Preface

The human brain is constantly exposed to variable sensory information about the surrounding environment. How does the brain integrate this information to make reliable inferences and predictions as the basis for decision making? If the available information is to be used in an optimal manner, sophisticated statistical methods need to be employed. The question which methods the brain uses to come to decisions and predictions is still unsolved and one of the most exciting questions of neuroscience today (Bach 2014; Lisman 2015).

In the statistical field, Bayesian decision theory combines all available information in an optimal fashion and thus offers a useful theoretical framework for explaining probabilistic inference by humans (Jaynes 2003; Robert 2007). In this framework prior beliefs are constantly updated to posterior beliefs in light of observed data according to Bayes' theorem (Baldi and Itti 2010). Thus, the Bayesian brain hypothesis, which states that the brain codes and computes Bayesian probabilities, has been proposed and is increasingly recognized as providing a framework for investigating cognitive brain functions (Kersten et al. 2004; Knill and Pouget 2004; Friston 2005; Doya et al. 2007; Gold and Shadlen 2007; Kopp 2008; Friston 2010; Bach and Dolan 2012).

Predictive coding theories of cortical functions and the free energy principle instantiate the Bayesian brain hypothesis (Friston 2002, 2010). They are widely applied frameworks for functional neuroimaging and electrophysiological studies of sensory cortical processing (Summerfield et al. 2006; Garrido et al. 2009; Summerfield and Egner 2009; Egner et al. 2010; Rauss et al. 2011; Winkler and Czigler 2011; Lieder et al. 2013). Put simply, these theories state that the brain tries to minimize any "surprise" or prediction error about sensory input. Specifically, predictive coding theories propose that the brain maintains an internal model of the world which it updates in dependence on the surprise about a current stimulus, in order to minimize the surprise about future stimuli (Friston 2002; Friston 2005; Spratling 2010). While earlier research provided results that are consistent with the Bayesian brain hypothesis (Hampton et al. 2006; Ostwald et al. 2012; Vilares et al. 2012; Lieder et al. 2013), definite empirical support is surprisingly scarce, and no unchallenged conclusion about the utility of the Bayesian brain hypothesis as a theoretical framework for explaining cognitive functions of the brain has been achieved so far (Clark 2013).

This work aims at filling this gap by collecting key experimental data in support of the Bayesian brain hypothesis. Which means are necessary to achieve this goal? First, some window into the brain is needed. The electroencephalogram (EEG), which is the signal of the electrical fields of the brain, was first presented nearly 90 years ago with the potential of providing this window (Berger 1929). Although the EEG data are severely corrupted by noise (Schimmel 1967), they provide signals of neural activity with a high temporal resolution (Makeig et al. 2004), which enables studies of brain signals in direct relation with complex cognitive tasks (da Silva 2013) and makes the EEG a useful tool for brain imaging (Michel and Murray 2012). From all the activity which can be seen in the brain, the so-called event-related potentials (ERPs) are particularly useful for a better understanding of brain functions. They are the “scalp-recorded neural activity that is generated in a given neuroanatomical module when a specific computation operation is performed.” (Luck 2014). This implies that by understanding the amplitude fluctuations of these ERPs, the computations of the brain themselves can be deduced.

Thus, the goal of this work is to better understand the probabilistic reasoning of humans by developing observer models to predict event-related potentials, select the models which best explains the EEG data using Bayesian model selection, and make deductions from the properties of the winning models. Note that this is a meta-Bayesian model-based analysis in the sense that Bayesian model selection is used to choose between Bayes-optimal observer models (Daunizeau et al. 2010; Lieder et al. 2013). Inference about the algorithms employed by the brain is then based on the winning models (Mars et al. 2012). A framework of Bayesian updating and predictive surprise is used for dissociating the functions underlying ERP amplitude fluctuations. Bayesian updating refers to changes in probability distributions given new observations and can be differentiated into Bayesian surprise, which constitutes the changes in beliefs about hidden states, and postdictive surprise, which represents the changes in predictions over observable events. In contrast, predictive surprise equals the surprise about observations under their current probabilities.

In a first step, the P300 component of event-related brain potentials is investigated. The P300 is a *positive* potential that is typically measured at parietal scalp regions in a time interval starting 300 ms after an unforeseeable stimulus is presented, and has long been in the focus of research concerned with the brain’s ability to infer statistical regularities of the environment (Sutton et al. 1965). It reflects the degree of surprise related to the processing of sensory input in a way that “surprising events elicit a large P300 component.” (Donchin 1981). A variant of the well-established oddball task (Ritter and Vaughan 1969) is used to collect the EEG data for developing and testing a digital filtering model (DIF), which fuses properties of the most popular and comprehensive model of P300 amplitude fluctuations with a completely computational model. These were proposed by Squires et al. (1976) and Mars et al. (2008), respectively. While Squires et al.’s model remains descriptive, Mars et al.'s observer simply integrates the sensory data over an infinitely long period of time and cannot explain the well-documented effects of the recent stimulus sequence on the P300 (Squires et al. 1976; Leuthold and Sommer 1993). The model selection results show that P300 amplitude fluctuations are best explained by predictive surprise based on the DIF model, which provides direct evidence for the coding of probability distributions in the human brain. Evidence for the updating of the probability distributions is, however, only implicit.

In the next step, the analyses are extended to enclose a total of four temporally and regionally distinguishable ERPs in order to find neural traces for the actual updating of the probability distributions as well as their presence. These ERPs are the frontocentrally distributed N250 (Hillyard and Picton 1987), the anteriorly distributed P3a, the parietally distributed P3b, and the posteriorly distributed slow wave (SW) (Matsuda and Nittono 2015). The P3a and P3b are dissociable components of the P300 (Polich 2007), while the P3a, P3b, and SW make up the so-called late positive complex (Sutton and Ruchkin 1984; Dien et al. 2004). A variant of the urn-ball task (Phillips and Edwards 1966) is introduced, which was specifically designed to represent Bayes' theorem. A Bayesian observer model is then proposed from which a belief distribution over hidden states and a prediction distribution over observable events are derived.

Additionally, it is investigated whether observer models that incorporate nonlinear probability weighting outperform their versions without weighting when predicting ERP amplitude fluctuations. This nonlinear weighting of probabilities was originally reported by prospect theory, which is a famous theory of economic decision behavior (Kahneman and Tversky 1979; Tversky and Kahneman 1992; Fox and Poldrack 2009). The model selection results show that the ERP components of the late positive complex (P3a, P3b, Slow Wave) provide dissociable measures of Bayesian updating and predictive surprise based on the Bayesian observer, while for the N250 predictive surprise based on the DIF model proved superior. These results indicate that the ERP components reflect distinct neural computations and provide evidence for the coding and computing of probability distributions.

The structure of this work is as follows: Chap. 1 introduces basic principles of ERP research which comprise the data acquisition methods used in this work, signal-to-noise ratio estimation for event-related potentials, and the important concept of circularity in data analyses. It further details the framework of Bayesian updating and predictive surprise and the concept of probability weighting functions. Chapter 2 introduces the parametric empirical Bayes methods and variational free energy used for model estimation and selection. It first motivates the use of these methods before thoroughly detailing their computation. In addition, an example experiment analyzes the performance of these methods for single subjects and group studies in light of the signal-to-noise ratio of the data.

Chapter 3 introduces the oddball task and the models taken from the literature before giving a detailed derivation of the digital filtering (DIF) model. The results are displayed and discussed for conventional ERP analyses as well as model-based trial-by-trial analyses. Next, Chap. 4 introduces the urn-ball task and the Bayesian observer model. Probability weighting functions are applied to the Bayesian observer model and to the DIF model. The discussion of the results for conventional ERP analyses and model-based trial-by-trial analyses concludes this chapter. Finally, Chap. 5 summarizes this work, draws the main conclusions, and closes with an outlook.

http://www.springer.com/978-3-319-32284-1

Computational Modeling of Neural Activities for Statistical Inference
Kolossa, A.
2016, XXIV, 127 p. 42 illus., 20 illus. in color., Hardcover
ISBN: 978-3-319-32284-1