1996). “A computational analysis of the relationship between neuronal and behavioral responses to visual motion”. In: *Journal of Neuroscience* 16.4, pp. 1486–1510.
- Shadlen, MN and WT Newsome (2001). “Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey”. In: *Journal of Neurophysiology* 86.4, pp. 1916–1936. DOI: 10.1152/jn.2001.86.4.1916.
- Shepherd, M, JM Findlay, and RJ Hockey (1986). “The relationship between eye movements and spatial attention”. In: *The Quarterly Journal of Experimental Psychology Section A* 38.3, pp. 475–491.
- Sherman, SM and R Guillery (1998). “On the actions that one nerve cell can have on another: distinguishing drivers from modulators”. In: *Proceedings of the National Academy of Sciences* 95.12, pp. 7121–7126. DOI: 10.1073/pnas.95.12.7121.
- Sigala, N, F Gabbiani, and NK Logothetis (2002). “Visual categorization and object representation in monkeys and humans.” In: *Journal of Cognitive Neuroscience* 14.2, pp. 187–98. ISSN: 0898-929X. DOI: 10.1162/089892902317236830.
- Sigala, N and NK Logothetis (2002a). “Visual categorization shapes feature selectivity in the primate temporal cortex”. In: *Nature* 415.6869, p. 318.
- Sigala, N and NK Logothetis (2002b). “Visual categorization shapes feature selectivity in the primate temporal cortex.” In: *Nature* 415.6869, pp. 318–20. ISSN: 0028-0836. DOI: 10.1038/415318a.

- Smith, EE and M Grossman (2008). “Multiple systems of category learning”. In: *Neuroscience & Biobehavioral Reviews* 32.2, pp. 249–264.
- Smith, PL and R Ratcliff (2004). “Psychology and neurobiology of simple decisions”. In: *Trends in Neurosciences* 27.3, pp. 161–168. ISSN: 01662236. DOI: 10.1016/j.tins.2004.01.006.
- Smith, SM and TE Nichols (2009). “Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference.” In: *Neuroimage* 44.1, pp. 83–98. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2008.03.061.
- Sommer, W, H Leuthold, and J Matt (1998). “The expectancies that govern the P300 amplitude are mostly automatic and unconscious”. In: *Behavioral and Brain Sciences* 21.1, pp. 149–150.
- Sorensen, T, S Hohenstein, and S Vasishth (2016). “Bayesian Linear Mixed Models using Stan: A tutorial for psychologists, linguists, and cognitive scientists”. In: *The Quantitative Methods for Psychology* 12.3, pp. 175–200. DOI: 10.20982/tqmp.12.3.p175.
- Spoida, K, D Eickelbeck, R Karapinar, T Eckardt, D Jancke, BV Ehinger, P König, D Dalkara, S Herlitze, and OA Masseck (2016). “Melanopsin variants as intrinsic optogenetic on and off switches for transient versus sustained activation of G protein pathways”. In: *Current Biology*. DOI: 10.1016/j.cub.2016.03.007.
- Srinivasan, M, S Laughlin, and A Dubs (1982). “Predictive coding: a fresh view of inhibition in the retina”. In: *Proceedings of the Royal Society of London* 216.4205, pp. 427–459.
- Stan Development Team (2016). *rstanarm: Bayesian applied regression modeling via Stan*. R package version 2.13.1. URL: <http://mc-stan.org/>.
- Stefanics, G, M Kimura, and I Czigler (2011). “Visual mismatch negativity reveals automatic detection of sequential regularity violation.” In: *Frontiers in human neuroscience* 5.May, p. 46. ISSN: 1662-5161. DOI: 10.3389/fnhum.2011.00046.
- Stodden, V, J Seiler, and Z Ma (2018). “An empirical analysis of journal policy effectiveness for computational reproducibility”. In: *Proceedings of the National Academy of Sciences* 115.11, pp. 2584–2589.
- Stolk, A, A Todorovic, JM Schoffelen, and R Oostenveld (2013). “Online and offline tools for head movement compensation in MEG.” In: *Neuroimage* 68, pp. 39–48. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2012.11.047.

- Sugase, Y, S Yamane, S Ueno, and K Kawano (1999). "Global and fine information coded by single neurons in the temporal visual cortex." In: *Nature* 400.6747, pp. 869–73. ISSN: 0028-0836. DOI: 10.1038/23703.
- Summerfield, C and FP de Lange (2014). "Expectation in perceptual decision making: neural and computational mechanisms". In: *Nature Reviews Neuroscience* October. ISSN: 1471-003X. DOI: 10.1038/nrn3838.
- Summerfield, C, EH Trittschuh, JM Monti, MM Mesulam, and T Egner (2008). "Neural repetition suppression reflects fulfilled perceptual expectations." In: *Nature neuroscience* 11.9, pp. 1004–6. ISSN: 1546-1726. DOI: 10.1038/nn.2163.
- Swaminathan, SK and DJ Freedman (2012). "Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex." In: *Nature neuroscience* 15.2, pp. 315–20. ISSN: 1546-1726. DOI: 10.1038/nn.3016.
- Szinte, M, D Jonikaitis, D Rangelov, and H Deubel (2018). "Pre-saccadic remapping relies on dynamics of spatial attention". In: *bioRxiv*. DOI: 10.1101/293886.
- Tadel, F, S Baillet, JC Mosher, D Pantazis, and RM Leahy (2011). "Brainstorm: a user-friendly application for MEG/EEG analysis." In: *Computational intelligence and neuroscience* 2011, p. 879716. ISSN: 1687-5273. DOI: 10.1155/2011/879716.
- Tanaka, JW and T Curran (2001). "A neural basis for expert object recognition." In: *Psychological science : a journal of the American Psychological Society / APS* 12, pp. 43–47. ISSN: 0956-7976. DOI: 10.1111/1467-9280.00308.
- Tassinari, H, TE Hudson, and MS Landy (2006). "Combining priors and noisy visual cues in a rapid pointing task." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26.40, pp. 10154–10163. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2779-06.2006.
- Tatler, BW, RJ Baddeley, and BT Vincent (2006). "The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task". In: *Vision Research* 46.12, pp. 1857–1862. ISSN: 00426989. DOI: 10.1016/j.visres.2005.12.005.
- Tatler, BW, JR Brockmole, and R Carpenter (2017). "LATEST: A model of saccadic decisions in space and time." In: *Psychological Review* 124.3, pp. 267–300. ISSN: 1939-1471. DOI: 10.1037/rev0000054.
- Tatler, BW and BT Vincent (2009). "The prominence of behavioural biases in eye guidance". en. In: *Visual Cognition* 17.6-7, pp. 1029–1054. ISSN: 1350-6285. DOI: 10.1080/13506280902764539.

- Thickbroom, GW, W Knezevi, WM Carroll, and FL Mastaglia (1991). “Saccade onset and offset lambda waves: relation to pattern movement visually evoked potentials”. In: *Brain Research* 551.1-2, pp. 150–156. ISSN: 00068993. DOI: 10.1016/0006-8993(91)90927-N.
- Thierry, G, CD Martin, P Downing, and AJ Pegna (2007). “Controlling for inter-stimulus perceptual variance abolishes N170 face selectivity.” In: *Nature Neuroscience* 10.4, pp. 505–11. ISSN: 1097-6256. DOI: 10.1038/nn1864.
- Thompson, E and FJ Varela (2001). “Radical embodiment: neural dynamics and consciousness”. In: *Trends in cognitive sciences* 5.10, pp. 418–425.
- Tickle, H, M Speekenbrink, K Tsetsos, E Michael, and C Summerfield (2016). “Near-optimal integration of magnitude in the human parietal cortex.” In: *Journal of cognitive neuroscience* 28.4, pp. 589–603. ISSN: 1530-8898. DOI: 10.1162/jocn_a_00918.
- Todorovic, A, F van Ede, E Maris, and FP de Lange (2011). “Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31.25, pp. 9118–23. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.1425-11.2011.
- Tong, F and SA Engel (2001). “Interocular rivalry revealed in the human cortical blind-spot representation”. In: *Nature* 411.6834, p. 195.
- Tootell, RB, E Switkes, MS Silverman, and SL Hamilton (1988). “Functional anatomy of macaque striate cortex. II. Retinotopic organization”. In: *Journal of neuroscience* 8.5, pp. 1531–1568.
- Trukenbrod, Ha and R Engbert (2014). “ICAT: a computational model for the adaptive control of fixation durations.” In: *Psychonomic bulletin & review* 21.4, pp. 907–34. ISSN: 1531-5320. DOI: 10.3758/s13423-013-0575-0.
- Tsai, AC, TP Jung, VSC Chien, AN Savostyanov, and S Makeig (2014). “Cortical surface alignment in multi-subject spatiotemporal independent EEG source imaging.” In: *Neuroimage* 87, pp. 297–310. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2013.09.045.
- Tsukamoto, H, Y Kubo, DL Farrens, M Koyanagi, A Terakita, and Y Furutani (2015). “Retinal Attachment Instability Is Diversified among Mammalian Melanopsins.” In: *The Journal of biological chemistry* 290.45, pp. 27176–87. ISSN: 1083-351X. DOI: 10.1074/jbc.M115.666305. URL: <http://www.jbc.org/content/290/45/27176.short>.
- Tzelepi, A, Q Yang, and Z Kapoula (2005). “The effect of transcranial magnetic stimulation on the latencies of vertical saccades”. In: *Experimental Brain Research* 164.1, pp. 67–77. DOI: 10.1007/s00221-005-2250-9.

- Underwood, G, JM Henderson, and A Hollingworth (1998). "Eye guidance in reading and scene perception". In: *Eye movements during scene viewing: An overview*. Ed. by G Underwood. Vol. 1. Elsevier. Chap. 12, pp. 269–293. ISBN: 0080433618.
- Updyke, B (1975). "The pattern of projection of cortical areas 17, 18, and 19 onto the laminae of the dorsal lateral geniculate nucleus in the cat". In: *Journal of Comparative Neurology* 163.4, pp. 377–395.
- VanRullen, R (2011). "Four common conceptual fallacies in mapping the time course of recognition." In: *Frontiers in Psychology* 2, p. 365. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2011.00365.
- VanRullen, R, L Reddy, and C Koch (2004). "Visual Search and Dual Tasks Reveal Two Distinct Attentional Resources". In: *Journal of Cognitive Neuroscience* 16.1, pp. 4–14. ISSN: 0898-929X. DOI: 10.1162/089892904322755502.
- Vaughan, J (1982). "Control of fixation duration in visual search and memory search: Another look". In: *Journal of Experimental Psychology: Human Perception and Performance* 8.5, pp. 709–723. ISSN: 0096-1523. DOI: 10.1037/0096-1523.8.5.709. URL: <http://www.ncbi.nlm.nih.gov/pubmed/6218231>.
- Vaughan, J and TM Graefe (1977). "Delay of stimulus presentation after the saccade in visual search". In: 22.2, pp. 201–205. DOI: 10.3758/bf03198755.
- Vickers, D (1970). "Evidence for an Accumulator Model of Psychophysical Discrimination". In: *Ergonomics* 13.1, pp. 37–58. ISSN: 0014-0139. DOI: 10.1080/00140137008931117.
- Vizioli, L, GA Rousselet, and R Caldara (2010). "Neural repetition suppression to identity is abolished by other-race faces." In: *Proceedings of the National Academy of Sciences of the United States of America* 107.46, pp. 20081–6. ISSN: 1091-6490. DOI: 10.1073/pnas.1005751107.
- Wacongne, C, E Labyt, V van Wassenhove, T Bekinschtein, L Naccache, and S Dehaene (2011). "Evidence for a hierarchy of predictions and prediction errors in human cortex." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.51, pp. 20754–9. ISSN: 1091-6490. DOI: 10.1073/pnas.1117807108.
- Wagner, AD, BJ Shannon, I Kahn, and RL Buckner (2005). "Parietal lobe contributions to episodic memory retrieval". In: *Trends in cognitive sciences* 9.9, pp. 445–453.
- Wahn, B and P König (2015a). "Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration."

- English. In: *Frontiers in psychology* 6, p. 1084. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2015.01084.
- Wahn, B and P König (2015b). “Vision and haptics share spatial attentional resources and visuotactile integration is not affected by high attentional load.” In: *Multisensory research* 28.3-4, pp. 371–392. ISSN: 2213-4794. DOI: 10.1163/22134808-00002482.
- Wahn, B and P König (2016). “Attentional resource allocation in visuotactile processing depends on the task, but optimal visuotactile integration does not depend on attentional resources.” In: *Frontiers in integrative neuroscience* 10, p. 13. ISSN: 1662-5145. DOI: 10.3389/fnint.2016.00013.
- Walshe, RC and A Nuthmann (2014). “Asymmetrical control of fixation durations in scene viewing.” In: *Vision research* 100, pp. 38–46. ISSN: 1878-5646. DOI: 10.1016/j.visres.2014.03.012.
- Walshe, RC and A Nuthmann (2015). “Mechanisms of saccadic decision making while encoding naturalistic scenes”. In: *Journal of Vision* 15, pp. 1–38. ISSN: 15347362. DOI: 10.1167/15.5.21.
- Walt, S van der, SC Colbert, and G Varoquaux (2011). “The NumPy Array: A Structure for Efficient Numerical Computation”. In: *Computing in Science & Engineering* 13.2, pp. 22–30. ISSN: 1521-9615. DOI: 10.1109/MCSE.2011.37.
- Warland, DK, P Reinagel, and M Meister (1997). “Decoding Visual Information From a Population of Retinal Ganglion Cells”. In: *Journal of Neurophysiology* 78.5, pp. 2336–2350. ISSN: 0022-3077. DOI: 10.1152/jn.1997.78.5.2336.
- Weber, RB and RB Daroff (1971). “The metrics of horizontal saccadic eye movements in normal humans”. In: *Vision Research* 11.9, 921–IN2. ISSN: 00426989. DOI: 10.1016/0042-6989(71)90212-4.
- Weiss, Y, EP Simoncelli, and EH Adelson (2002). “Motion illusions as optimal percepts”. In: *Nature neuroscience* 5.6, p. 598.
- White, BJ, DJ Berg, JY Kan, RA Marino, L Itti, and DP Munoz (2017). “Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video”. In: *Nature Communications* 8, p. 14263. ISSN: 2041-1723. DOI: 10.1038/ncomms14263.
- Wichmann, F, J Drewes, P Rosas, and K Gegenfurtner (2010). “Animal detection in natural scenes: Critical features revisited”. In: *Journal of Vision* 10.4, pp. 1–27. DOI: 10.1167/10.4.6.Introduction.
- Wickham, H (2007). “Reshaping Data with the {reshape} Package”. In: *Journal of Statistical Software* 21.12, pp. 1–20. DOI: 10.18637/jss.v021.i12.

- Wickham, H (2015). *stringr: Simple, Consistent Wrappers for Common String Operations*.
- Wilcox, RR (2012). *Introduction to robust estimation and hypothesis testing*. DOI: 10.2307/2669876.
- Wilming, N, T Betz, TC Kietzmann, and P König (2011). “Measures and Limits of Models of Fixation Selection”. In: *PLOS ONE* 6.9. Ed. by T Wennekers, e24038. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0024038.
- Wilming, N, S Harst, N Schmidt, and P König (2013). “Saccadic Momentum and Facilitation of Return Saccades Contribute to an Optimal Foraging Strategy”. In: *PLoS Computational Biology* 9.1. Ed. by O Sporns, e1002871. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002871.
- Wilming, N, S Onat, JP Ossandón, A Açk, TC Kietzmann, K Kaspar, RR Gameiro, A Vormberg, and P König (2017). “An extensive dataset of eye movements during viewing of complex images.” In: *Scientific data* 4, p. 160126. ISSN: 2052-4463. DOI: 10.1038/sdata.2016.126.
- Wolf, E and A Morandi (1962). “Retinal sensitivity in the region of the blind spot”. In: *JOSA* 52.11, pp. 806–812. DOI: 10.1364/josa.52.000806.
- Wolpert, D, Z Ghahramani, and M Jordan (1995). “An internal model for sensorimotor integration”. In: *Science* 269, pp. 1880–1882. DOI: 10.1126/science.7569931.
- Wu, CC and E Kowler (2013). “Timing of saccadic eye movements during visual search for multiple targets”. In: *Journal of Vision* 13.11, pp. 1–21. ISSN: 15347362. DOI: 10.1167/13.11.11.doi.
- Wyart, V, V de Gardelle, J Scholl, and C Summerfield (2012). “Rhythmic fluctuations in evidence accumulation during decision making in the human brain.” In: *Neuron* 76.4, pp. 847–58. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.09.015.
- Zago, L, MJ Fenske, E Aminoff, and M Bar (2005). “The rise and fall of priming: how visual exposure shapes cortical representations of objects.” In: *Cerebral cortex (New York, N.Y. : 1991)* 15.11, pp. 1655–65. ISSN: 1047-3211. DOI: 10.1093/cercor/bhi060.
- Zimmer, M and G Kovács (2011). “Position specificity of adaptation-related face aftereffects.” In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 366.1564, pp. 586–95. ISSN: 1471-2970. DOI: 10.1098/rstb.2010.0265.

Appendix: Thesis Art

Throughout my PhD thesis, I got to work with many talented students. They worked very hard to earn their degree, but their final theses stayed hidden in a file drawer - in most cases never to be looked at again. I wanted to change this and decided to transform each thesis into a unique piece of art, which every student (and I myself) can hang on the walls in their homes. Making this "thesis-art" interesting enough will query friends and colleagues to understand what is going on and spark discussions. Two other benefits are apparent as well: I perform a type of public outreach and I can properly thank my students for their hard work.

In the following I will shortly comment on the six pieces for six students.

Katja Häusser

Katja wrote her Bachelor's Thesis with me. She worked in the project described in Chapter 2 and is a coauthor in Ehinger, Häusser, et al. (2017). In her thesis-art, the stimulus used in the study is visualized by the words of her thesis. That is, the whole thesis is represented on the poster.

Lilli Kaufhold

Lilli wrote her Master's Thesis with me. She worked on the project described in Chapter 4 and is a coauthor in Ehinger, Kaufhold, et al. (2018). In her thesis-art, I recorded my eye-movements while reading her thesis. These eye-movement patterns are printed as a matrix of multiple multiples, describing all spatial-decisions one made where to look next, while reading the thesis.

Jiameng Wu

Meng wrote her Bachelor's Thesis with me. She worked on the project described in Chapter 2. She was looking at the effect of saliency on the choice time. Therefore, I calculated several saliency features (e.g. Contrast, Brightness) of her thesis and overlaid them. This concept is an artistic approximation to a saliency map.

Edoardo Pinzuti

Edoardo wrote his Master's thesis with me. His work is not described in this thesis. He implemented a matlab toolbox to analyse directed interactions between complex systems. Because his work touches Chaotic Attractors and Takens Theorem, I simulated a dynamic attractor (a Lorenz System) and visualized it using the text from his thesis.

Judith Schepers

Judith wrote her Bachelor's Thesis with me. She worked on the project described in Chapter 4. In her thesis-art, I visualized the

guided-bubble paradigm. Because she generalized the paradigm to more than five bubbles, many more bubbles are visible in the thesis-art.

Maria Sokotushchenko

Maria wrote her Master's Thesis with me. She worked on a successor project to Chapter 2. In her thesis-art, I artistically visualized the brain's surprise response to a unexpected stimulus change. This response is sorted by how fast subjects responded (late on top, fast in the bottom)

Lisa-Marie Vortmann

Lisa wrote her Master's Thesis with me. She worked on an unrelated project based on SSVEP's and Multiple Object Tracking. In her thesis-art, I show several EEG signals filtered at the respective SSVEP frequencies. Due to the discrete sampling, distinct lines seem to appear - but in reality they do not exist. This is related to Nyquist's Theorem.

PSYCHOLOGICAL STUDY ON THE TEMPORAL AND NASAL VISUAL IDENTIFIERS

INTRODUCTION: The present study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants.

METHODS: The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants.

RESULTS: The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants.

CONCLUSIONS: The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants. The study was designed to investigate the temporal and nasal visual identifiers. The study was conducted in a laboratory setting and involved 20 participants.

Psychomusicology study on the Temporal and Nasal Visual Identifiers



The Influence of Fixation Durations on the Initiation of Saccades A Gaze-Dependent Eye Tracking Study on Bubble Stimuli

Table of Contents

- 1. Introduction
- 2. Literature Review
- 3. Method
- 4. Results
- 5. Discussion
- 6. Conclusion

Chapter 1
Introduction

1.1 Introduction

1.2 Statement of research objectives

1.3 Rationale for the study

1.4 Scope of the study

1.5 Structure of the dissertation

Chapter 2
Literature Review

2.1 Introduction

2.2 Eye movements

2.3 Gaze-contingent viewing

2.4 Exploration-exploitation dilemma

2.5 Multi-armed bandit problem

2.6 Reinforcement learning

2.7 Summary

Chapter 3
Method

3.1 Introduction

3.2 Participants

3.3 Apparatus & Stimuli

3.4 Procedure

3.5 Data Collection

3.6 Summary



Chapter 4
Results

4.1 Introduction

4.2 Results of the 3x3 grid task

4.3 Results of the 2x2 grid task

4.4 Results of the 4x4 grid task

4.5 Summary



Chapter 5
Discussion

5.1 Introduction

5.2 Summary of findings

5.3 Implications for eye movements

5.4 Implications for the exploration-exploitation dilemma

5.5 Summary

Chapter 6
Conclusion

6.1 Introduction

6.2 Summary of findings

6.3 Implications for eye movements

6.4 Implications for the exploration-exploitation dilemma

6.5 Summary

Chapter 7
Bibliography

7.1 Introduction

7.2 List of references

7.3 Summary

Chapter 8
Appendix

8.1 Introduction

8.2 Appendix A: Ethics approval

8.3 Appendix B: Questionnaire

8.4 Appendix C: Raw data

8.5 Appendix D: Summary



Chapter 9
References

9.1 Introduction

9.2 List of references

9.3 Summary



Chapter 10
Appendix

10.1 Introduction

10.2 Appendix A: Ethics approval

10.3 Appendix B: Questionnaire

10.4 Appendix C: Raw data

10.5 Appendix D: Summary

Chapter 11
References

11.1 Introduction

11.2 List of references

11.3 Summary



Chapter 12
Appendix

12.1 Introduction

12.2 Appendix A: Ethics approval

12.3 Appendix B: Questionnaire

12.4 Appendix C: Raw data

12.5 Appendix D: Summary

Chapter 13
References

13.1 Introduction

13.2 List of references

13.3 Summary



Chapter 14
Appendix

14.1 Introduction

14.2 Appendix A: Ethics approval

14.3 Appendix B: Questionnaire

14.4 Appendix C: Raw data

14.5 Appendix D: Summary

Chapter 15
References

15.1 Introduction

15.2 List of references

15.3 Summary



Chapter 16
Appendix

16.1 Introduction

16.2 Appendix A: Ethics approval

16.3 Appendix B: Questionnaire

16.4 Appendix C: Raw data

16.5 Appendix D: Summary

Chapter 17
References

17.1 Introduction

17.2 List of references

17.3 Summary



Chapter 18
Appendix

18.1 Introduction

18.2 Appendix A: Ethics approval

18.3 Appendix B: Questionnaire

18.4 Appendix C: Raw data

18.5 Appendix D: Summary

Chapter 19
References

19.1 Introduction

19.2 List of references

19.3 Summary

Chapter 20
Appendix

20.1 Introduction

20.2 Appendix A: Ethics approval

20.3 Appendix B: Questionnaire

20.4 Appendix C: Raw data

20.5 Appendix D: Summary

Chapter 21
References

21.1 Introduction

21.2 List of references

21.3 Summary



Chapter 22
Appendix

22.1 Introduction

22.2 Appendix A: Ethics approval

22.3 Appendix B: Questionnaire

22.4 Appendix C: Raw data

22.5 Appendix D: Summary

Chapter 23
References

23.1 Introduction

23.2 List of references

23.3 Summary



Chapter 24
Appendix

24.1 Introduction

24.2 Appendix A: Ethics approval

24.3 Appendix B: Questionnaire

24.4 Appendix C: Raw data

24.5 Appendix D: Summary

Chapter 25
References

25.1 Introduction

25.2 List of references

25.3 Summary



Chapter 26
Appendix

26.1 Introduction

26.2 Appendix A: Ethics approval

26.3 Appendix B: Questionnaire

26.4 Appendix C: Raw data

26.5 Appendix D: Summary

Chapter 27
References

27.1 Introduction

27.2 List of references

27.3 Summary



Chapter 28
Appendix

28.1 Introduction

28.2 Appendix A: Ethics approval

28.3 Appendix B: Questionnaire

28.4 Appendix C: Raw data

28.5 Appendix D: Summary

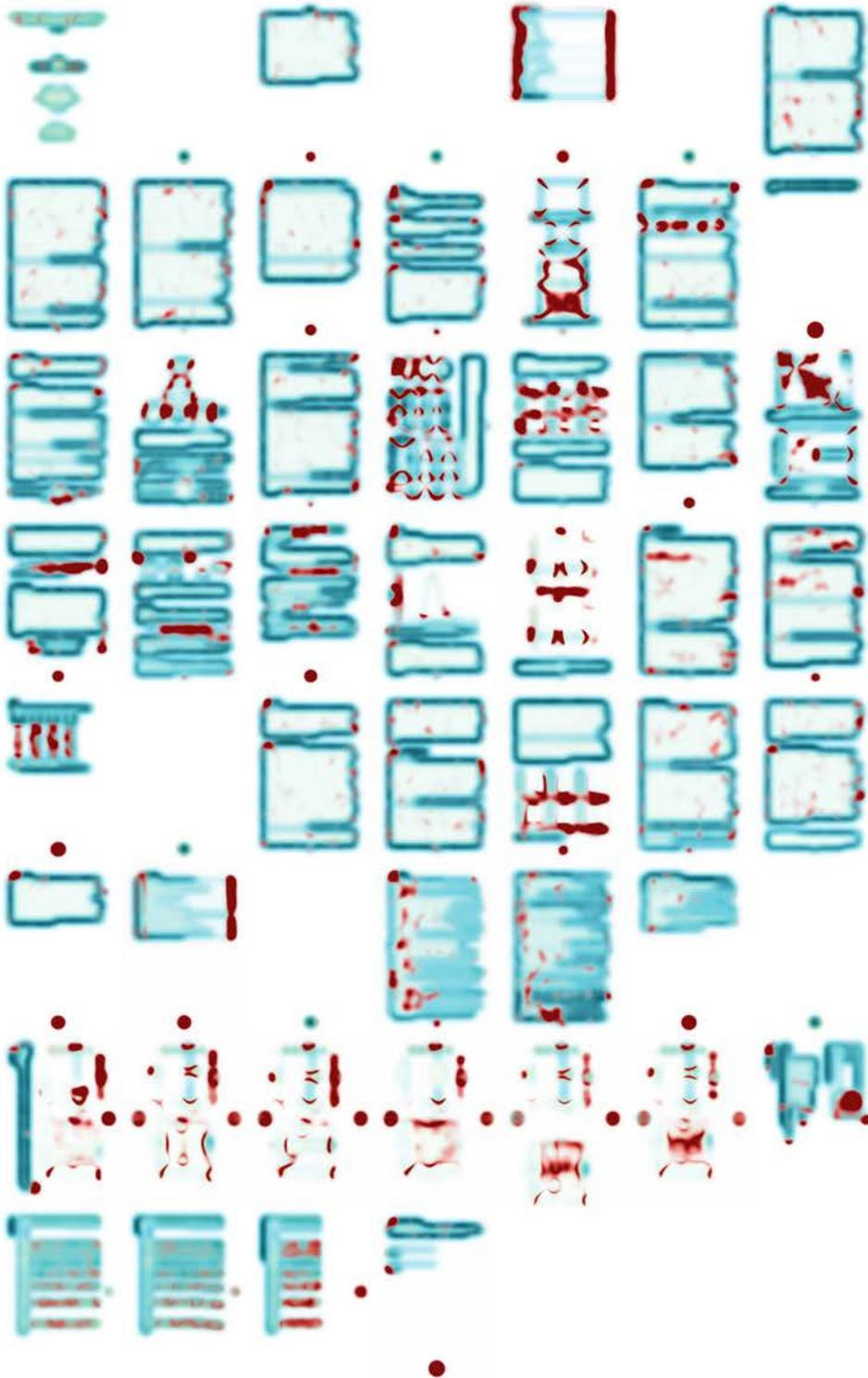
Chapter 29
References

29.1 Introduction

29.2 List of references

29.3 Summary

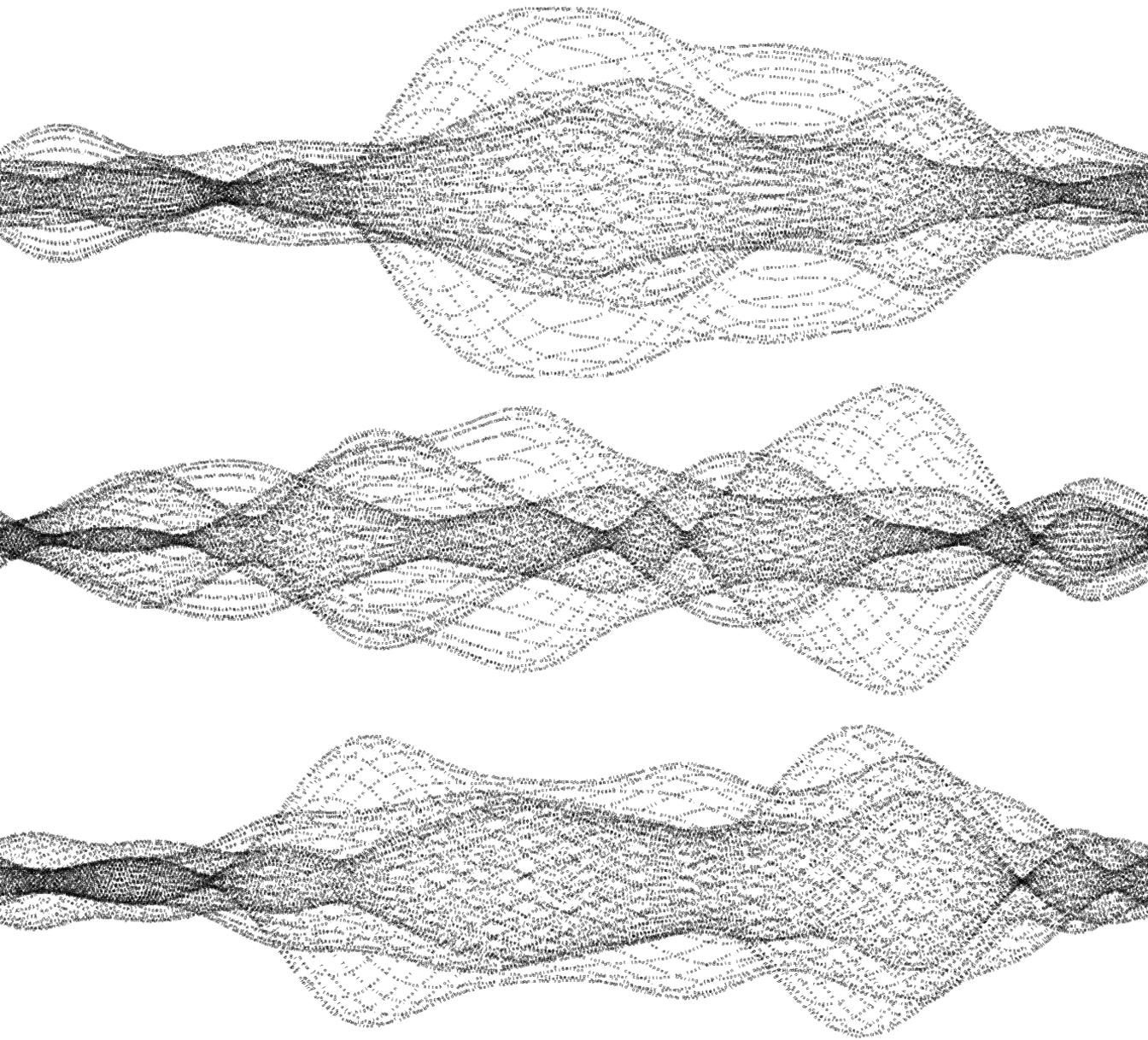
Investigating eye movements as an exploration-exploitation dilemma using a new gaze-contingent viewing task



The Influence of Saliency on the Initiation of Saccades
in a Guided Viewing Paradigm with Bubble-Stimuli

P300 prediction errors in veridical and inferred information processing: an EEG study

Maria Sokotushchenko
March 2018



Unsupervised Decoding of Visuospatial Attention
Performance in a Multiple Object Tracking Task using
EEG Steady-State Analysis

Lisa-Marie Vortmann
May 2018

Acknowledgements

Foremost, I want to thank my favorite colleague *Anna Gert* for her great ideas, steady support and immense patience (also in life). I especially thank Peter König for his outstanding supervision throughout my BSc, MSc and this dissertation. Even though Tim Kietzmann and José Ossandón supervised me in theses long past, I continue to learn and get inspired by their lessons.

Many colleagues influenced my scientific thinking and integrity more than they know: Anne Urai, Arthur Czeszumski, Ashima Keshava, Basil Wahn, Gautam Argawal, Greta Häberle, Holger Finger, Johannes Keyser, Lilian Weber, Niklas Wilming, Olaf Dimigen, Petra Fischer, Raquel London, Selim Onat, Silja Timm, just to name a few.

I am very grateful to be allowed to work with so many talented students. I learned a lot from you all! Thanks to Edoardo Pinzuti, Jiameng Wu, Judith Schepers, Katharina Groß, Katja Häusser, Lena Gschossmann, Lilli Kaufhold, Lisa-Marie Vortmann, Maria Sokotushchenko & Ule Diallo.

I am lucky to have an amazing science family: Thanks to Veronika, Konstantin, Sebastian, Petra & Albert.

