

P300 amplitude variations, prior probabilities, and likelihoods: A Bayesian ERP study

Bruno Kopp¹ · Caroline Seer¹ · Florian Lange¹ · Anouck Kluytmans² · Antonio Kolossa³ · Tim Fingscheidt³ · Herbert Hoijtink²

Published online: 12 July 2016
© Psychonomic Society, Inc. 2016

Abstract The capability of the human brain for Bayesian inference was assessed by manipulating probabilistic contingencies in an urn-ball task. Event-related potentials (ERPs) were recorded in response to stimuli that differed in their relative frequency of occurrence (.18 to .82). Averaged ERPs with sufficient signal-to-noise ratio (relative frequency of occurrence > .5) were used for further analysis. Research hypotheses about relationships between probabilistic contingencies and ERP amplitude variations were formalized as (in-)equality constrained hypotheses. Conducting Bayesian model comparisons, we found that manipulations of prior probabilities and likelihoods were associated with separately modifiable and distinct ERP responses. P3a amplitudes were sensitive to the degree of prior certainty such that higher prior probabilities were related to larger frontally distributed P3a waves. P3b amplitudes were sensitive to the degree of likelihood certainty such that lower likelihoods were associated with larger parietally distributed P3b waves. These ERP data suggest that these antecedents of Bayesian inference (prior probabilities and likelihoods) are coded by the human brain.

Keywords Bayesian brain · Urn-ball task · Informative hypotheses · Signal-to-noise ratio (SNR) · P300 · P3b · P3a

Probabilistic inference is a well-established research area in cognitive psychology (see Anderson, 2015, for a textbook introduction; Kahneman, Slovic, & Tversky, 1982). The topic has only recently begun to receive attention in cognitive neuroscience (Bach & Dolan, 2012; Glimcher, 2004). Current research focuses on the neural implementations of probabilistic inference in two main areas. At the cellular level, studies examine neural codes for probabilistic inference (Gerstner & Frenaux, 2015; Kappel, Habenschuss, Legenstein, & Maass, 2015; Pouget, Beck, Ma, & Latham, 2013). At the level of large-scale neural circuits, studies identify neural response variability related to probabilistic inference with high spatial resolution (functional magnetic resonance imaging; fMRI) or with high temporal resolution (event-related potentials; ERP).

Statement of the Problem

Bayesian inference (Jaynes, 2003; Robert, 2007) is the focus of much of this research. Bayes' theorem updates the probability for a hypothesis as more evidence becomes gradually available. Bayes' theorem derives posterior probabilities for the hypothesis as consequences of two antecedents, that is, the prior probabilities for the hypothesis and the likelihoods for the evidence should a hypothesis prove true. Figure 1 provides a geometric visualization of Bayes' theorem. Readers who are unfamiliar with Bayes' theorem are advised to inspect Fig. 1 before proceeding. A more formal introduction to Bayes' theorem can be found below.

According to the Bayesian brain hypothesis, the brain codes and computes Bayesian parameters such as prior

Electronic supplementary material The online version of this article (doi:10.3758/s13415-016-0442-3) contains supplementary material, which is available to authorized users.

✉ Bruno Kopp
kopp.bruno@mh-hannover.de

¹ Department of Neurology, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany

² Department of Methods and Statistics, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands

³ Institute for Communications Technology, Technische Universität Braunschweig, Schleinitzstr. 22, 38106 Braunschweig, Germany

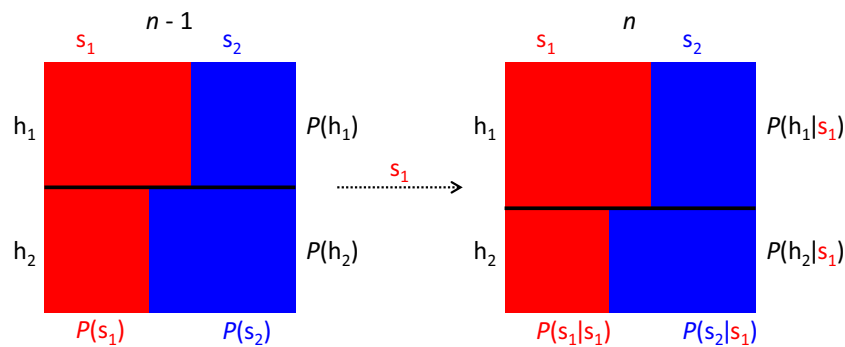


Fig. 1 A geometric illustration of Bayesian inference. In our study, the ‘Mondrian diagram’ on the left-hand side shows probability distributions and likelihoods on trial $n - 1$. In our study, trial $n - 1$ refers to the situation before the first piece of evidence has been received, implying that these parameters were instructed. The ‘Mondrian diagram’ on the right-hand side shows the same parameters on trial n , that is, after a first piece of evidence was observed (i.e., in the example here, an s_1 stimulus, indicated by the dotted black arrow). The receipt of this piece of evidence triggers Bayesian inference. Solid black horizontal lines define prior (left

‘Mondrian diagram’) and posterior (right ‘Mondrian diagram’) probabilities. Relative proportions of colored rectangles along the horizontal axis represent likelihoods. Bayesian inference thus equals shifts of the solid black horizontal line along the vertical axis. In the given example, the black horizontal line was dragged downwards from trial $n - 1$ to trial n , such that $P(h_1|s_1) > P(h_1)$ and $P(h_2|s_1) < P(h_2)$. The degree of belief in hypothesis h_1 being true increased, and accordingly the degree of belief in hypothesis h_2 being true decreased, after obtaining evidence s_1 due to unbalanced likelihoods, that is, $P(s_1|h_1) > P(s_1|h_2)$. (Color figure online)

probabilities and likelihoods (Doya, Ishii, Pouget, & Rao, 2007; Friston, 2005; Knill & Pouget, 2004). Hence, fMRI or ERP discoveries showing that the manipulation of prior probabilities and likelihoods is associated with separately modifiable and distinct neural responses would strengthen the support for the Bayesian brain hypothesis.

Overview of the Existing fMRI Research

We briefly review the literature in chronological order. Summerfield and Koechlin (2008) showed multiple representations of prior expectations within the visual hierarchy in extrastriate and anterior temporal lobe regions with fMRI. Stern, Gonzalez, Welsh, and Taylor (2010) measured fMRI while subjects accumulated sequential pieces of probabilistic evidence and then made a decision under uncertainty. Increased uncertainty during evidence accumulation was associated with activity in dorsal cingulate cortex, whereas greater uncertainty when executing a decision was related to enhanced lateral frontal and parietal activity. These results suggest that neural mechanisms of uncertainty depend on the parameters and on the stage of decision making. Participants in the study by Furl and Averbeck (2011) viewed sequences of bead colors drawn from hidden urns and attempted to infer the majority bead color in each urn. When viewing each bead color, participants chose either to seek more evidence about the urn by drawing another bead (draw choices) or to infer the urn contents (urn choices). When faced with urns that had bead color splits closer to chance (i.e., 60/40 vs. 80/20), participants increased their evidence-seeking behavior. Urn choices evoked larger hemodynamic responses than did draw choices in the insula, striatum, anterior cingulate, and parietal

cortex. These parietal responses were greater for participants who sought more evidence on average and for participants who increased their evidence seeking when draws came from 60/40 urns.

Vilares, Howard, Fernandes, Gottfried, and Körding (2012) designed a sensorimotor decision-making task in which human subjects had to estimate positions of hidden visual targets. Uncertainty (a parameter that is also referred to as precision in other publications) was systematically varied in a two-by-two factorial design, with the factors prior and likelihood uncertainty, such that the conditions were matched for performance accuracy but differed in prior and likelihood uncertainty. Their fMRI study revealed that prior and likelihood uncertainty had distinct neural representations. Whereas likelihood uncertainty activated brain regions along the visuomotor pathway, representations of prior uncertainty were identified in brain areas outside this pathway, including putamen, amygdala, insula, and orbitofrontal cortex. O’Reilly et al. (2013) separately modified surprise, that merely reflects the unexpectedness of an observation and shifts in beliefs or Bayesian updating (see Fig. 1), in a saccadic planning task. They observed that distinct brain regions are associated with surprise and Bayesian updating. Whereas surprise was associated with a hemodynamic response in the posterior parietal cortex, the anterior cingulate cortex was activated during Bayesian updating.

D’Acromont, Schultz, and Bossaerts (2013) asked their subjects to estimate the probability of uncertain stimuli based on prior knowledge as well as on repeated empirical observations. They found hemodynamic correlates of prior knowledge in the striatum, whereas areas in the default mode network (angular gyri, posterior cingulate cortex, and medial prefrontal cortex) encoded stimulus frequencies. Schwartenbeck,

FitzGerald, and Dolan (2016) developed a novel paradigm that allowed decomposing surprise and Bayesian updating (akin to O'Reilly et al., 2013). Using fMRI, they showed that dopamine-rich midbrain regions encode Bayesian updating, whereas surprise is encoded in prefrontal regions, including the presupplementary motor area and dorsal cingulate cortex. In the context of the dopaminergic system, another line of research has shown that this system conveys stochastic parameters of upcoming rewards, such as reward uncertainty (Fiorillo, Tobler, & Schultz, 2003; Preusschoff, Bossaerts, & Quartz, 2006; Preusschoff, Quartz, & Bossaerts, 2008). Trapp, Lepsien, Kotz, and Bar (2016) addressed the neural underpinnings of prior probability in category-specific areas (i.e., the fusiform face area and the parahippocampal place area). Their findings showed that hemodynamic activity in the fusiform face area was higher when faces had higher prior probability.

The available fMRI research shows that various probabilistic parameters may in fact be associated with separately modifiable and anatomically distinct hemodynamic responses. With regard to the exact localization of these responses, the data are less consistent, putatively because of the enormous heterogeneity of the tasks considered. The available data seem to converge in fronto-striatal representations of prior parameters and predominantly parietal representations of likelihood parameters, although there are exceptions from this rule. Furthermore, these parameters are not necessarily equivalent to proper prior probabilities and likelihoods because a majority of tasks that were used in these studies must be considered as being only indirectly related to Bayes' theorem.

Overview of the Existing ERP Research

The relevant ERP research is very much focused on the P300 component (see Kopp, 2008; Polich, 2007, for an overview). Donchin (1981) argued that P300 amplitude is not crucially determined by the inherent attributes of eliciting events. He ascertained that “surprising events elicit a large P300 component” (p. 498). Squires, Wickens, Squires, and Donchin, (1976) had presented a model of trial-by-trial P300 (which is also referred to as P3b, i.e., the parietally distributed P300) amplitude variations, based on the concept of expectancy, which was thought to be determined by multiple factors, most notably by the memory for event frequency and by the “global” event probability. Mars et al. (2008) indeed found that surprise that merely reflects the unexpectedness of an observation was encoded in the parietally distributed P3b (see also Seer, Lange, Boos, Dengler, & Kopp, 2016).

Achtziger, Alós-Ferrer, Hügelschäfer, and Steinhauser (2014) measured ERPs during performance of a belief-updating task and found evidence that participants who overweight new information display a lower sensibility to conflict detection, captured by the N2 (see Folstein & Van Petten,

2008, for overview), whereas participants who overweight prior probabilities initiated action choice well before new information was presented (captured by lateralized readiness potentials; see Eimer, 1998, for an overview). Ostwald et al. (2012) found evidence for spatiotemporal and functional segregation in human somatosensory processing. Their data suggest that early processing stages (somatosensory cortex) may implement short-term Bayesian perceptual learning, intermediate processing stages (right inferior frontal cortex) index active stimulus engagement, and late processing (cingulate cortex) may reflect learning-induced updating of top-down attentional control mechanisms (as expressed in scalp-recorded P3b responses).

Kolossa, Fingscheidt, Wessel, and Kopp (2013) modeled the surprise-sensitivity of trial-by-trial P3b amplitude variations as originating from the operation of multiple digital filters in the brain (i.e., two parallel first-order infinite impulse response low-pass filters and an additional fourth-order finite impulse response high-pass filter). Kolossa, Kopp, and Fingscheidt (2015) replicated the surprise-sensitivity of P3b amplitude variations and showed that the anteriorly distributed P3a encodes Bayesian updating, a finding that was also reported by Bennett, Murawski, and Bode (2015).

The available ERP research shows in a remarkably consistent pattern that various probabilistic parameters may in fact be associated with separately modifiable and functionally distinct electrophysiological responses. Multiple studies showed that amplitude variations in the frontally distributed P3a are related to Bayesian updating, hence to changes in prior probabilities as a result of the gradual receipt of new evidence (as illustrated in Fig. 1). In contrast, amplitude variations in the parietally distributed P3b are related to surprise that reflects the unexpectedness of the evidence.

The Current Study

We sought to bring additional insight into how the human brain produces Bayesian inference by searching for potential ERP correlates of Bayesian parameters. Our study aims to contribute to the existing literature in the following ways. First, previous ERP studies were mainly interested in revealing neural mechanisms underlying the dynamics of Bayesian updating and less attention was given to the neural correlates of the two antecedents of Bayesian inference (i.e., the prior probabilities for the hypothesis and the likelihoods for the evidence should a hypothesis prove true). Second, our study examined potential ERP correlates of preexperiential forms of *instructed* Bayesian parameters (FitzGerald, Seymour, Bach, & Dolan, 2010; Hertwig, Barron, Weber, & Erev, 2004; Hertwig & Erev, 2009). Third, we also advocate an explicitly Bayesian approach to statistical inference in place of classical methods. Bayesian inferential approaches have been proposed

as early as 1985 for psychophysiological research (O'Connor, 1985), but have not gained importance in the relevant literature until recently (see Maris, 2012, for an overview; Constantino & Daw, 2015).

We will formalize our research hypotheses (see [Derivation of Eight Hypotheses of Interest](#)) as so-called informative—that is, (in-)equality constrained—hypotheses (Hojtink, 2012) that are quite distinct from traditional null and alternative hypotheses. Subsequently, support in the data (in the form of Bayes factors and posterior model probabilities) for each of these hypotheses is computed. This is a purely confirmatory inferential approach in which the researcher's degrees of freedom are limited to the specification of the informative hypotheses of interest. Hence, we also try to demonstrate the utility of applying a Bayesian statistics approach to ERP research.

The Urn-Ball Task

Figure 2 shows our experimental paradigm, which employed an urns-and-balls (or urn-ball) task (Phillips & Edwards, 1966). Samples of balls were drawn from two types of urns. We manipulated probabilistic contingencies at two levels. First, we introduced various degrees of uncertainty (Bach & Dolan, 2012) about the type of urn by sampling urns from two probability distributions with different uncertainty parameters. Second, we manipulated the proportion of ball colors within the two types of urns at different levels of uncertainty. We refer to these probabilistic contingencies as prior probabilities and likelihoods, respectively. Figure 2b shows that the urn-ball task requested binary decisions about the colors of all balls across the entire sequence of four draws of balls. After four draws of balls, subjects had to infer the origin of these samples of balls (i.e., they were forced to choose the type of urn from which they believed these four balls had actually been sampled).

Non-Bayesian Interpretation of the Urn-Ball Task

The non-Bayesian analysis of the urn-ball task rests upon the sampling of balls from the population of balls. As shown in Fig. 2a, a non-Bayesian could ignore the existence of the urns in the tableau, and he or she could simply count or estimate the overall numbers of red and blue balls in order to predict the result of the first draws (see Fig. 2b). These quantities correspond to the marginal probabilities $P(s_1)$ and $P(s_2)$ in Fig. 1. The reader who wants to know more about these marginal probabilities in the urn-ball task is referred to Kolossa et al. (2015). It will suffice to point out that these marginal probabilities correspond to the unexpectedness of the evidence (i.e., surprise).

Bayesian Formulation of the Urn-Ball Task

As illustrated in Fig. 1, the Bayesian analysis of the urn-ball task evaluates the probability of a hypothesis (such as that a particular sample of balls was drawn from one type of urn). It rests upon the idea that these probabilities are updated in the light of new evidence, and Bayes' theorem prescribes that posterior probabilities are proportional to the product of likelihoods (distribution of ball colors, given one urn) and prior probabilities (distribution of urns). Given prior probabilities $P(h_1) = 1 - P(h_2)$ for two hypothetical urns $h \in \{h_1, h_2\}$ and the availability of one piece of evidence or a stimulus (i.e., an observed ball) $s \in \{s_1, s_2\}$, we can compute the posterior probabilities as

$$P(h_1|s) = \frac{P(s|h_1) \cdot P(h_1)}{P(s)} \quad (1)$$

$$P(h_2|s) = \frac{P(s|h_2) \cdot P(h_2)}{P(s)}. \quad (2)$$

and

Formulating these posterior probabilities in the log-odds form (Jaynes, 2003), using the notation $\text{Lo}(h_1|s) = \ln(P(h_1|s)/P(h_2|s))$ for the posterior log-odds that favor h_1 , given evidence s , we obtain

$$\text{Lo}(h_1|s) = \ln \frac{P(s|h_1) \cdot P(h_1)}{P(s|h_2) \cdot P(h_2)} = \ln \frac{P(s|h_1)}{P(s|h_2)} + \ln \frac{P(h_1)}{P(h_2)}. \quad (3)$$

Hence, Bayesian probability updating (in log-odds form and for binary decisions) simply equals adding the logarithm of the likelihood ratio, that is, $\ln(P(s|h_1)/P(s|h_2))$, to the logarithm of the prior odds, that is, $\ln(P(h_1)/P(h_2))$.

Derivation of Predictions for ERP Amplitudes

We recorded 20-channel electroencephalographic data from human participants who performed the urn-ball task. We deliberately restricted our analyses to cortical responses to just the initial items of the samples of balls. Figure 2B illustrates that this confinement served to identify potential ERP correlates of the unexpectedness of the evidence, prior probabilities, and likelihoods as they were *instructed* via the tableaux in our urn-ball task (FitzGerald et al., 2010; Hertwig et al., 2004; Hertwig & Erev, 2009). We measured cortical responses to

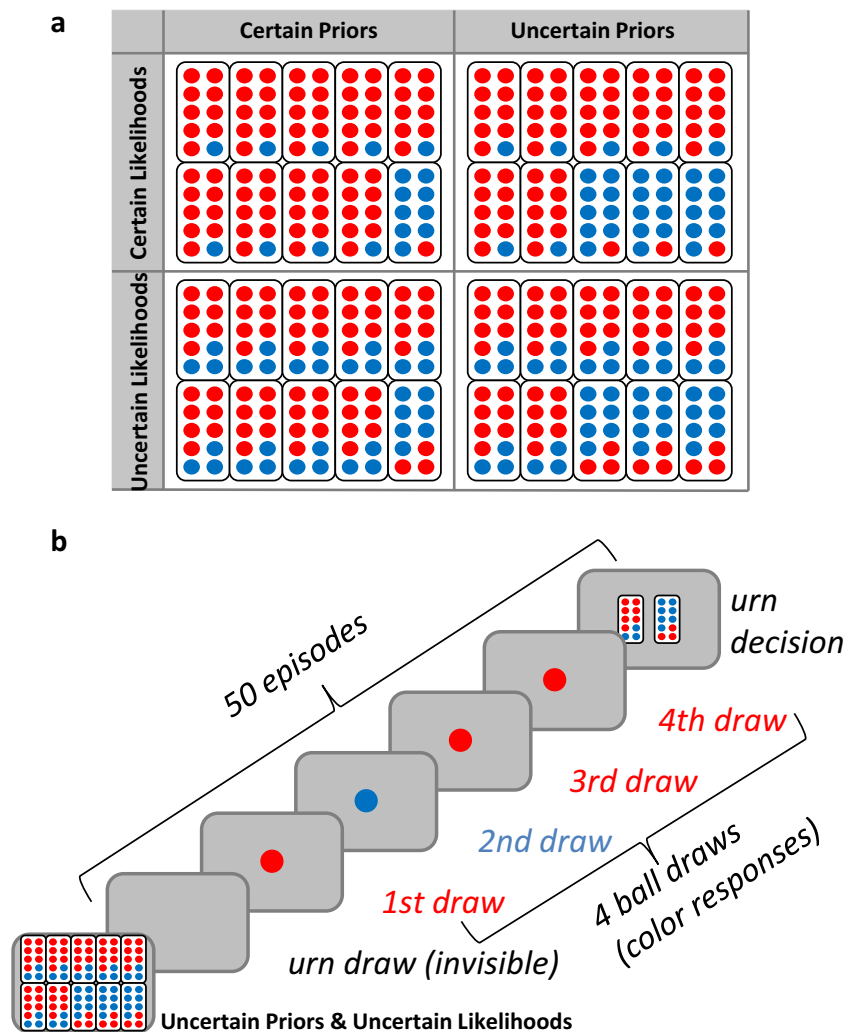


Fig. 2 (a). The four tableaux that were shown to participants in the four probability conditions of the urn-ball task—certain priors (Pc: $P=0.9$, $P=0.1$), uncertain priors (Pu: $P=0.7$, $P=0.3$), certain likelihoods (Lc: $P=0.9$, $P=0.1$), uncertain likelihoods (Lu: $P=0.7$, $P=0.3$); see text for details. (b). Each of the four probability conditions (here, uncertain priors & uncertain likelihoods condition, PuLu, serves as an example)

comprised 50 consecutive episodes of sampling that consisted of the invisible drawing of one urn, the visible drawing of a sample of four balls from that urn (sequential drawing with replacement), and a forced choice on the type of urn from which this sample might have originated (Color figure online)

simple colored stimuli that were processed for probabilistic inference at a later time. We aimed to test whether instructed unexpectedness of the evidence, prior probabilities, and likelihoods separately modify distinct neural responses.

We hypothesized that the unexpectedness of the evidence, prior probabilities, and likelihoods should reveal itself as amplitude variations in P300 waves. We expected that P3a amplitudes are modified by the manipulation of prior probabilities and that P3b amplitudes are modified by the manipulation of the unexpectedness of the evidence and/or likelihoods. The P3a prediction is related to empirical and theoretical grounds for expecting sensitivity to prior probabilities. Empirically, some if not most of the cited fMRI work may be used to derive this prediction if one assumes that frontal sources contribute to the generation

of scalp-recorded P3a waves (Polich, 2007). Theoretically, Bayesian updating can be described as transforming prior probabilities to posterior probabilities (see Fig. 1). Because the cited ERP work has shown that amplitude variations in the P3a are related to Bayesian updating, they may also be related to prior probabilities. Surprise reflects the unexpectedness of the evidence—see the marginal probabilities $P(s_1)$ and $P(s_2)$ in Fig. 1—and relationships between surprise and P3b amplitude variations are manifest in many previous ERP studies. The concept of likelihood has hitherto been neglected in ERP studies, yet amplitude variations in the P3b may be sensitive to likelihoods, too. Most of the cited fMRI work may be used to derive this prediction if one assumes that parietal sources contribute to the generation of scalp-recorded P3b waves (Polich, 2007).

Method

Participants

Sixteen undergraduate psychology students participated to gain course credits (15 female, 1 male). Their age ranged from 19 to 50 years ($M = 24.7$, $SD = 9.3$ years of age). One participant was left-handed and two were ambidextrous. All participants indicated having normal or corrected-to-normal vision. The procedure was approved by the local ethics committee.

Task Procedures

As shown in Fig. 2a, we manipulated the task's probabilistic contingencies at two levels. We refer to these probabilistic contingencies as prior probabilities and likelihoods, respectively, and to their unique combination as a probability condition. At the beginning of each probability condition, a tableau of 10 urns containing a total of 100 colored balls was shown to participants. First, we introduced varying degrees of uncertainty about the type of urn by manipulating their prior probabilities. Second, we manipulated the proportion of ball colors within each urn, thereby manipulating likelihoods. Visualization of prior probabilities and likelihoods in the form of these tableaus allowed participants to build internal representations of probabilistic contingencies.

Uncertainty in prior distributions was manipulated by presenting 10 urns, composed of different numbers of type h_1 and type h_2 urns. On *uncertain* prior probability conditions (Pu), seven type h_1 urns and three type h_2 urns—that is, $P(h_1) = .7$, $P(h_2) = .3$ —were presented. Note that usage of the term *more uncertain* would be technically accurate; however, we use the technically inaccurate, but simpler *uncertain* throughout the article. On the *certain* prior probability condition (Pc), nine type h_1 urns and one type h_2 urn—that is, $P(h_1) = .9$, $P(h_2) = .1$ —were presented. Note that usage of the term *more certain* would be technically accurate; however, we use the technically inaccurate, but simpler *certain* throughout the article. Uncertainty in likelihoods was manipulated in the following way: On *uncertain* likelihood conditions (Lu), urn type h_1 contained seven red (s_1) and three blue (s_2) balls— $P(s_1|h_1) = .7$, $P(s_2|h_1) = .3$ —whereas urn type h_2 contained three red and seven blue balls—that is, $P(s_1|h_2) = .3$, $P(s_2|h_2) = .7$. On *certain* likelihood conditions (Lc), urn type h_1 contained nine red balls and one blue ball—that is, $P(s_1|h_1) = .9$, $P(s_2|h_1) = .1$ —whereas urn type h_2 contained one red ball and nine blue balls—that is, $P(s_1|h_2) = .1$, $P(s_2|h_2) = .9$. We chose these quite strongly unbalanced prior probabilities and likelihoods to optimize the discriminability of the informative hypotheses (see [Derivation of Eight Hypotheses of Interest](#)). The results may have been biased by the selection of these values. Ball colors were counterbalanced across participants. To avoid confusion, s_1

(s_2) always refers to that ball color in a tableau that was present in more (fewer) cases, regardless of the actual color. We refer to the frequent stimulus as s_1 and to the rare stimulus as s_2 in the remainder of this article.

Figure 2b illustrates the urn-ball paradigm and outlines the experimental setup for all four probability conditions. Repeated sampling consisted of the following sequence of events: First, one of the 10 urns was selected randomly, but the outcome of this selection remained hidden to the participants. Subsequently, a random sample of four balls was sequentially drawn, with replacement, from the selected urn, and balls were presented visually, one after the other, on the computer screen. Visual stimuli (i.e., the sampled balls) were presented at the center of a computer screen (Eizo FlexScan T766 19-in.; Hakusan, Ishikawa, Japan) against gray background (stimulus size = one degree, stimulus duration = 100 ms, stimulus onset asynchrony = 2,500 ms). Cortical responses to these visual stimuli constituted the brain activations that we measured. Participants were asked to indicate the color of each ball by pressing a key corresponding to one of the two colors. Participants also repeatedly had to perform probabilistic inverse inference. That is, they were asked after each sample of four balls to decide from which type of urn this sample might have originated (forced choice). No feedback as to the correctness of the choices was provided.

Each of the four probability conditions $c \in \{\text{PcLc}, \text{PcLu}, \text{PuLc}, \text{PuLu}\}$ comprised 50 episodes of random sampling, each consisting of four draws, yielding a total of 800 sequentially presented colored ball stimuli. We restricted our analyses to cortical responses to just the initial draws within the 200 (i.e., four times 50) episodes of sampling in an attempt to identify potential neural correlates of prior probabilities and likelihoods on the processing of the *initial* ball of each sample. Prior probabilities and likelihoods could solely be learned by instruction (i.e., inspection of the tableaus), whereas the expected values of relative frequencies of red and blue balls—PcLc: $P(s_1) = .82$, $P(s_2) = .18$; PcLu & PuLc: $P(s_1) = .66$, $P(s_2) = .34$; PuLu: $P(s_1) = .58$, $P(s_2) = .42$, respectively—could be learned by inspection of the tableaus as well as from experience. This holds true for large numbers of episodes of sampling where observed relative frequencies of red and blue balls should approximate their expected values.

Each participant completed four practice episodes of sampling under the supervision of and in a task-related dialogue with the experimenter to become accustomed to the task at the beginning of the experiment. The tableau of each practice sampling consisted of one single h_1 urn and one single h_2 urn (yielding uniform prior probabilities). Successful completion of these practice episodes of samplings demonstrated that the participants understood the procedure and their task. Each participant performed all four probability conditions with breaks of approximately 2 minutes between conditions. The

order of probability conditions was counterbalanced across participants. The experiment was run using Presentation® (Neurobehavioral Systems, Albany, CA).

Instructions read as follows: “You will experience a series of chance experiments. You will see ten urns, each of which is filled with 10 balls. Urns predominantly contain red or blue balls. One of the urns will be selected by chance, but you will never see which urn was sampled. Four balls will be randomly sampled with replacement from the selected urn, and these balls will be shown to you one after the other. You will have to identify the color of each of these balls by appropriate responding. After having seen a sample of balls, you will have to infer from which type of urn (the predominantly red or blue type) that particular sample may have originated. These episodes of sampling will be repeated many times.”

Electrophysiological Recordings and Data Preanalysis

The electroencephalogram (EEG) was recorded continuously, using a QuickAmps-72 amplifier (Brain Products, Gilching, Germany) and the Brain Vision Recorder Version 1.02 software (Brain Products, Gilching, Germany) from frontal (F7, F3, Fz, F4, F8), fronto-central (FCz), central (T7, C3, Cz, C4, T8), parietal (P7, P3, Pz, P4, P8), occipital (O1, O2), and mastoid (M1, M2) sites. Ag-AgCl EEG electrodes were used, which were mounted on an EasyCap (EasyCap, Herrsching-Breitbrunn, Germany). Electrode impedance was kept below 10 k Ω . All EEG electrodes were referenced to average reference during the recording. Participants were informed about the problem of noncerebral artifacts, and they were encouraged to reduce the occurrence of movement artifacts. Ocular artifacts were monitored by means of bipolar pairs of electrodes positioned at the sub- and supraorbital ridges (vertical electrooculogram; vEOG) and at the external ocular canthi (horizontal electrooculogram; hEOG). The EEG and EOG channels were subject to a band-pass filter of .01 Hz to 30 Hz and digitized at 250 Hz sampling rate with t as discrete time index and a resulting sampling period $T = 4$ ms.

Initial off-line analysis of the EEG data was performed by means of the Brain Vision Analyzer Version 2.0.1 software (Brain Products, Gilching, Germany). EEG electrodes were offline rereferenced to average mastoid reference. Careful manual artifact rejection was performed before averaging to discard trials during which eye movements or any other noncerebral artifact except blinks had occurred. One hundred and eleven trials of the 3,200 trials that were originally recorded had to be rejected (i.e., around 3.5%) so that 3,089 trials entered subsequent levels of analysis. Deflections in the averaged EOG waveforms were small, indicating that fixation was well maintained on those trials that survived the manual artifact

rejection process. Semi-automatic blink detection and the application of an established method for blink artifact removal were employed for blink correction (Gratton, Coles, & Donchin, 1983). Error trials were excluded so that solely artifact-free trials on which stimuli were correctly identified entered the final level of analysis—error rates for frequent (s_1) stimuli: PcLc = .01, PcLu = .02, PuLc = .02, PuLu = .02; for rare (s_2) stimuli: PcLc = .05, PcLu = .04, PuLc = .04, PuLu = .04.

Data Reduction

The continuous EEG obtained at the finally selected electrodes Fz, F4, and Pz (cf. Selection of EEG/ERP Data for Statistical Analysis) was divided for each individual participant $i \in \{1, \dots, N\}$ with $N = 16$ into epochs $x_{i,c,s,m}(t)$ of 700 ms duration starting 100 ms before stimulus onset with $m \in \{1, \dots, M_{i,c,s}\}$, where $M_{i,c,s}$ denotes the number of epochs of stimulus type $s \in \{s_1, s_2\}$ presented to participant i in probability condition $c \in \{PcLc, PcLu, PuLc, PuLu\}$. Epochs were corrected using the interval $[-100, 0]$ ms before stimulus presentation as the baseline.

Event-related potential (ERP) waveforms $\bar{x}_{i,c,s}(t)$ were created for each individual participant, separately for each stimulus type and for each probability condition, by averaging over epochs $\bar{x}_{i,c,s}(t) = \frac{1}{M_{i,c,s}} \sum_{m=1}^{M_{i,c,s}} x_{i,c,s,m}(t)$. Because of enormous interindividual variability in ERP waveforms from individual participants, the corresponding stimulus-type-specific and probability-condition-specific waves were separately averaged over participants yielding grand-average waveforms $x_{c,s}(t) = \frac{1}{N} \sum_{i=1}^N \bar{x}_{i,c,s}(t)$.

A moving average $\tilde{x}_{c,s}(t) = \frac{1}{9} \sum_{\tau=-4T}^{4T} x_{c,s}(t + \tau)$ was applied to these grand average ERP waveforms to reduce high-frequency noise to find the time of maximum variance, $t_{\sigma_{\max}^2,s}$, between probability conditions.

Following

$$t_{\sigma_{\max}^2,s} = \arg \max_t \left\{ \text{Var} \left(\tilde{x}_{PcLc,s}(t), \tilde{x}_{PcLu,s}(t), \tilde{x}_{PuLc,s}(t), \tilde{x}_{PuLu,s}(t) \right) \right\},$$

ERP latencies with the maximum variance between the four probability conditions were separately identified for frequently (s_1) and rarely (s_2) occurring stimuli for $t \in [290, 400]$ ms corresponding to a wide latency range for measuring the P300 (Polich, 2007).

In the next step, participant-specific single-trial voltages were registered at the emergent grand-average latencies of maximum variance $t_{\sigma_{\max}^2,s}$ (Fz: s_1 : 312 ms, s_2 : 356 ms; F4: s_1 : 332 ms, s_2 : 352 ms; Pz: s_1 : 360 ms, s_2 : 364 ms; visual inspection of the variance over time confirmed these points in time as distinctive points of maximum between-condition variability) separately for each probability condition. These single-trial voltages were smoothed in order to reduce high frequency

noise following

$$\tilde{x}_{i,c,s,m} = \frac{1}{9} \sum_{\tau=-4T}^{4T} x_{i,c,s,m} \left(t_{\sigma_{\max}^2, s}^2 + \tau \right),$$

and subsequently averaged over epochs $\bar{x}_{i,c,s} = \frac{1}{M_{i,c,s}} \sum_{m=1}^{M_{i,c,s}} \tilde{x}_{i,c,s,m}$. Bayesian model selection was conducted based on these means (see below).

Signal-to-Noise Ratio Estimation

The signal-to-noise ratio (SNR) for ERPs can be defined as the ratio of the ERP power P_{ERP} to the noise power P_{D} (beim Graben, 2001; Kolossa, 2016)

$$\text{SNR} = \frac{P_{\text{ERP}}}{P_{\text{D}}},$$

which can be specified in decibel (dB) according to

$$\text{SNR [dB]} = 10 \log_{10} \left(\frac{P_{\text{ERP}}}{P_{\text{D}}} \right).$$

We calculated the condition- and stimulus-specific power estimates, $\hat{P}_{\text{ERP},i,c,s}$ and $\hat{P}_{\text{D},i,c,s}$, respectively, for individual participants based on the single trial ERP estimates $\tilde{x}_{i,c,s,m}$ following Möcks, Gasser, and Köhler (1988). These authors suggested that

$$\hat{P}_{\text{D},i,c,s} = \frac{1}{M_{i,c,s}-1} \sum_{m=1}^{M_{i,c,s}} (\tilde{x}_{i,c,s,m} - \bar{x}_{i,c,s})^2,$$

and

$$\hat{P}_{\text{ERP},i,c,s} = \frac{1}{M_{i,c,s}} \left(\sum_{m=1}^{M_{i,c,s}} (\tilde{x}_{i,c,s,m})^2 - \hat{P}_{\text{D},i,c,s} \right).$$

Under the assumption that noise in any epoch is not correlated with the ERP nor with noise in any other epoch, averaging over $M_{i,c,s}$ epochs improves the condition- and stimulus-specific $\text{SNR}_{i,c,s}$ of ERP average waveforms for individual participants by $\sqrt{M_{i,c,s}}$, yielding

$$\text{SNR}_{i,c,s} [\text{dB}] = 10 \log_{10} \left(\frac{M_{i,c,s} \cdot \hat{P}_{\text{ERP},i,c,s}}{\hat{P}_{\text{D},i,c,s}} \right).$$

The mean SNR over individual participants further simply follows

$$\text{SNR}_{c,s} [\text{dB}] = \frac{1}{N} \sum_{i=1}^N \text{SNR}_{i,c,s} [\text{dB}].$$

Table 1 shows the condition- and stimulus-specific mean $\text{SNR}_{c,s}$ [dB] as well as the standard deviation, minimum and maximum values of individual $\text{SNR}_{i,c,s}$ [dB]

separately at the finally selected electrodes Fz, F4 and Pz (cf. Selection of EEG/ERP Data for Statistical Analysis). This is a theoretical upper bound on the SNR under ideal assumptions, which are not met in real data (beim Graben, 2001). Actual SNRs are expected to be lower. Table 1 presents SNRs, mean, minimum and maximum number of trials per probability condition and stimulus type, $M_{i,c,s}$, that went into ERP averaging at the level of individual participants.

Selection of EEG/ERP Data for Statistical Analyses

Late positive potentials of the ERP are decomposable into at least two separable components (Dien, Spencer, & Donchin, 2004; Polich, 2007; Sutton & Ruchkin, 1984). Those ERP components are a parietally distributed positivity (P3b, with maximum peak at electrode Pz) and a frontally distributed positivity (P3a, with maximum peak at more anterior midline electrodes). Three principles for selecting data for statistical analyses governed our initial attempts to reveal statistically significant effects of prior probabilities and likelihoods on P3a and P3b amplitude variations: (1) Analyze EEG data as recorded in single trials, (2) analyze EEG data recorded at midline electrodes, and (3) analyze all eight measured conditions (four probability conditions times two stimulus types). Application of these principles was associated with difficulties to consistently disprove the null hypothesis, which we attributed to the abundantly large random variation in these measurements, mostly because of the notoriously low SNR of single-trial EEG data.

Therefore, the principles for selecting data for statistical analyses were changed, as follows: (1) Analyze averaged ERP data with sufficient SNRs, and (2) analyze ERP data recorded at the electrodes where prior probabilities and likelihoods exerted maximum effects on face value. Application of these principles led to the following decisions: (1) As shown in Table 1, only ERP averages in response to the frequent stimuli s_1 , but not those in response to the rare stimuli s_2 , possessed sufficient SNRs throughout all four probability conditions. The SNR of ERP averages in response to the rare stimuli s_2 in the PcLc probability condition fell below acceptable levels because of insufficient numbers of trials that entered these ERP averages (fewer than 10 trials on average). Based on these data, we decided to restrict the statistical analyses to the s_1 ERP averages that were measured in the four probability conditions. (2) Fig. 3 (middle panels) shows topographic maps of difference waves (likelihood difference waves and prior probability difference waves). Inspection of Fig. 3 (middle panels) reveals that the likelihood manipulation exerted maximum effects on face value at electrode Pz. The prior probability manipulation exerted maximum effects on face value at electrode F4 in close proximity to electrode Fz. Therefore, in

Table 1 The mean, standard deviation, minimum and maximum values of signal-to-noise ratios (SNRs) at electrodes Fz, F4, and Pz are presented separately for each stimulus type *s* and probability condition *c*.Additionally, the respective means as well as minimum and maximum numbers of trials $M_{i,c,s}$ are given

		Stimulus type and probability condition							
		s_1				s_2			
		PcLc	PcLu	PuLc	PuLu	PcLc	PcLu	PuLc	PuLu
$M_{i,c,s}$	<i>M</i>	40.63	31.63	30.75	27.19	8.06	16.81	17.25	20.75
	<i>min</i>	34	26	25	13	2	10	13	9
	<i>max</i>	46	36	37	38	14	23	22	35
$SNR_{c,s}$ at Fz	<i>M</i>	16.08	14.93	14.92	14.54	8.99	11.98	12.91	13.47
	<i>SD</i>	0.43	0.53	0.76	1.31	1.85	1.43	1.51	0.97
	<i>min</i>	15.11	13.80	13.79	11.45	5.36	10.01	10.41	12.04
	<i>max</i>	16.65	15.43	16.91	16.41	12.53	15.06	16.17	15.28
$SNR_{c,s}$ at F4	<i>M</i>	16.14	15.04	14.83	14.38	8.80	12.08	12.39	13.32
	<i>SD</i>	0.42	0.63	0.70	1.52	2.22	1.34	1.13	1.22
	<i>min</i>	15.26	13.91	13.62	10.85	3.07	10.14	10.50	11.88
	<i>max</i>	16.95	16.19	16.30	17.45	13.16	14.27	14.05	15.75
$SNR_{c,s}$ at Pz	<i>M</i>	17.27	16.62	16.32	15.68	10.22	13.78	13.91	14.84
	<i>SD</i>	1.20	1.30	1.65	1.50	5.90	1.96	2.02	1.77
	<i>min</i>	15.34	14.19	14.01	13.72	-10.08	10.65	10.43	12.17
	<i>max</i>	20.81	19.26	19.51	20.14	16.44	17.43	19.32	17.69

Note. PcLc: Certain prior probability, certain likelihood; PcLu: Certain prior probability, uncertain likelihood; PuLc: Uncertain prior probability, certain likelihood; PuLu: Uncertain prior probability, uncertain likelihood

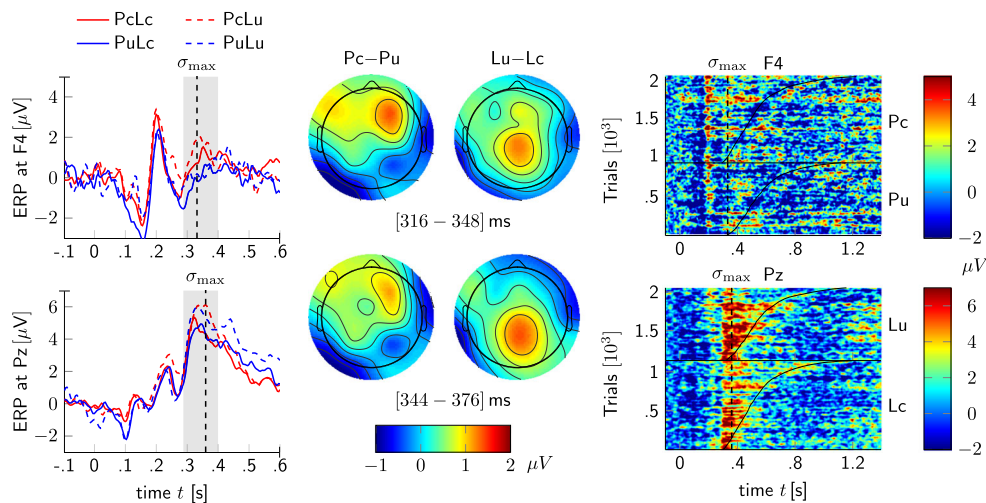


Fig. 3 (Left panels). Grand-average ERP wave shapes at electrodes F4 (upper panel) and Pz (lower panel) in response to frequent stimuli s_1 in the four probability conditions (all Pc waves in red lines; all Pu waves in blue lines; all Lc waves in solid lines; all Lu waves in dashed lines). Both diagrams show the latency range for measuring the P300 (i.e., from 290 to 400 ms; in light gray) as well as the latencies of maximum variance as defined in the text (i.e., F4: 332 ms; Pz: 360 ms). **(Middle panels).** Topographic maps of difference waves (on the left: prior probability difference waves, i.e., Pc–Pu; on the right: likelihood difference waves, i.e., Lu–Lc). The prior probability manipulation exerted maximum effects at electrode F4 in close proximity to electrode Fz around the latency of

maximum variance (i.e., 316–348 ms). The likelihood manipulation exerted maximum effects at electrode Pz around the latency of maximum variance (i.e., 344–376 ms). **(Right panels).** ERP images at F4 and Pz electrodes in response to frequent stimuli s_1 from the four probability conditions. ERP images at F4 (upper panel) show color-coded single-trial EEG data separately for certain (Pc) and uncertain (Pu) prior probability conditions, sorted according to the response times on these trials (from fastest to slowest). ERP images at Pz (lower panel) show color-coded single-trial EEG data separately for uncertain (Lu) and certain (Lc) likelihood conditions, sorted according to the response times on these trials (from fastest to slowest). (Color figure online)

addition to electrodes Fz and Pz, electrode F4 was selected for statistical analyses.

Derivation of Eight Hypotheses of Interest

We used information-theoretic measures (see Fig. 4a; see also MacKay, 2003) such as surprise (Kolossa et al., 2013), entropy (Kopp & Lange, 2013), and redundancy (Barlow, 2001) as response functions for modeling the measured ERP responses (Friston, 2005). *Surprise* (I) equals a logarithmic measure of the probability $p(x)$ that a binary random variable X takes on the value x . Therefore, $I(x) = -\log_2 p(x)$; $I(x) \in [0, +\infty]$ (with minimum surprise ($I(x) = 0$) at $p(x) = 1.0$ and maximum surprise ($I(x) \rightarrow +\infty$ at $p(x) = .0$). *Entropy* (H) equals the summation of surprise values over all the possible values x that X can take multiplied with their respective probability. For a binary random variable X it is $H(X) = -\sum_{k=1}^2 p(x_k) \log_2(p(x_k))$; $H(X) \in [0, 1]$ (with minimum entropy ($H(X) = 0$) at $p(x_k) = .0$ and at $p(x_k) = 1.0$ and maximum entropy ($H(X) = 1$) at $p(x_k) = .5$). *Redundancy* (R) is sort of a reciprocal value to entropy, and it equals $R(X) = 1 - H(X)$; $R(X) \in [0, 1]$ (with minimum redundancy ($R(X) = 0$) at $p(x_k) = .5$ and maximum redundancy ($R(X) = 1$ at $p(x_k) = .0$ and at $p(x_k) = 1.0$).

Informative hypotheses pose specific (in-)equality constraints on the measured cortical responses that we obtained from the frequent stimulus s_1 in the four probability conditions (PcLc, PcLu, PuLc, PuLu). We use the following notation here: Certain condition (.9 and .1 probability distribution), uncertain condition (.7 and .3 probability distribution), $s_{cc} = s_1$ in the certain prior and certain likelihood condition; $s_{cu} = s_1$ in the certain prior and uncertain likelihood condition; $s_{uc} = s_1$ in the uncertain prior and certain likelihood condition; $s_{uu} = s_1$ in the uncertain prior and uncertain likelihood condition. Equivalently, μ_{cc} , μ_{cu} , μ_{uc} , and μ_{uu} correspond to the predicted condition-specific mean ERP voltages. There are three groups of conceivable informative hypotheses.

The first group of hypotheses subsumes non-Bayesian hypotheses. We refer to this group of hypotheses as non-Bayesian hypotheses because their (in-)equality constraints are based on the unexpectedness of the evidence or surprise—that is, the marginal probability $P(s_1)$ in Fig. 1. The unexpectedness of s_1 draws varies between .82 and .58 on the four experimental conditions, with $P(s_{cc}) = .82$, $P(s_{cu}) = .66$, $P(s_{uc}) = .66$, $P(s_{uu}) = .58$ (see Fig. 2a). Application of the three information-theoretic transformations (I , H , R ; see Fig. 4b) to these probabilities yields two distinct hypotheses, namely:

$$\begin{aligned} H_1: \mu_{uu} > \{\mu_{cu} = \mu_{uc}\} > \mu_{cc} \\ [(I;H) \text{ hypotheses: } I_{.58} > \{I_{.66} = I_{.66}\} > I_{.82}; H_{.58} > \{H_{.66} = H_{.66}\} > H_{.82}] \\ H_2: \mu_{uu} < \{\mu_{cu} = \mu_{uc}\} < \mu_{cc} \\ [(R) \text{ hypothesis: } R_{.58} < \{R_{.66} = R_{.66}\} < R_{.82}]. \end{aligned}$$

Fig. 4 (a). Surprise (I , red curve), entropy (H , blue curve) and redundancy (R , green curve) over the probability P of a binary random variable. (b). Ordinal inequality predictions of P300 amplitude variations for the frequent stimulus s_1 in the four probability conditions (PcLc, PcLu, PuLc, PuLu) as a function of surprise (I , red curve), entropy (H , blue curve) and redundancy (R , green curve) with $P =$ unexpectedness of s_1 at the abscissa. (c). Ordinal inequality predictions of P300 amplitude variations for the frequent stimulus s_1 in the four probability conditions as a function of surprise (I , red curve), entropy (H , blue curve) and redundancy (R , green curve) with $P =$ prior probability at the abscissa. (d). Ordinal inequality predictions of P300 amplitude variations for the frequent stimulus s_1 in the four probability conditions as a function of surprise (I , red curve), entropy (H , blue curve) and redundancy (R , green curve) with $P =$ likelihood at the abscissa. (Color figure online)

Surprise- and entropy-transformed non-Bayesian hypotheses are associated with indistinguishable (in-)equality constraints, but their predictions can be differentiated from those of the redundancy-transformed non-Bayesian hypothesis.

The second group of hypotheses subsumes two prior probability hypotheses. Recall that the condition-specific prior probabilities amounted to $P(s_{cc}) = .90$, $P(s_{cu}) = .90$, $P(s_{uc}) = .70$, $P(s_{uu}) = .70$. Application of the three information-theoretic transformations (I , H , R ; see Fig. 4c) to these prior probabilities yields two distinct hypotheses, namely:

$$\begin{aligned} H_3: \{\mu_{uu} = \mu_{uc}\} > \{\mu_{cu} = \mu_{cc}\} \\ [(I;H) \text{ hypotheses: } \{I_{.7} = I_{.7}\} > \{I_{.9} = I_{.9}\}; \{H_{.7} = H_{.7}\} > \{H_{.9} = H_{.9}\}] \\ H_4: \{\mu_{uu} = \mu_{uc}\} < \{\mu_{cu} = \mu_{cc}\} \\ [(R) \text{ hypothesis: } \{R_{.7} = R_{.7}\} < \{R_{.9} = R_{.9}\}]. \end{aligned}$$

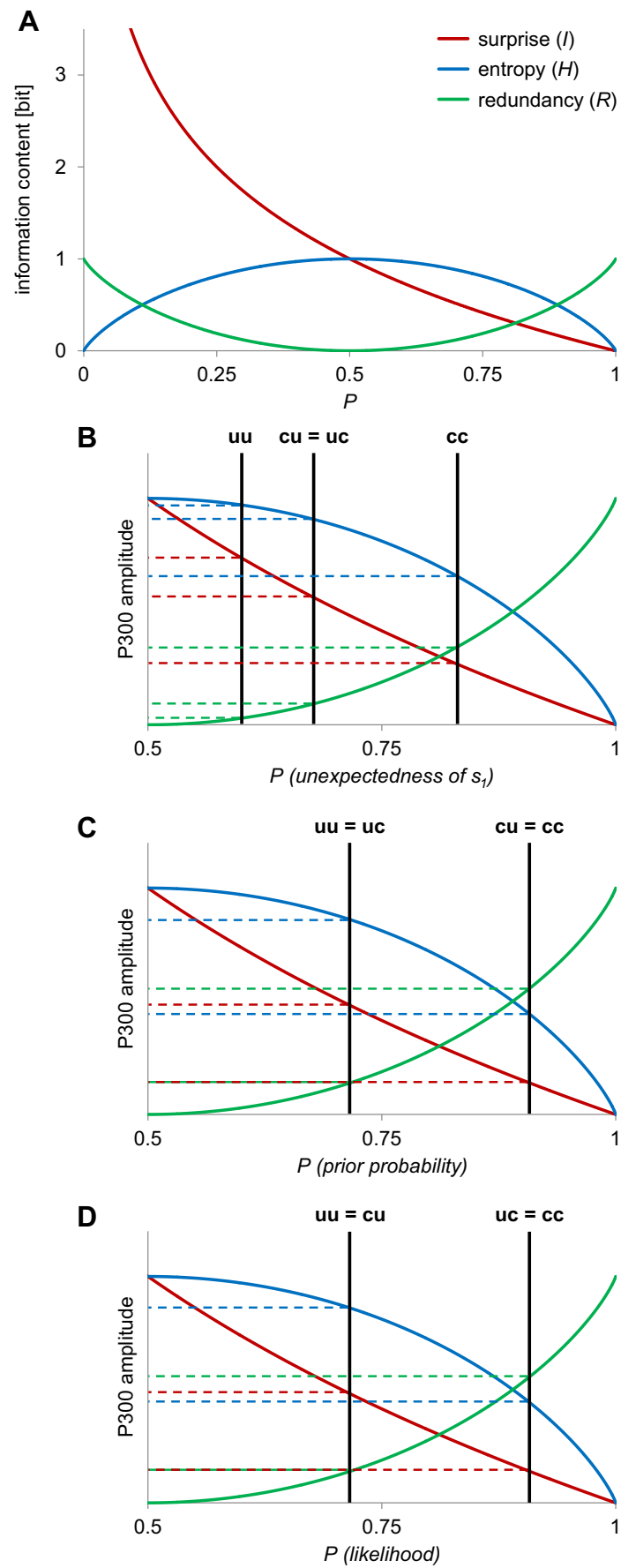
Surprise- and entropy-transformed prior probability hypotheses are associated with indistinguishable (in-)equality constraints, but their predictions can be differentiated from those of the redundancy-transformed prior probability hypothesis.

The third group of hypotheses subsumes two likelihood hypotheses. Recall that the condition-specific likelihoods amounted to $P(s_{cc}) = .90$, $P(s_{cu}) = .70$, $P(s_{uc}) = .90$, $P(s_{uu}) = .70$. Application of the three information-theoretic transformations (I , H , R ; see Fig. 4d) of the likelihoods yields two hypotheses, namely:

$$\begin{aligned} H_5: \{\mu_{uu} = \mu_{cu}\} > \{\mu_{uc} = \mu_{cc}\} \\ [(I;H) \text{ hypotheses: } \{I_{.7} = I_{.7}\} > \{I_{.9} = I_{.9}\}; \{H_{.7} = H_{.7}\} > \{H_{.9} = H_{.9}\}] \\ H_6: \{\mu_{uu} = \mu_{cu}\} < \{\mu_{uc} = \mu_{cc}\} \\ [(R) \text{ hypothesis: } \{R_{.7} = R_{.7}\} < \{R_{.9} = R_{.9}\}] \end{aligned}$$

Surprise- and entropy-transformed likelihood hypotheses are associated with indistinguishable (in-)equality constraints, but their predictions can be differentiated from those of the redundancy-transformed likelihood hypothesis.

The null hypothesis equals the traditional equality constraint, i.e.,



$$H_0: \{\mu_{uu} = \mu_{cu} = \mu_{uc} = \mu_{cc}\}$$

In contrast, the unconstrained hypothesis does not impose any constraints, i.e.,

$$H_u: \{\mu_{uu}, \mu_{cu}, \mu_{uc}, \mu_{cc}\}$$

Bayesian Evaluation of the Eight Hypotheses of Interest

For each of the Fz, F4, and Pz electrodes, averaged ERP voltage measurements are obtained for each of 16 participants in the four probability conditions (see section Selection of EEG/ERP Data for Further Analyses). *Averaged* means that within each condition, single-trial measurements for each participant are averaged for the frequent stimulus (s_1). This results for each electrode in a data matrix consisting of 16 rows (one for each participant) containing the averaged ERP voltage measurements $\bar{x}_{i,c}$ in each of four conditions.

Eight informative hypotheses (Hojtink, 2012, pp. 5–8, 23) with respect to the structure in the means have been formulated. These hypotheses will be evaluated using Bayes factors (BF ; Hoijtink, 2012, pp. 51–52) that can be transformed into posterior model probabilities (PMP s; Hoijtink, 2012, pp. 52–53). This is done using five steps (see Appendix A for a technical elaboration of the description given below):

- Step 1. Model each data matrix with a specific instance of the multivariate normal linear model (Hojtink, 2012, pp. 22–23), namely, the repeated-measures model (Hojtink, 2012, pp. 29–32). The main parameters of the repeated measures model are μ_{cc} , μ_{cu} , μ_{uc} , and μ_{uu} , that is, the mean of the 16 measurements in each condition and Σ the covariance matrix of the voltage measurements in each of the four conditions (see Appendix A, Section 1).
- Step 2. Specify the unconstrained prior distribution (Hojtink, 2012, pp. 46–48) of the parameters under H_u , that is, a low informative prior distribution for the means and a noninformative prior distribution for the covariance matrix (see Appendix A, Section 2).
- Step 3. Combine the density of the data (Hojtink, 2012, pp. 44–45) of the repeated-measures model with the prior distribution under H_u to obtain the unconstrained posterior distribution (Hojtink, 2012, pp. 49–51) of the parameters under H_u (see Appendix A, Section 3).
- Step 4. Use the result from Hoijtink (2012, pp. 51–52) to compute $BF_{mu} = f_m / c_m$ for $m = 0, \dots, 6$, that is, the Bayes factor quantifying the relative support in the data for H_m and H_u , where f_m denotes the fit of H_m (the proportion of the posterior distribution in agreement with

the constraints in H_m) and c_m denotes the complexity of H_m (the proportion of the prior distribution in agreement with the constraints in H_m). A Bayes factor of 5 indicates that after observing the data the support for H_m is 5 times higher than the support for H_u if the fit and complexity of both hypotheses are accounted for (see Appendix A, Section 4).

- Step 5. To facilitate the interpretation of the seven resulting Bayes factors for H_0 to H_6 , they can be transformed into posterior model probabilities PMP_m (Hojtink, 2012, pp. 52–53) for each of the 7 constrained hypotheses and H_u . Assuming that a priori each hypothesis is equally likely, $PMP_m = BF_{mu} / (1 + \sum_{m'} BF_{m'u})$ for $m' = 0, \dots, 6$, and $PMP_u = 1 / (1 + \sum_{m'} BF_{m'u})$ for H_u . These posterior model probabilities are numbers on a scale from 0 to 1 that quantify the support in the data for each of the hypotheses under consideration (see Appendix A, Section 5).

BF s and PMP s properly weighting the fit and complexity of each hypothesis under consideration can be computed using the software package BIEMS (Mulder, Hoijtink, & de Leeuw, 2012). The results are displayed in Table 2. As can be seen in the last two columns, for example, both H_1 (with a Bayes factor of 1.63) and H_5 (with a Bayes factor of 10.31) receive more support than H_u . This implies that the constraints used to formulate both hypotheses are supported by the data. By transforming the Bayes factors to posterior model probabilities, it can be seen that the support for H_5 (with a posterior model probability of .71) is much larger than for H_1 (with a posterior model probability of .11). The next section will provide an elaborate interpretation of the results in Table 2.

Results

Descriptive ERP Results

Figure 3 (left panels) depicts grand-average ERP wave shapes at F4 and Pz electrodes in response to frequent stimuli s_1 in the four probability conditions (all Pc waves in red lines; all Pu waves in blue lines; all Lc waves in solid lines; all Lu waves in dashed lines). Both diagrams show the latency range for measuring the P300 (i.e., from 290 to 400 ms; in light gray) as well as the latencies of maximum variance as defined above (i.e., F4: 332 ms; Pz: 360 ms). Inspection of the ERP wave shapes at electrode F4 reveals that the ERP waves in response to s_1 stimuli in the certain prior probability condition (Pc) were more positive than the ERP waves in response to s_1 stimuli in the uncertain prior probability condition (Pu) at the latency of maximum variance, identifiable as a difference between red and blue ERP wave shapes. Inspection of the ERP wave shapes at electrode Pz reveals that the ERP waves in response

Table 2 Bayes factors (*BFs*) and posterior model probabilities (*PMPs*) for the eight hypotheses tested at electrodes Fz, F4, and Pz

	Class of hypotheses	Transformation	Fz		F4		Pz	
			<i>BF</i>	<i>PMP</i>	<i>BF</i>	<i>PMP</i>	<i>BF</i>	<i>PMP</i>
H_1	non-Bayesian hypotheses	surprise, entropy	0.33	.06	0.04	.01	1.63	.11
H_2		redundancy	0.41	.07	0.44	.06	0.03	< .01
H_3	prior hypotheses	surprise, entropy	0.22	.04	0.05	.01	0.40	.03
H_4		redundancy	1.55	.26	5.01	.68	0.26	.02
H_5	likelihood hypotheses	surprise, entropy	1.52	.25	0.46	.06	10.31	.71
H_6		redundancy	0.22	.04	0.10	.01	0.11	.01
H_0	reference hypotheses		0.77	.12	0.31	.04	0.76	.05
H_u			1	.17	1	.14	1	.07

Note. *Bolface terms* indicate the hypotheses that received maximum support

to s_1 stimuli in the uncertain likelihood condition (Lu) were more positive than the ERP waves in response to s_1 stimuli in the certain likelihood condition (Lc) at the latency of maximum variance, identifiable as a difference between dashed and solid ERP wave shapes.

Figure 3 (middle panels) shows topographic maps of difference waves (likelihood difference waves, i.e., Lu–Lc, and prior probability difference waves, i.e., Pc–Pu). Inspection of the maps reveals that the likelihood manipulation exerted maximum effects at electrode Pz around the latency of maximum variance (i.e., from 344 to 376 ms). The prior probability manipulation exerted maximum effects at electrode F4 in close proximity to electrode Fz around the latency of maximum variance (i.e., from 316 to 348 ms).

Figure 3 (right panels) depicts ERP images at F4 and Pz electrodes in response to the frequent stimulus s_1 from the four probability conditions. The ERP images at F4 (upper panel) show color-coded single-trial EEG data separately for certain (Pc) and uncertain (Pu) prior probability conditions, sorted according to the response times on these trials (from fastest to slowest). It can be seen that there is a trend toward more positive single-trial EEG deflections in response to s_1 stimuli in certain as compared to uncertain prior probability conditions at the latency of maximum variance, thus ruling out the possibility that the apparent prior probability effect that was observed in grand-average ERP wave shapes (see upper left panel) might simply be related to boosted latency jitter in the uncertain prior probability condition in comparison to the certain prior probability condition. The ERP images at Pz (lower panel) show color-coded single-trial EEG data separately for uncertain (Lu) and certain (Lc) likelihood conditions, sorted according to the response times on these trials (from fastest to slowest). It can be seen that there is a trend toward more positive single-trial EEG deflections in response to s_1 stimuli in uncertain as compared to certain likelihood conditions at the latency of maximum variance, thus ruling out the possibility that the

apparent likelihood effect that was observed in grand-average ERP wave shapes (see lower left panel) might simply be related to boosted latency jitter in the certain likelihood condition in comparison to the uncertain likelihood condition.

Bayesian Hypothesis Evaluation

Table 2 displays *BFs* and *PMPs* for each of the eight competing hypotheses H_1 through H_6 as well as H_0 and H_u separately at each of the three electrodes that were examined (i.e., Fz, F4, Pz). Guidelines for interpreting the magnitudes of *BFs* can be found in Kass and Raftery (1995). *BF* values between 1 and 3 denote a degree of support that is barely worth mentioning. *BF* values between 3 and 10 denote substantial support for the hypothesis under investigation (H_i) compared to a reference hypothesis such as for example H_u . *BF* values between 10 and 100 denote strong evidence in favor of H_i compared to H_u , whereas *BF* values larger than 100 denote decisive evidence. *BF* values smaller than one denote marginal to decisive evidence against H_i . As can be seen from inspection of Table 2, the non-Bayesian hypotheses H_1 and H_2 did generally not receive support from the data ($.03 \leq BFs \leq 1.63$, below $.01 \leq PMPs \leq .11$).

The results at electrode Fz remain inconclusive. The prior probability (redundancy) hypothesis H_4 receives the most support from the data, with $BF_4 = 1.55$ and $PMP_4 = .26$. One should be reluctant to conclude from these data that ERP amplitude variations at Fz are associated with prior probabilities. Recall that $BF < 3$ denotes a degree of support for any H_i that is barely worth mentioning. The support for H_4 is just $.26/.12 = 2.2$ times higher than the support for H_0 , and the support for H_4 is just $.26/.17 = 1.53$ times higher than the support for H_u . The support for the best likelihood (surprise, entropy) hypothesis H_5 , was virtually indistinguishable from that for hypothesis H_4 , with $BF_5 = 1.52$ and $PMP_5 = .25$.

The results at electrode F4 lend substantial support for prior probability (redundancy) hypothesis H_4 , with $BF_4 = 5.01$ and $PMP_4 = .68$. Recall that $3 < BF < 10$ denotes a substantial degree of support for any H_i . The support for H_4 is $.68/.04 = 17$ times higher than the support for H_0 , and the support for H_4 is $.68/.14 = 4.86$ times higher than the support for H_u . The support for the prior probability (redundancy) hypothesis H_4 was 11.3 times higher than the support for the best likelihood (surprise, entropy) hypothesis H_5 , with $BF_5 = .46$ and $PMP_5 = .06$. One can conclude from these data that ERP amplitude variations at F4 are associated with prior probabilities.

The results at electrode Pz provide strong support for likelihood (surprise, entropy) hypothesis H_5 , with $BF_5 = 10.31$ and $PMP_5 = .71$. Recall that $BF > 10$ denotes a strong degree of support for any H_i . The support for H_5 is $.71/.05 = 14.2$ times higher than the support for H_0 , and the support for H_5 is $.71/.07 = 10.14$ times higher than the support for H_u . The support for the likelihood (surprise, entropy) hypothesis H_5 was 11.3 times higher than the support for the best prior probability (surprise, entropy) hypothesis H_3 , with $BF_3 = .40$ and $PMP_3 = .03$. One can conclude from these data that ERP amplitude variations at Pz are associated with likelihoods.

Discussion

Our data show that the manipulation of prior probabilities and likelihoods is associated with separately modifiable and distinct ERP responses, thereby demonstrating that the antecedents of Bayesian inference are coded by the brain. We found that the manipulation of prior probabilities was associated with P3a amplitude variations, whereas the manipulation of likelihoods was related to P3b amplitude variations. P3a amplitudes were sensitive to the degree of prior certainty such that more certain prior probabilities were related to larger frontally distributed P3a waves, irrespective of the level of certainty in the likelihoods. P3b amplitudes were sensitive to the degree of likelihood certainty such that less certain likelihoods were associated with larger parietally distributed P3b waves, irrespective of the level of prior certainty.

The Current Study in the Context of the Existing Literature

Two Bayesian parameters were instructed, that is, prior probabilities and likelihoods (FitzGerald et al., 2010; Hertwig et al., 2004; Hertwig & Erev, 2009). Therefore, our findings are novel because there are no previous studies that demonstrate ERP correlates of instructed Bayesian parameters. This may help to understand why our P3b negative results with regard to the unexpectedness of the evidence seem to be in disagreement with the existing P3b literature (see Kolossa

et al., 2013; Kolossa et al., 2015). Most P3b studies show that P3b amplitude variations are sensitive to the unexpectedness of the evidence or surprise, but none of these studies has examined instructed unexpectedness or surprise against instructed likelihoods.

The available fMRI research suggests fronto-striatal representations of prior parameters and predominantly parietal representations of likelihood parameters. These broad anatomical distinctions are in accordance with our findings if one assumes that frontal sources contribute to the generation of scalp-recorded P3a waves and that parietal sources contribute to the generation of scalp-recorded P3b waves (Polich, 2007). The P3a finding needs further specification because previous ERP studies revealed an association between Bayesian updating and P3a amplitudes rather than between prior probabilities and P3a amplitudes (Bennett et al., 2015; Kolossa et al., 2015). The P3a waves in this study were more anteriorly distributed than the P3a waves in the previous studies, but this issue clearly needs to be further examined by future studies. It also remains open why higher prior certainty was associated with larger frontal P3a amplitudes rather than with smaller P3a amplitudes. The direction of the effect of likelihood uncertainty on P3b amplitude variations is congruent with the results that were obtained from fMRI studies because higher likelihood uncertainty was repeatedly associated with larger parietal activation (Furl & Averbeck, 2011; Stern et al., 2010).

The P3b Component in the Urn-Ball Task

Manipulation of the likelihoods resulted in P3b amplitude variations at electrode Pz. Specifically, s_1 stimuli in uncertain likelihood conditions evoked more prominent P3b responses compared to s_1 stimuli in certain likelihood conditions. This finding is in accord with previous research on the P3b ERP component, which revealed that P3b amplitude variations are related to the uncertainty or surprise of its eliciting events (Donchin, 1981; Kolossa et al., 2013; Sutton, Braren, Zubin, & John, 1965).

Likelihoods express conditional probabilities, that is, $P(s|h_1)$ and $P(s|h_2)$ in the current urn-ball task. Applied to stimuli s_1 , likelihoods amounted to $P(s_1|h_1) = .7$ and $P(s_1|h_2) = .3$ in uncertain likelihood conditions, and they amounted to $P(s_1|h_1) = .9$ and $P(s_1|h_2) = .1$ in certain likelihood conditions. The urn-ball task can be understood as a binary classification test, and sensitivity [here, $P(s_1|h_1) = .7$ vs. $P(s_1|h_1) = .9$] and specificity [here, $P(s_1|h_2) = .3$ vs. $P(s_1|h_2) = .1$] are well-known statistical measures of the performance of such tests (e.g., Green & Swets, 1966). Low sensitivity was coupled with low specificity in uncertain likelihood conditions, thereby minimizing the information that can be gained from an observed s_1 event in this condition. High sensitivity was coupled with high specificity in certain likelihood conditions, thereby maximizing the information

that can be gained from an observed s_1 event in this condition. P3b amplitude variations in the urn-ball task reflect variations in sensitivity and specificity of an observed s_1 event because we measured more prominent P3b responses under relatively low levels of sensitivity and specificity in comparison to higher levels of these probabilistic quantities. It is easy to see how sensitivity and specificity can be combined to compute the logarithm of the likelihood ratio (logLR), that is, $\ln(P(s|h_1)/P(s|h_2))$, that was formalized in Equation 3.

Studies of perceptual decisions in primates (Kira, Yang, & Shadlen, 2015; Yang & Shadlen, 2007) and in rodents (Hanks et al., 2015) showed that the firing rates of neurons in the parietal cortex reflected the gradual accumulation of logLR when different logLRs were assigned to various sensory cues. Although the underlying neural mechanisms are less understood, a similar framework explains a variety of P3b-like phenomena that are related to perceptual decisions in humans (Kelly & O'Connell, 2013; O'Connell, Dockree, & Kelly, 2012). Our P3b findings can thus be viewed as reflecting activities of a neural module, putatively located in the parietal cortex, which traces the accumulation of sensory evidence in terms of logLR. Our P3b results also imply that this posterior probabilistic module does not seem to be sensitive to prior probabilities, and this question was never addressed before.

The P3a Component in the Urn-Ball Task

Manipulation of prior probabilities resulted in P3a amplitude variations at electrode F4. Specifically, s_1 stimuli in certain prior probability conditions evoked more prominent P3a responses compared to s_1 stimuli in uncertain prior probability conditions. There is not much previous knowledge about the relationships between P3a amplitude variations and probabilistic quantities, such as uncertainty and surprise. Two earlier studies from our group suggested that P3a amplitude variations are related to the resolution of uncertainty (Kopp & Lange, 2013; Lange, Seer, Finke, Dengler, & Kopp, 2015).

Prior probabilities amounted to $P(h_1) = .7$ and $P(h_2) = .3$ in uncertain prior probability conditions, and they amounted to $P(h_1) = .9$ and $P(h_2) = .1$ in certain prior probability conditions. P3a amplitude variations in the urn-ball task reflected these variations in prior probabilities at the point in time when the initial stimuli s_1 in the samples of balls were processed. This finding establishes a connection between prior probabilities and the brain activations that we measured. Equation 3 tells us that the logarithm of the prior odds, that is, $\ln(P(h_1)/P(h_2))$, is the second summand that contributes to Bayesian inference. Given that P3b responses code logLR (see above), the discovery that P3a amplitude variations in the urn-ball task reflect variations in prior probabilities provides the missing

piece of evidence that the human brain codes *both* summands that are required for Bayesian inference.

Although one cannot draw strong conclusions from 20-channel electroencephalographic data about the underlying neural mechanisms, the present P3a findings do justify some propositions. Our P3a findings can be viewed as reflecting neural activities of a neural module, which traces prior probabilities (anterior probabilistic module), putatively in terms of log prior odds. The anterior probabilistic module is dissociable from the posterior probabilistic module, both in terms of the cortical areas that are involved and in terms of their functions: While the posterior probabilistic module is sensitive to logLR, but insensitive to log prior odds, the anterior probabilistic module is sensitive to log prior odds, but insensitive to logLR.

Prefrontal regions do contribute to the generation of scalp-recorded P3a wave shapes (Polich, 2007), suggesting that the anterior probabilistic module might be located in the prefrontal cortex. It is not possible to draw firm conclusions about the underlying neural mechanisms from the fact that the manipulation of prior probabilities exerted maximum effects at electrode F4 (i.e., right-lateralized). There exist some hints that the right cerebral hemisphere might be dominant for statistical processing in general (Anderson, 2008; Danckert, Stöttinger, Quehl, & Anderson, 2012; Roser, Fiser, Aslin, & Gazzaniga, 2011; Shaqiri & Anderson, 2013).

Implications and Suggestions

We found evidence for two dissociable probabilistic modules in the human brain that code instructed prior probabilities and likelihoods. Our ERP data demonstrate that all information necessary for Bayesian inference is coded by the human brain (i.e., that the *antecedent* conditions for Bayesian inference are met). The results suggest that the human brain computes Bayesian updating on the basis of these probabilistic quantities. A demonstration that the human brain computes Bayesian updating requires tracing the dynamic evolution of Bayesian probabilities and brain activations over the complete sequences of ball samples (Kolossa et al., 2015).

Another implication of our results is that the brain uses the information that was conveyed by the tableaux (i.e., instruction-based knowledge). Neither likelihoods nor prior probabilities could be learned through statistical learning (i.e., through accumulation of experience across episodes of sampling within probability conditions). In that regard it is worth mentioning that participants never received feedback as to the correctness of their urn choices. Hence, statistical learning yields an estimate of the expected value of relative frequencies of ball colors (plus sampling error), whereas neither likelihoods nor prior probabilities are observable quantities. This aspect of our results suggests that the viewing of the tableaux led to the formation of probabilistic models of ball

sampling, and that both probabilistic modules form part of these probabilistic models. The measured brain activations may disclose neural correlates of a model-based probabilistic system for guiding behavior at choice points (Dolan & Dayan, 2013; Doll, Duncan, Simon, Shohamy, & Daw, 2015; O'Doherty, Lee, & McNamee, 2015; Tolman, 1948).

Limitations and Future Directions

The SNR is an important, yet often neglected, factor in ERP research. As detailed in the Signal-to-Noise Ratio Estimation paragraph of the **Method** section, the current urn-ball task was associated with insufficient SNRs of some average ERPs because of insufficient numbers of trials for averaging (see Table 1). Insufficient SNRs led to the exclusion of all ERPs that were recorded in response to rarely occurring stimuli s_2 . Future urn-ball studies should take the SNR problem into account. In that regard, it might be of interest that Paukkunen, Leminen, and Sepponen (2010) suggested an adaptive method that is based on monitoring the SNR online on a trial-by-trial basis. Their method allows continuing ERP recordings until the SNR quality of an accumulating ERP average meets some predefined criterion, such that the sufficiency of the number of trials for ERP averaging can be better guaranteed.

The stability and validity of the results need to be examined through appropriate replication and generalization studies. *Decisive* support that the brain activations that we measured code prior probabilities and likelihoods was not achieved. The results at electrode F4 lent *substantial* support for a prior probability hypothesis, and the results at electrode Pz provided *strong* support for a likelihood hypothesis. Replication in an independent, preferentially larger sample is desirable. An examination of the generalization across different probability distributions and task domains (urn-ball problems vs. real-world-like problems) is warranted.

Future studies should aim to integrate these and related findings with the broader theoretical concepts that have been developed for an understanding of P300 amplitude variation. Polich (2007) attributed P3a amplitude variations to attentional mechanisms in the service of evaluating incoming stimuli, and subsequent studies should clarify how Bayesian and other schemes of information processing could be brought together to provide novel theoretical accounts of P3a amplitude variation. P3b amplitude variation is traditionally understood as being related to memory storage in the service of context updating (Donchin & Coles, 1988; Polich, 2007). It remains to be delineated how our findings can be reconciled with this more traditional view. P300 amplitude variation is known to be related to many other variables, such as the difficulty of a task (Kopp, Kizilirmak, Liebscher, Runge, & Wessel, 2010) or the specific temporal regimes on a task

(Gonsalvez, Barry, Rushby, & Polich, 2007). A comprehensive theory of P300 amplitude variation still needs to be developed (see Johnson, 1986; Polich, 2007, for corresponding attempts).

Conclusions

Our approach allowed us to assess relationships between antecedents of probabilistic inference and electrophysiological measures. Our results open new windows onto the neural correlates of probabilistic inference in humans by isolating distinct electrophysiological signatures of how the human brain processes instructed prior probabilities and likelihoods. These signals could be measured with a minimum of signal processing and with high temporal resolution. The influence of instructed prior probabilities and likelihoods is dissociable and may be mapped onto the classical distinction between anteriorly and posteriorly distributed scalp-recorded P300 waves. The picture that emerges is one of a Bayesian brain as revealed by the sensitivity of anteriorly distributed P3a waves for variations in prior probabilities, as well as by the sensitivity of posteriorly distributed P3b waves for variations in likelihoods.

Acknowledgments This work was supported by the Deutsche Forschungsgemeinschaft (DFG) via a Future Fund to T. F. (725 508 02). F. L. receives a PhD fellowship from the German National Academic Foundation. This work was also supported by a grant from the Netherlands Organization of Scientific Research (NWO-406-12-001).

Author contributions B. K. designed, performed and supervised the research. C. S., F. L., and A. Ko. analyzed the data. H. H. contributed new analytic tools. A. Kl. and H. H. contributed to the project after summaries of the raw data have been constructed. B. K., C. S., F. L., A. Kl., A. Ko., T. F., and H. H. wrote the manuscript.

References

- Achtziger, A., Alós-Ferrer, C., Hügelschäfer, S., & Steinhauser, M. (2014). The neural basis of belief updating and rational decision making. *Social Cognitive and Affective Neuroscience*, 9(1), 55–62. doi:10.1093/scan/nss099
- Anderson, B. (2008). Neglect as a disorder of prior probability. *Neuropsychologia*, 46(5), 1566–1569. doi:10.1016/j.neuropsychologia.2007.12.006
- Anderson, J. R. (2015). *Cognitive psychology and its implications* (8th ed.). New York, NY: Worth.
- Bach, D. R., & Dolan, R. J. (2012). Knowing how much you don't know: A neural organization of uncertainty estimates. *Nature Reviews Neuroscience*, 13(8), 572–586. doi:10.1038/nrn3289
- Barlow, H. (2001). Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12(3), 241–253. doi:10.1080/net.12.3.241.253
- beim Graben, P. (2001). Estimating and improving the signal-to-noise ratio of time series by symbolic dynamics. *Physical Review E*, 64(5), 051104. doi:10.1103/PhysRevE.64.051104

- Bennett, D., Murawski, C., & Bode, S. (2015). Single-trial event-related potential correlates of belief updating. *eNeuro*, 2(5), e0076–15.2015. doi:[10.1523/ENEURO.0076-15.2015](https://doi.org/10.1523/ENEURO.0076-15.2015)
- Constantino, S. M., & Daw, N. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, and Behavioral Neuroscience*, 15(4), 837–853. doi:[10.3758/s13415-015-0350-y](https://doi.org/10.3758/s13415-015-0350-y)
- d'Acremont, M., Schultz, W., & Bossaerts, P. (2013). The human brain encodes event frequencies while forming subjective beliefs. *Journal of Neuroscience*, 33(26), 10887–10897. doi:[10.1523/JNEUROSCI.5829-12.2013](https://doi.org/10.1523/JNEUROSCI.5829-12.2013)
- Danckert, J., Stöttinger, E., Quehl, N., & Anderson, B. (2012). Right hemisphere brain damage impairs strategy updating. *Cerebral Cortex*, 22(12), 2745–2760. doi:[10.1093/cercor/bhr351](https://doi.org/10.1093/cercor/bhr351)
- Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: Mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, 41(5), 665–678. doi:[10.1111/j.1469-8986.2004.00193.x](https://doi.org/10.1111/j.1469-8986.2004.00193.x)
- Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, 80(2), 312–325. doi:[10.1016/j.neuron.2013.09.007](https://doi.org/10.1016/j.neuron.2013.09.007)
- Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, 18(5), 767–772. doi:[10.1038/nn.3981](https://doi.org/10.1038/nn.3981)
- Donchin, E. (1981). Surprise! ... surprise? *Psychophysiology*, 18(5), 493–513. doi:[10.1111/j.1469-8986.1981.tb01815.x](https://doi.org/10.1111/j.1469-8986.1981.tb01815.x)
- Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11(3), 357–427.
- Doya, K., Ishii, S., Pouget, A., & Rao, R. P. N. (2007). *Bayesian brain: Probabilistic approaches to neural coding*. Cambridge, MA: MIT Press.
- Eimer, M. (1998). The lateralized readiness potential as an on-line measure of central response activation processes. *Behavior Research Methods, Instruments, & Computers*, 30(1), 146–156. doi:[10.3758/BF03209424](https://doi.org/10.3758/BF03209424)
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614), 1898–1902. doi:[10.1126/science.1077349](https://doi.org/10.1126/science.1077349)
- FitzGerald, T. H., Seymour, B., Bach, D. R., & Dolan, R. J. (2010). Differentiable neural substrates for learned and described value and risk. *Current Biology*, 20(20), 1823–1829. doi:[10.1016/j.cub.2010.08.048](https://doi.org/10.1016/j.cub.2010.08.048)
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1), 152–170. doi:[10.1111/j.1469-8986.2007.00602.x](https://doi.org/10.1111/j.1469-8986.2007.00602.x)
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 360(1456), 815–836. doi:[10.1098/rstb.2005.1622](https://doi.org/10.1098/rstb.2005.1622)
- Furl, N., & Averbeck, B. B. (2011). Parietal cortex and insula relate to evidence seeking relevant to reward-related decisions. *Journal of Neuroscience*, 31(48), 17572–17582. doi:[10.1523/JNEUROSCI.4236-11.2011](https://doi.org/10.1523/JNEUROSCI.4236-11.2011)
- Gerstner, W., & Fremaux, N. (2015). Neuromodulated spike-timing-dependent plasticity and theory of three-factor learning rules. *Frontiers in Neural Circuits*, 9, 85. doi:[10.3389/fncir.2015.00085](https://doi.org/10.3389/fncir.2015.00085)
- Glimcher, P. W. (2004). *Decisions, uncertainty, and the brain: The science of neuroeconomics*. Cambridge, MA: MIT Press.
- Gonzalez, C. J., Barry, R. J., Rushby, J. A., & Polich, J. (2007). Target-to-target interval, intensity, and P300 from an auditory single-stimulus task. *Psychophysiology*, 44(2), 245–250.
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55(4), 468–484. doi:[10.1016/0013-4694\(83\)90135-9](https://doi.org/10.1016/0013-4694(83)90135-9)
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.
- Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., & Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520, 220–223. doi:[10.1038/nature14066](https://doi.org/10.1038/nature14066)
- Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12), 517–523. doi:[10.1016/j.tics.2009.09.004](https://doi.org/10.1016/j.tics.2009.09.004)
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8), 534–539. doi:[10.1111/j.0956-7976.2004.00715.x](https://doi.org/10.1111/j.0956-7976.2004.00715.x)
- Hofjink, H. (2012). *Informative hypotheses: Theory and practice for the behavioral and social scientists*. Boca Raton, FL: Chapman and Hall CRC Press.
- Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge, UK: Cambridge University Press.
- Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23(4), 367–384.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kappel, D., Habenschuss, S., Legenstein, R., & Maass, W. (2015). Network plasticity as Bayesian inference. *PLoS Computational Biology*, 11(11), e1004485. doi:[10.1371/journal.pcbi.1004485](https://doi.org/10.1371/journal.pcbi.1004485)
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795. doi:[10.1080/01621459.1995.10476572](https://doi.org/10.1080/01621459.1995.10476572)
- Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience*, 33(50), 19434–19441. doi:[10.1523/JNEUROSCI.3355-13.2013](https://doi.org/10.1523/JNEUROSCI.3355-13.2013)
- Kira, S., Yang, T., & Shadlen, M. N. (2015). A neural implementation of Wald's sequential probability ratio test. *Neuron*, 85(4), 861–873. doi:[10.1016/j.neuron.2015.01.007](https://doi.org/10.1016/j.neuron.2015.01.007)
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), 712–719. doi:[10.1016/j.tins.2004.10.007](https://doi.org/10.1016/j.tins.2004.10.007)
- Kolossa, A. (2016). *Computational modeling of neural activities for statistical inference*. Cham, Switzerland: Springer.
- Kolossa, A., Fingscheidt, T., Wessel, K., & Kopp, B. (2013). A model-based approach to trial-by-trial P300 amplitude fluctuations. *Frontiers in Human Neuroscience*, 6, 359. doi:[10.3389/fnhum.2012.00359](https://doi.org/10.3389/fnhum.2012.00359)
- Kolossa, A., Kopp, B., & Fingscheidt, T. (2015). A computational analysis of the neural bases of Bayesian inference. *NeuroImage*, 106, 222–237. doi:[10.1016/j.neuroimage.2014.11.007](https://doi.org/10.1016/j.neuroimage.2014.11.007)
- Kopp, B. (2008). The P300 component of the event-related brain potential and Bayes' theorem. In M. K. Sun (Ed.), *Cognitive sciences at the leading edge* (pp. 87–96). New York, NY: Nova Science.
- Kopp, B., & Lange, F. (2013). Electrophysiological indicators of surprise and entropy in dynamic task-switching environments. *Frontiers in Human Neuroscience*, 7, 300. doi:[10.3389/fnhum.2013.00300](https://doi.org/10.3389/fnhum.2013.00300)
- Kopp, B., Kizilirmak, J., Liebscher, C., Runge, J., & Wessel, K. (2010). Event-related brain potentials and the efficiency of visual search for vertically and horizontally oriented stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, 10(4), 523–540.
- Lange, F., Seer, C., Finke, M., Dengler, R., & Kopp, B. (2015). Dual routes to cortical orienting responses: Novelty detection and uncertainty reduction. *Biological Psychology*, 105, 66–71. doi:[10.1016/j.biopsycho.2015.01.001](https://doi.org/10.1016/j.biopsycho.2015.01.001)
- MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge, UK: Cambridge University Press.
- Maris, E. (2012). Statistical testing in electrophysiological studies. *Psychophysiology*, 49(4), 549–565. doi:[10.1111/j.1469-8986.2011.01320.x](https://doi.org/10.1111/j.1469-8986.2011.01320.x)

- Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *Journal of Neuroscience*, 28(47), 12539–12545. doi:10.1523/JNEUROSCI.2925-08.2008
- Möcks, J., Gasser, T., & Köhler, W. (1988). Basic statistical parameters of event-related potentials. *Journal of Psychophysiology*, 2(1), 61–70.
- Mulder, J., Hooijink, H., & de Leeuw, C. (2012). BIEMS: A Fortran90 program for calculating Bayes factors for inequality and equality constrained models. *Journal of Statistical Software*, 46(2), 1–39. Retrieved from <http://www.jstatsoft.org/v46/i02>
- O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience*, 15(12), 1729–1735. doi:10.1038/nn.3248
- O'Connor, K. (1985). The Bayesian-inferential approach to defining response processes in psychophysiology. *Psychophysiology*, 22(4), 464–479. doi:10.1111/j.1469-8986.1985.tb01633.x
- O'Doherty, J. P., Lee, S. W., & McNamee, D. (2015). The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, 1, 94–100. doi:10.1016/j.cobeha.2014.10.004
- O'Reilly, J. X., Schüffelen, U., Cuell, S. F., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 110(38), E3660–E3669. doi:10.1073/pnas.1305373110
- Ostwald, D., Spitzer, B., Guggenmos, M., Schmidt, T. T., Kiebel, S. J., & Blankenburg, F. (2012). Evidence for neural encoding of Bayesian surprise in human somatosensation. *NeuroImage*, 62(1), 177–188. doi:10.1016/j.neuroimage.2012.04.050
- Paukkunen, A. K., Leminen, M. M., & Sepponen, R. (2010). Development of a method to compensate for signal quality variations in repeated auditory event-related potential recordings. *Frontiers in Neuroengineering*, 3, 2. doi:10.3389/fneng.2010.00002
- Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72(3), 346–354. doi:10.1037/h0023653
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. doi:10.1016/j.clinph.2007.04.019
- Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: Knowns and unknowns. *Nature Neuroscience*, 16(9), 1170–1178. doi:10.1038/nn.3495
- Preusschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, 51(3), 381–390. doi:10.1016/j.neuron.2006.06.024
- Preusschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, 28(11), 2745–2752. doi:10.1523/JNEUROSCI.4286-07.2008
- Robert, C. P. (2007). *The Bayesian choice: From decision-theoretic foundations to computational implementation*. New York, NY: Springer.
- Roser, M. E., Fiser, J., Aslin, R. N., & Gazzaniga, M. S. (2011). Right hemisphere dominance in visual statistical learning. *Journal of Cognitive Neuroscience*, 23(5), 1088–1099. doi:10.1162/jocn.2010.21508
- Schwartenbeck, P., FitzGerald, T. H., & Dolan, R. (2016). Neural signals encoding shifts in beliefs. *NeuroImage*, 125, 578–586. doi:10.1016/j.neuroimage.2015.10.067
- Seer, C., Lange, F., Boos, M., Dengler, R., & Kopp, B. (2016). Prior probabilities modulate cortical surprise responses: A study of event-related potentials. *Brain & Cognition*, 106, 78–89.
- Shaqiri, A., & Anderson, B. (2013). Priming and statistical learning in right brain damaged patients. *Neuropsychologia*, 51(13), 2526–2533. doi:10.1016/j.neuropsychologia.2013.09.024
- Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 193(4258), 1142–1146. doi:10.1126/science.959831
- Stern, E. R., Gonzalez, R., Welsh, R. C., & Taylor, S. F. (2010). Updating beliefs for a decision: Neural correlates of uncertainty and underconfidence. *Journal of Neuroscience*, 30(23), 8032–8041. doi:10.1523/JNEUROSCI.4729-09.2010
- Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron*, 59(2), 336–347. doi:10.1016/j.neuron.2008.05.021
- Sutton, S., & Ruchkin, D. S. (1984). The late positive complex. *Annals of the New York Academy of Sciences*, 425(1), 1–23. doi:10.1111/j.1749-6632.1984.tb23520.x
- Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700), 1187–1188. doi:10.1126/science.150.3700.1187
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189–208. doi:10.1037/h0061626
- Trapp, S., Lepsien, J., Kotz, S. A., & Bar, M. (2016). Prior probability modulates anticipatory activity in category-specific areas. *Cognitive, Affective, and Behavioral Neuroscience*, 16(1), 135–144. doi:10.3758/s13415-015-0373-4
- Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A., & Körding, K. P. (2012). Differential representations of prior and likelihood uncertainty in the human brain. *Current Biology*, 22(18), 1641–1648. doi:10.1016/j.cub.2012.07.010
- Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, 447(7148), 1075–1080. doi:10.1038/nature05852