



Electroencephalographic correlates of temporal Bayesian belief updating and surprise

Antonino Visalli^{a,b,*}, Mariagrazia Capizzi^a, Ettore Ambrosini^{b,c}, Bruno Kopp^d,
Antonino Vallesi^{c,e,*}

^a Department of Neuroscience, University of Padova, 35128 Padova, Italy

^b Department of General Psychology, University of Padova, 35131 Padova, Italy

^c Department of Neuroscience & Padova Neuroscience Center, University of Padova, 35131 Padova, Italy

^d Department of Neurology, Hannover Medical School, 30625 Hannover, Germany

^e Brain Imaging and Neural Dynamics Research Group, IRCCS San Camillo Hospital, 30126 Venice, Italy

ARTICLE INFO

Keywords:

Bayesian brain
Foreperiod paradigm
Hazard function
Surprise
Belief updating
P3

ABSTRACT

The brain predicts the timing of forthcoming events to optimize responses to them. Temporal predictions have been formalized in terms of the hazard function, which integrates prior beliefs on the likely timing of stimulus occurrence with information conveyed by the passage of time. However, how the human brain updates prior temporal beliefs is still elusive. Here we investigated electroencephalographic (EEG) signatures associated with Bayes-optimal updating of temporal beliefs. Given that updating usually occurs in response to surprising events, we sought to disentangle EEG correlates of updating from those associated with surprise. Twenty-six participants performed a temporal foreperiod task, which comprised a subset of surprising events not eliciting updating. EEG data were analyzed through a regression-based massive approach in the electrode and source space. Distinct late positive, centro-parietally distributed, event-related potentials (ERPs) were associated with surprise and belief updating in the electrode space. While surprise modulated the commonly observed P3b, updating was associated with a later and more sustained P3b-like waveform deflection. Results from source analyses revealed that neural encoding of surprise comprises neural activity in the cingulo-opercular network (CON) and parietal regions. These data provide evidence that temporal predictions are computed in a Bayesian manner, and that this is reflected in P3 modulations, akin to other cognitive domains. Overall, our study revealed that analyzing P3 modulations provides an important window into the Bayesian brain. Data and scripts are shared on OSF: <https://osf.io/ckqa5/>

1. Introduction

The Bayesian brain hypothesis assumes that brain functions can be modeled as Bayesian inference (Doya et al., 2007; Friston, 2012; Kersten et al., 2004; Knill and Pouget, 2004). Bayesian inference is very effective in modeling information processing under uncertainty, and it represents a general framework increasingly being used for investigation of human perception, cognition and action (Chater et al., 2010; De Lange et al., 2018; Friston et al., 2017). More specifically, the Bayesian brain hypothesis supposes that the brain represents information as posterior probabilities, which are derived by integrating likelihoods of newly gained information with prior probabilities, that is, previously held (prior) beliefs. A general brain function hence is to *update* prior beliefs to build accurate predictive models that minimize discrepancies (i.e., *surprise*) – which may impose cognitive costs (Zénon et al., 2019) – between newly gained information and prior beliefs.

The current study focuses on a fundamental aspect of perception that until now has received much less attention within the Bayesian framework as compared to other perceptual processes, namely, *temporal expectations* (Coull, 2009; Niemi and Naatanen, 1981; Vallesi, 2010). Temporal expectations can be considered as beliefs about the timing of occurrence of relevant events. One explanation for the paucity of Bayesian research on this key ability could be ascribed to the success of previous literature in modeling temporal expectations in terms of the hazard function. Hazard functions describe the change in the probability that an event will occur given it has not yet occurred (Nobre and van Ede, 2018). To illustrate the hazard function, consider a task in which the time interval (foreperiod, FP) between a warning signal and a target can assume one of three equally probable durations [$P(FP_{short}) = P(FP_{medium}) = P(FP_{long}) = .33$]. If the target does not appear at the short FP, the probability of its occurrence at the medium or long FP will change to .5, and if still it does not appear at the medium FP,

* Corresponding authors at: Department of Neuroscience, University of Padova, 35128 Padova, Italy
E-mail addresses: antonino.visalli@unipd.it (A. Visalli), antonino.vallesi@unipd.it (A. Vallesi).

<https://doi.org/10.1016/j.neuroimage.2021.117867>

Received 29 September 2020; Received in revised form 2 February 2021; Accepted 9 February 2021

Available online 13 February 2021

1053-8119/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

the probability of its occurrence at the long FP will equal 1. As long as the (generative) probability of target occurrence remains stable, hazard functions are valid models for describing both behavioral and neural responses in FP paradigms (Janssen and Shadlen, 2005; Buetti et al., 2010; Herbst et al., 2018). However, hazard functions do not capture potential changes in the generative probability of FP durations. As a consequence, hazard functions are not well suited to model updating of temporal beliefs¹.

In a previous study, we combined a Bayesian computational approach with functional magnetic resonance imaging (fMRI) in an attempt to understand neural correlates of temporal belief updating (Visalli et al., 2019). To that aim, we implemented a FP paradigm (adapted from O'Reilly et al., 2013) that was designed to identify separable neural correlates of belief updating and surprise. Here, by surprise we refer to the information content of a particular event, such that the lower its probability of occurrence is, the higher the level of surprise is. Under common conditions, more surprising events typically give rise to strong changes in beliefs. However, updating and surprise reflect qualitatively different types of information (see Methods section 2.4, for further details). Belief updating is often quantified by the Kullback-Leibler divergence (D_{KL} ; Baldi and Itti, 2010), whereas Shannon's Information quantifies surprise (I_S ; Itti and Baldi, 2009)². Although these two measures are usually highly correlated, surprising events do not necessarily trigger belief updating (see the "white snow paradox" in Itti and Baldi, 2009, and the "two models" example in Barto et al., 2013). The extent to which a surprising event induces a change in beliefs depends, for example, on the precision (confidence or inverse uncertainty) of our beliefs (Courville et al., 2006; Mathys et al., 2011), the volatility of the environment (Behrens et al., 2007; Nassar et al., 2010) or reward (McGuire et al., 2014). Furthermore, it is also possible to imagine scenarios in which belief updating might occur after events with very low surprise (O'Reilly et al., 2013), a phenomenon that can be more easily appreciated in a FP paradigm. Indeed, a target that occurs after a FP with low a priori probability potentially induces a change in beliefs. At the same time, if the target appears later in time, even if the associated a priori probability is very low, its level of surprise is also low, since probability will increase during the trial as a function of the elapsed time (i.e., high hazard rate). Going back to our previous study (Visalli et al., 2019), to lower the correlation between updating and surprise, which is usually high in experimental paradigms, our FP paradigm introduced an experimental manipulation for which a subset of temporally surprising events not requiring belief updating emerged (see the Methods Section 2.2, for details). Consistent with other recent studies (Kobayashi and Hsu, 2017; Schwartenbeck et al., 2016), our fMRI results (see the Methods Section 2.7.2, for details) showed that updating of temporal beliefs and surprise modulated trial-by-trial activity in two key brain networks supporting top-down cognitive control (i.e. the ability to coordinate information processing for the pursuit of current goals; Cocchi et al., 2013) in a multitude of cognitive tasks: the fronto-parietal

¹ The terms *belief* (more common in the Bayesian brain literature) and *expectation* (more common in the FP literature) are usually considered as synonyms. Hereafter, we will use *temporal beliefs* to refer to prior beliefs about FP duration, whereas *expectations* will refer to beliefs about FP duration modulated within a trial according to the hazard function.

² The D_{KL} between an observer's prior and posterior beliefs is sometimes referred to as *Bayesian Surprise*, whereas the negative log probability (I_S) of a particular event as *Surprisal*. As reported in the main text, although the terminology might be confusing, I_S and D_{KL} have distinct computational definitions: for the former definition a highly unlikely event has high surprise, whereas for the latter definition the degree of surprise of an event depends on how much it changes internal beliefs.

(FPN) and the cingulo-opercular (CON) networks (Dosenbach et al., 2008). Overall, our results showed that Bayesian inference is suitable for modeling temporal belief updating and, in line with the above-mentioned studies (Kobayashi and Hsu, 2017; Schwartenbeck et al., 2016), they provided further evidence for casting top-down control networks within a Bayesian brain framework (Pezzulo, 2012).

In the present study, we moved a step forward using electroencephalography (EEG) to overcome the poor temporal resolution of fMRI and to better describe the dynamics of neural processes of temporal updating and surprise. Following the EEG-based research on Bayesian belief updating (see Kopp et al., 2016, for an overview), we predicted to differentiate updating and surprise at the level of the P3 event-related potential (ERP) family (Donchin and Coles, 1988; Polich, 2007; Sutton et al., 1965). It is a well-replicable fact that P3 amplitudes are inversely related to stimulus probabilities (Donchin and Coles, 1988). Therefore, notions of expectancy, uncertainty, subjective probability, surprise, or updating were introduced as functional hypotheses of the P3 early on (Donchin, 1981; Squires et al., 1976; Sutton et al., 1965). A major shortcoming of the traditional P3 literature lies in the fact that these functional notions remained undefined in formal terms. The relevant research needs to be qualified as being non-computational because P3 research was grounded on the goal of semantic agreement, which was never achieved (see the controversy between Verleger (1988) and Donchin and Coles (1988) as a prime example). Interest in P3 research decreased in the years after 1988 until the rise of computational formalization in cognitive neuroscience revitalized the field. Our study is another attempt to understand P3 amplitude fluctuations from the vantage point of the Bayesian brain hypothesis. Explicit linkage between Bayesian inference and P3 amplitude variability did only appear in relatively recent years (Kopp, 2008; Friston, 2005). Since then, a relatively small number of studies relied on computational formalization and modeling of functional P3 hypotheses (e.g., Mars et al., 2008; Kolossa et al. 2013, 2015). A consistent finding emerging from these studies is that distinct P3 sub-components can be differentiated in terms of the above-mentioned D_{KL} and I_S measures. In particular, a parietal scalp-distributed P3b has been associated with I_S (Kolossa et al., 2013; Mars et al., 2008), whereas an earlier and more anteriorly scalp-distributed P3a has been linked to D_{KL} (Bennett et al., 2015; Kolossa et al., 2015). Of note, none of these previous Bayesian P3 studies was designed to differentiate between belief updating and surprise. Using the same task employed in our previous study (Visalli et al., 2019), we addressed this limitation in the present EEG study by applying an appropriate experimental manipulation.

All previous Bayesian P3 studies investigated correlations between D_{KL} and/or I_S and EEG single-trial amplitude measures extracted from time windows and electrodes that were selected on the basis of some a-priori assumptions and/or inspection of grand-average ERP waveforms. The main disadvantage of such an approach is that it does not provide temporally detailed information about the waveforms of interest (Smith and Kutas, 2015). Indeed, neural activities related to belief updating and surprise may occur in close spatial and temporal proximity, potentially overlapping in scalp-recorded resulting EEG signals. Finally, other modulations outside the P3 latency range and/or its typical scalp topography could be associated with updating and surprise, but these potential ERP modulations may have been overlooked in those previous Bayesian P3 studies.

To overcome these limitations, we adopted a regression-based massive approach that considered the entire spatio-temporal EEG recordings in the present study (Ehinger and Dimigen, 2019). Moreover, we performed source reconstruction of the EEG signal within some regions of interest (ROIs) selected on the basis of our related fMRI study (Visalli et al., 2019) in order to gather further information on our previous findings.

In sum, the present Bayesian ERP study targeted three goals: (i) To identify ERP correlates of temporal belief updating and distinguish them from ERP correlates of surprise through appropriate experimental manipulations; (ii) To link the results from previous Bayesian P3 studies with those from the present study on temporal belief updating. At the same time, by applying some theoretical (i.e., the distinction between updating and surprise through appropriate experimental manipulation) and methodological (i.e., mass analysis of EEG signals) refinements, we tried to improve the analysis and the computationally informed understanding of the P3; (iii) To further scrutinize our previous fMRI finding showing the involvement of cognitive-control brain networks in temporal belief updating and encoding of surprise by the analysis of EEG signals in the source space.

2. Methods

2.1. Participants

The study included an initial sample of 26 participants. One participant was excluded due to low compliance with task instructions (22% of premature/missing responses) and replaced with an additional participant. Therefore, the final sample still comprised 26 participants [10 males; mean age: 23.4 (SD = 3), range: 19-33 years old]. All of them were right-handed [Edinburgh Handedness Inventory (Oldfield, 1971) average score: 81.8 (SD = 16)] except one who showed weak right-handedness (EHI = 20). An a-priori sensitivity power analysis (G*Power 3 software; Faul et al., 2007) revealed that our sample size was large enough to have a statistical power of .80 to detect the significant ($\alpha = .05$) effects of interest with a medium effect size $d = .50$ (Cohen, 1988), as estimated from our previous study (Visalli et al., 2019).

All participants reported no current (or history of) neurological or psychiatric disorders, normal color vision and normal or corrected-to-normal visual acuity. The procedures involved in this study were approved by the Bioethical Committee of the Azienda Ospedaliera di Padova. Participants gave their written informed consent before the experiment, in accordance with the Declaration of Helsinki, and were reimbursed 20 Euros for their time.

2.2. Task and procedure

The experimental paradigm – the same variable FP task used in our previous study (Visalli et al., 2019; see also O'Reilly et al., 2013) – was implemented in MATLAB (The MathWorks, Inc., Natick, Massachusetts, United States) using the PSYCHOPHYSICS TOOLBOX 3 (Brainard and Vision, 1997; Kleiner et al., 2007). The task is available in the shared folder *task.zip*. On each trial, participants had to respond as quickly as possible to the onset of a target that was separated from a neutral warning signal by variable FPs (Fig. 1A). The warning signal was a black fixation cross displayed centrally against a gray background for the whole FP duration. The target was a colored circle with a diameter equal to the length of the cross arms and centrally displayed for 1000 ms. Participants responded by pressing the spacebar with their index finger (half of the participants used their right hand, whereas the other half their left hand). The inter-trial interval was randomly jittered between 1250 and 1750 ms (rectangular distribution). The onset of the majority of the targets (80% of 790 trials) was predictable (predictable trials), since the target appeared after a FP drawn from a Gaussian distribution (normal FP) having constant mean and standard deviation during a block of trials (Fig. 1A and 1B). However, FP means and standard deviations changed abruptly between blocks of trials. The blocks were not temporally separated, and participants were explicitly instructed that a change in target color (vermillion, reddish purple, bluish green, or blue) signaled the beginning of a new block (updating trial). On a few trials (20%; uniform trials), interspersed with the other trial types, the target appeared after a FP drawn from a uniform distribution ranging from 250 to 2500

ms (half of them had a FP below the expected mean and the other half above). Therefore, the generative probability density function over FP duration for each block was:

$$p(FP) = .80 p(FP|FP \sim N(\mu, \sigma)) + .20 p(FP|FP \sim U(250 \text{ ms}, 2500 \text{ ms})) \quad (1)$$

that is, the sum of the probability of normal FPs, $p(FP|FP \sim N(\mu, \sigma))$, scaled by 80% and the probability of uniform FPs, $p(FP|FP \sim U(250 \text{ ms}, 2500 \text{ ms}))$, scaled by 20%.

Importantly, “uniform” targets were always white, signaling that the current trial was not igniting a new block. Although update and uniform trials were both highly surprising, participants were instructed to use the color information in order to avoid updating after uniform trials. Summarizing, the color manipulation created three types of trials: (i) “predictable” trials in which the target onset could be easily predicted; (ii) “updating” trials in which target onsets were generally surprising (lower hazard rate) and led to strongly update the internal model; (iii) “uniform” trials that, like updating trials, were generally surprising, but in which updating was minimized (see Supplementary Figure S1, for a behavioral confirmation of task strategy). It follows that the color manipulation allowed having a subset of surprising events not eliciting updating in order to partially differentiate between processes involved in trial-to-trial belief updating and within-trial preparatory processes to the current trial. The number of blocks was 72 (plus an initial familiarization block, which was excluded from analyses). In each block, predictable FPs were drawn from a different Gaussian distribution, whose parameters were derived from an orthogonal combination of 9 means (500, 700, 900, 1100, 1300, 1500, 1700, 1900, and 2100 ms) and 8 standard deviations (25, 50, 75, 100, 125, 150, 175 and 200 ms). Each block included a number of trials ranging between 8 and 13 (mean: 10.5). In half of the blocks, the Gaussian distribution had a mean at least 3 SD higher than the previous one, whereas in the other half it was at least 3 SD lower. The blocks were grouped in four runs with self-paced breaks in between. Participants were informed that each new run started from the same FP distribution as that used in the last block of the previous run (trials of a new block presented at the beginning of a new run were discarded from analyses). Before the EEG cap application, participants familiarized themselves with the task. They performed a shorter version of the experimental task comprising three runs of two blocks each. In the first two runs, they were presented only with normal FP, in order to realize the arbitrary association between colors and FP distributions and to be made aware that after a pause (beginning of a new run) the new block resumed the same FP distribution as in the last block. The third practice run introduced uniform FPs and the difference between them and the normal ones was carefully explained to participants.

2.3. Ideal Bayesian observer

The model was introduced in Visalli and colleagues (2019; see also O'Reilly et al., 2013) and is now available in the shared folder *bayesian_model*. Briefly, on a trial-by-trial basis, the model estimates the posterior probability of putative pairs of μ and σ values of the parameters of the Gaussian distribution from which normal FPs are drawn. The range of possible values of μ and σ (parameter space) considered by the model is 300×25 (respectively, μ from 10 to 3000 ms in steps of 10 ms, and σ from 10 to 250 ms in steps of 10 ms). Following task instructions given to participants, the model updates prior beliefs according to the type of trial.

Predictable trial. The posterior at trial n is calculated with Bayes' rule by multiplying the likelihood of the new observation with posterior at trial $n-1$:

$$p(FP \sim N(\mu_n, \sigma_n) | FP_{1:n}) \propto p(FP_n | FP \sim N(\mu, \sigma)) \times p(FP \sim N(\mu_n, \sigma_n) | FP_{1:n}) \quad (2)$$

where $p(FP \sim N(\mu_n, \sigma_n) | FP_{1:n})$ indicates the posterior probability of each μ and σ parameter pair at trial n , $p(FP_n | FP \sim N(\mu, \sigma))$ indicates the likeli-

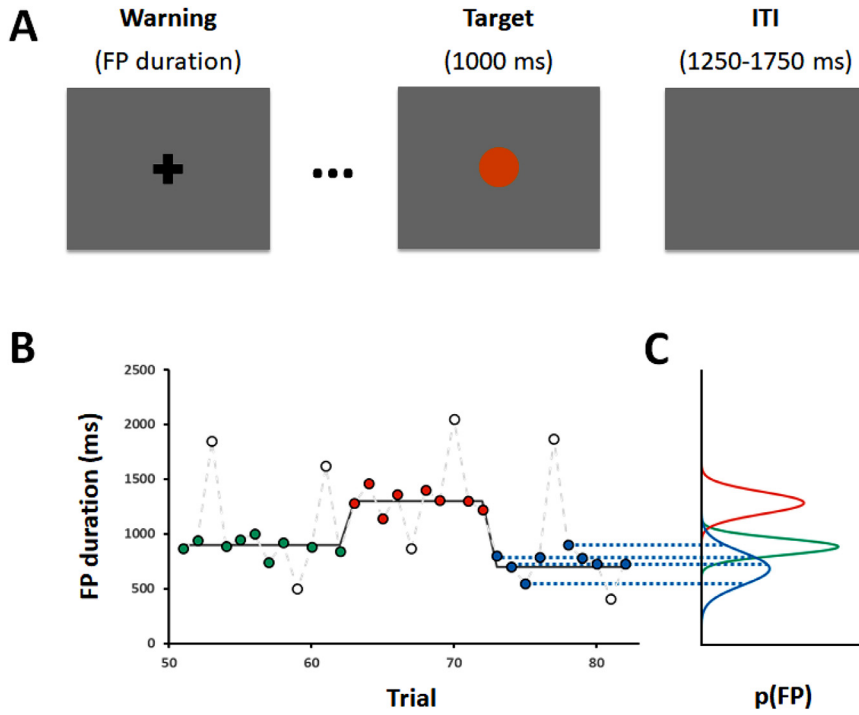


Fig. 1. Depiction of the task paradigm. (A) Example of a trial: Participants had to respond to target onsets that were separated from a neutral warning signal by variable FPs. (B) FP durations in three exemplary blocks. In most of the trials (i.e., 80%), target onsets were predictable since they occurred after FP (colored dots) drawn from a Gaussian distribution (as suggested by dotted lines connecting normal-FP values of the third block in panel B with the corresponding Gaussian distribution in panel C). FP mean (solid grey line) and standard deviation were kept constant within a block. On 20% of trials, targets occurred after FP (white dots) extracted from a uniform distribution. (C) Gaussian distributions over FP duration in the three exemplary blocks.

hood of the observed FP duration, and $p(\text{FP} \sim N(\mu_n, \sigma_n) | \text{FP}_{1:n-1})$ indicates the posterior probability at trial $n-1$.

Uniform trial. The posterior at trial n is derived from the posterior at trial $n-1$ without modification.

Update trial. The posterior at trial n is calculated with Bayes' rule by multiplying the likelihood of the new observation with a uniform distribution over parameter space:

$$p(\text{FP} \sim N(\mu_n, \sigma_n) | \text{FP}_{1:n}) \propto p(\text{FP}_n | \text{FP} \sim N(\mu, \sigma)) \times p(\text{FP} \sim U(\mu, \sigma)) \quad (3)$$

where $p(\text{FP} \sim U(\mu, \sigma))$ indicates the uniform distribution over the considered parameter space.

Posteriors over parameter space are then translated into probability density functions over normal FP duration by computing a sum of all possible Gaussian distributions whose μ , σ belonged to the parameter space weighted according to their posterior probability:

$$p(\text{FP} | \text{FP}_{1:n}) = \sum_{\mu, \sigma} p(\text{FP} | \text{FP} \sim N(\mu, \sigma)) \times p(\text{FP} \sim N(\mu_n, \sigma_n) | \text{FP}_{1:n}) \quad (4)$$

where $p(\text{FP} | \text{FP} \sim N(\mu, \sigma))$ is the probability distribution over FP duration for each Gaussian belonging to the parameter space.

At the beginning of each trial, participants know they can be presented with targets appearing either after predictable FPs or after random FPs (uniform/update trials). Therefore, a prior probability density function over FP duration for trial $n+1$ is obtained as a weighted mixture between the posterior Gaussian distribution over time at trial n and a uniform distribution:

$$p(\text{FP}_{n+1} | \text{FP}_{1:n}) = p(\text{predictable}_{nb+1}) \cdot p(\text{FP} | \text{FP}_{1:n}) + p(\text{uniform}_{nb+1} \cup \text{update}_{nb+1}) \cdot p(\text{FP} | U(10, 3000 \text{ ms})) \quad (5)$$

where $p(\text{predictable}_{nb+1})$ and $p(\text{uniform}_{nb+1} \cup \text{update}_{nb+1})$ represent the probability of incurring, respectively, in a predictable or in a uniform/update trial at the next trial of the current block (nb indicates the trial number within a block, which differs from n that indicates

trial number referred to the whole task). For simplicity, $p(\text{uniform}_{nb+1} \cup \text{update}_{nb+1})$ was the true proportion of incurring, respectively, in a predictable or a uniform/update trial at each trial within a block, smoothed with a moving average in order to have a monotonic increase in the probability of having an update trial during a block. Concerning $p(\text{predictable}_{nb+1})$, it was $1 - p(\text{uniform}_{nb+1} \cup \text{update}_{nb+1})$.

2.4. Model-based measures of updating and surprise

As mentioned in the Introduction, two measures from information theory were used to formally quantify updating of temporal expectations and surprise. Such information theory measures were computed with reference to the ideal Bayesian observer, and used as regressors for both behavioral and EEG analyses.

Following *Itti and Baldi (2009)*, updating was trial-wise quantified in terms of the Kullback-Leibler divergence (D_{KL}) between prior and posterior on each trial:

$$D_{KL} = \sum_{\text{FP}=10}^{3000} p(\text{FP} | \text{FP}_{1:n-1}) \log \frac{p(\text{FP} | \text{FP}_{1:n-1})}{p(\text{FP} | \text{FP}_{1:n})} \quad (6)$$

where the prior $p(\text{FP} | \text{FP}_{1:n-1})$ and the posterior $p(\text{FP} | \text{FP}_{1:n})$ are defined analogously as in Eq. 5.

Since during a trial, the prior probability of target onset changed as a function of the elapse of time (*Janssen and Shadlen, 2005*), surprise was quantified as the Shannon information (I_S), i.e. the negative log value of the hazard function at target onset:

$$I_S = -\log \frac{p(\text{FP}_n | \text{FP}_{1:n-1})}{1 - p(\text{FP} < \text{FP}_n | \text{FP}_{1:n-1})} \quad (7)$$

where $p(\text{FP}_n | \text{FP}_{1:n-1})$ is the probability of FP duration at trial n given the prior, and $p(\text{FP} < \text{FP}_n | \text{FP}_{1:n-1})$ is cumulative probability up to that FP. The correlation between D_{KL} and I_S was $r = 0.30$ ($R^2 = 0.09$). An output of the model and the resulting model-based measures of updating and surprise is presented in Supplementary Figure S2.

2.5. RT data analysis

Raw data and scripts used for this analysis are available in the shared folder *behavioral_analysis*.

The behavioral analysis is the same as the one conducted in Visalli and colleagues (2019). RTs were log-transformed to mitigate the influence of non-normally distributed and skewed data. Log-RTs were then analyzed by means of Linear Mixed Models (LMM) using the *lme4* library (Bates et al., 2015) in R (<http://www.R-project.org/>). Data from error trials (i.e., anticipated and/or missing response to the target) and post-error trials were not included [mean percentage of excluded trials: 4.4% (SD = 3.2%), range: 0.5-14.3%]. First, we specified a full LMM as the following Wilkinson-notation (Wilkinson and Rogers, 1973) formula:

$$\log(RT) \sim Trial + Preceding RT + D_{KL} * I_S + (Trial + Preceding RT + D_{KL} + I_S | ID) \quad (8)$$

In detail, we entered D_{KL} and I_S (and their interaction) along with the rank-order of each trial (Trial) and the log-RT at the preceding trial (Preceding RT) as fixed-effect terms. The random structure (between parentheses in Eq. 8) included correlated by-subject random intercepts and by-subject random slopes for Trial, Preceding RT, I_S and D_{KL} . All these continuous predictors were standardized using Z-score in order to facilitate model convergence. Following Baayen and Milin (2010), the variables Trial and Preceding RT were included to control for the temporal dependencies between successive trials that generally need to be addressed when modeling RTs. Specifically, TRIAL was included to capture possible effects of learning and fatigue, while PRECEDING RT was used to capture the trial-by-trial RT autocorrelation. Model selection procedure from the full model to a minimal adequate model was conducted through the function *step* of the *lmerTest* R-library (Kuznetsova et al., 2017), which performs backwards elimination of non-significant random and fixed effects of LMM (Kuznetsova et al., 2015).

2.6. EEG data acquisition

The EEG was recorded using BrainAmp amplifiers (Brain Products, Munich, Germany) from 64 Ag/AgCl electrodes that were mounted on an elastic cap (EASYCAP GmbH, Germany) according to the international 10-10-system. Electrooculographic activity was recorded through an electrode placed under the right eye and was also monitored through the scalp electrodes placed in the proximity of both eyes. Before recording, impedance for all electrodes was checked and adjusted until it was lower than 10 k Ω before testing. All electrodes were referenced to FCz during the recording, and an electrode positioned at AFz served as the ground. EEG activity was digitized at a sampling rate of 500 Hz.

2.7. EEG data analysis

Raw EEG datasets, data and scripts used for the EEG data analysis are available in the shared folder *eeg_analysis*.

2.7.1. EEG pre-processing

Offline processing of EEG signal was performed using custom MATLAB scripts, which included functions from the EEGLAB environment (version 14.1.2b, Delorme and Makeig, 2004) and the *FastICA* algorithm (Hyvarinen and Oja, 2000).

As a preliminary step for ICA decomposition (used for EEG artefact removal), a band-pass filter using a one-pass non-causal zero-phase Kaiser windowed sinc FIR filter (cut-off frequencies = 2 and 40 Hz; beta = 5.6526; transition bandwidth = 4 and 20 Hz for the high- and low-pass filters, respectively) was applied to continuous EEG data. The high cut-off for the high-pass filter was applied to remove low-frequency drifts in order to improve the results of ICA (Winkler et al., 2015). The *clean_rawdata* function was used to remove noisy channels (channel criterion = .8) and short-time bursts (burst criterion = 20 SD) from the

data. A maximum of 3 channels per subject (mean = 0.6, SD = 1) were removed. Then, the *FastICA* algorithm was applied to obtain ICA weight matrices and sphering matrices.

Since the use of a high-pass filter with such a high (2-Hz) cut-off may attenuate ERP effects and introduce distortions (Tanner et al., 2015), the ICA solution calculated on the 2-Hz high-pass filtered data was then applied on continuous EEG data band-pass filtered with a high-pass cut-off frequency of 0.1 (transition bandwidth = .2). Before applying the ICA solution, noisy channels identified in 2-Hz high-pass filtered data were removed. Subsequently, the EEGLAB plugin *SASICA* (Chaumon et al., 2015) was used to guide the identification and exclusion of artifactual ICA components (e.g., blinks, eye movements, muscle activity, misconnected channels). Finally, removed channels were interpolated using spherical splines (Perrin et al., 1989) and continuous EEG data were re-referenced to the average of all of the EEG electrodes adding FCz channel back to the data.

2.7.2. EEG source analysis

The neuronal source of the scalp-based EEG signal was estimated using Brainstorm toolbox (Tadel et al., 2011). In particular, we estimated the current strength dynamics of the EEG cortical sources using the depth-weighted minimum norm estimation approach (Baillet et al., 2001) and a boundary element methods (BEM) conductive head model generated with OpenMEEG (Gramfort et al., 2010; Kybic et al., 2005) using the adaptive integration method. The solution space was constrained to the cerebral cortex, which was modeled as a three-dimensional grid of 15002 elementary current dipoles with unconstrained orientations based on the FreeSurfer brain template (FSAverage; see Fischl et al., 1999). Based on the results of our previous fMRI study (Visalli et al., 2019), we created ten regions of interest (ROIs) centered on the ten volumes of interest that we used in the correlational psychophysiological interaction analysis of fMRI data (see Supplementary Table S1 in Visalli et al., 2019) and covering a cortical surface of approximately 5 cm² (range: 4.99-5.07 cm²).

The ten ROIs were located in the following FPN and CON regions: bilateral inferior frontal gyri (IIFG and rIFG), bilateral posterior parietal cortices (IPPC, IPPC2, and rPPC), precuneus (PCun), posterior cingulate cortex (PCC), bilateral insulae (lAI and rAI), and dorsal anterior cingulate cortex (dACC). Activity in such regions was found to be modulated by updating and/or surprise with the following distinctions: (i) medial parietal regions within the FPN (i.e., PCun and PCC) were exclusively modulated by belief updating; (ii) lateral frontoparietal regions within the FPN (i.e., PPC and IFG) were instead modulated by both updating and surprise; (iii) CON regions (i.e., dACC, and AI) were modulated by surprise only.

The time course of the current strength of each ROI was computed by performing a principal component analysis and taking the first component explaining most of the signal variance.

2.7.3. EEG inferential statistics

For both scalp- and source-reconstructed EEG signals, first-level (subject-specific) analysis was performed using the *Unfold* toolbox (Ehinger and Dimigen, 2019) in MATLAB. This toolbox allows performing regression-based EEG analysis by integrating a mass-univariate approach with linear deconvolution. Deconvolution tries to disentangle overlapping electrophysiological responses from subsequent events (e.g., from stimulus onset and from button press). This aspect is very helpful in our paradigm to analyze the neural response associated with target onset. Indeed, targets were preceded by variable foreperiod intervals in which several processes likely occurred. Since targets appeared at different stages of these foreperiod processes depending on their status (i.e., predictable vs. surprising), the resolution of electrophysiological correlates of such processes differently affected target-related ERPs. Deconvolution helped us to face this problem by (at least partially) isolating specific ERPs associated with target onset. For a methodological description of the “Unfold” analysis steps, the reader is referred to the

original work (Ehinger and Dimigen, 2019). For the current analysis, we specified three events: cue onset, target onset and button press. Error trials were discarded from the analysis. Target onsets were modeled according to the following Wilkinson-notation formula: $y = 1 + D_{KL} + I_S$. As suggested by Ehinger and Dimigen (2019), cue onset and button press were modeled using a simple intercept term ($y = 1$) to remove ERP evoked by events of no interest from the estimation of the target evoked response. The design matrix was time expanded from -1 s to 1 s around each event using a set of spline functions (1 s is the target stimulus duration and, therefore, the maximum time allowed to respond). Before fitting the model, artifact intervals in the scalp-EEG signal were identified using a peak-to-peak voltage threshold of 75 μV and removed from the design matrix (i.e., set to 0). The same bad intervals identified in the scalp-EEG signal were also marked as artifactual in the source-reconstructed EEG signal and removed. The estimated regression coefficients (betas) of each predictor (D_{KL} and I_S), which can be visualized as ERP-like waveforms or scalp topographies (aka “regression-ERPs”, rERPs), were then entered in a second-level analysis for group inference.

Second-level analysis was performed using the ept- threshold-free cluster enhancement (TFCE) toolbox (Mensen and Khatami, 2013) in MATLAB. One-sample t-tests were performed on estimated betas of D_{KL} and I_S parameters³ in the data space channels \times epoch time points (0–1000 ms) and results were corrected for multiple comparisons using the TFCE approach. For scalp-EEG analysis, since performed at the whole scalp level, we used a conservative alpha-level = .001 (number of permutations = 20000), whereas for the source-reconstructed EEG analysis, since performed on *a priori* ROIs, we used the more common alpha-level = .05 (number of permutations = 5000).

3. Results

3.1. RT data results

The final model after backward selection is specified as follows:

$$\log(RT) \sim Trial + Preceding RT + D_{KL} * I_S + (Trial + Preceding RT + I_S | ID) \quad (9)$$

Visual inspection of the residuals showed that they were skewed. As suggested by Baayen and Milin (2010), trials with absolute standardized residuals higher than 2.5 SD were considered outliers and removed (1.6% of the trials). After outlier trials removal, the model was refitted achieving reasonable closeness to normality. Marginal R^2 of the final model, which represents the variance explained by the fixed effects, was 0.24; conditional R^2 , which is the variance explained by both fixed and random effects, was 0.58.

Table 1 shows the statistical results of the type III ANOVA. The analysis revealed a significant effect of I_S . Indeed, on average, RTs were longer for higher values of I_S . Importantly, this effect was further qualified by a significant interaction between D_{KL} and I_S . Fig. 2 shows that the increase of RTs with increasing surprise was slightly augmented with high D_{KL} in a multiplicative way. These results replicated our previous findings (see Table 1 and Fig. 2 in Visalli et al., 2019). In other words, at negligible levels of belief updating (a situation that is equivalent to common FP paradigms), participants were faster in responding to predictable compared to surprising target onsets (grey dots in Fig. 2; see also the left cloud of dots in supplementary Figure S3). Moreover, slowness for very surprising targets (around 2 SD above the mean surprise)

³ Whether baseline correction is an always appropriate EEG preprocessing step is still a matter of debate (Alday, 2019), especially in tasks with stimulus expectations or response preparation effects (Maess et al., 2016). For this reason, here we report results of TFCE analysis on no-baseline corrected betas. An analysis conducted on betas baseline corrected to the average betas calculated over the 50 ms preceding and following the target onset (Correa and Nobre, 2008; Kononowicz and van Rijn, 2014) yielded similar results.

Table 1

Analysis of variance with type-III sums of squares.

Fixed Effects	Sum Sq	Num. df	Den. df	F	p	β
Trial	0.15	1	24.9	4.1	.055	-0.04
Preceding RT	1.14	1	23.4	30.9	< .001	0.56
D_{KL}	2.84	1	18364.7	76.7	< .001	-0.10
I_S	9.27	1	24.5	250.2	< .001	0.26
$I_S:D_{KL}$	4.57	1	18364.3	123.4	< .001	0.12

Notes: F-statistics and associated p-values were calculated using Satterthwaite’s approximation of degrees of freedom. Additionally, standardized regression coefficients (β) are shown. D_{KL} and I_S represent the main effects of updating and surprise, respectively, while $D_{KL}:I_S$ represent their interaction.

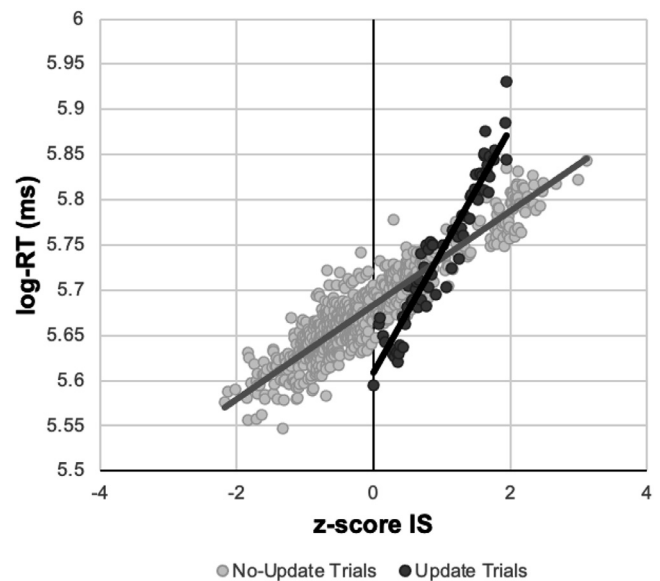


Fig. 2. Interaction plot for log-transformed RTs. The plot shows the effect of surprise (I_S) for update trials (black line and dots) and no-update (i.e., uniform and predictable; grey line and dots) trials. Fitted log-RT values were extracted from the final model using the *fitted.merMod* function in the *lme4* library. Then, separately for each trial type (i.e., update and no-update), they were averaged by I_S values across participants (black and grey dots represent average I_S values for update and no-update trials, respectively).

was greater at higher levels of belief updating, that is, in update trials (black dots in Fig. 2).

3.2. Scalp EEG data results

As predicted, the results of the TFCE analysis on updating and surprise (Fig. 3) showed that both processes were associated with distinct P3-like modulations but also with other earlier effects.

Specifically, we mainly observed three significant D_{KL} effects (Fig. 4). An early positive deflection was present over frontal electrodes (Fp1, Fpz, Fp2, AF3 AF4, AF8, F2, F4, F6, FC4) approximately between 230 and 280 ms. A second positive deflection emerged over posterior electrodes between 310 and 400 ms. The last significant effect consisted of a late and sustained positive modulation extending approximately between 480 and 820 ms over parietal scalp sites (CP4, CP6, P1, P3, P5, Pz, P2, P4, P6, P8, PO3, POz, PO4, PO8).

Concerning significant I_S effects (Fig. 5), an early positive deflection was observed roughly between 70 and 140 ms over posterior electrodes (CP6, P4, P6, PO3, POz, PO4, O1, Oz, O2). A later significant modulation was observed as a positivity that developed over centro-parietal scalp sites (C4, Cp1, Cp3, Cpz, Cp4, P1, P3, Pz, P2, P4, PO3, POz, PO4) approximately between 360 and 640 ms. Such P3-like modulation

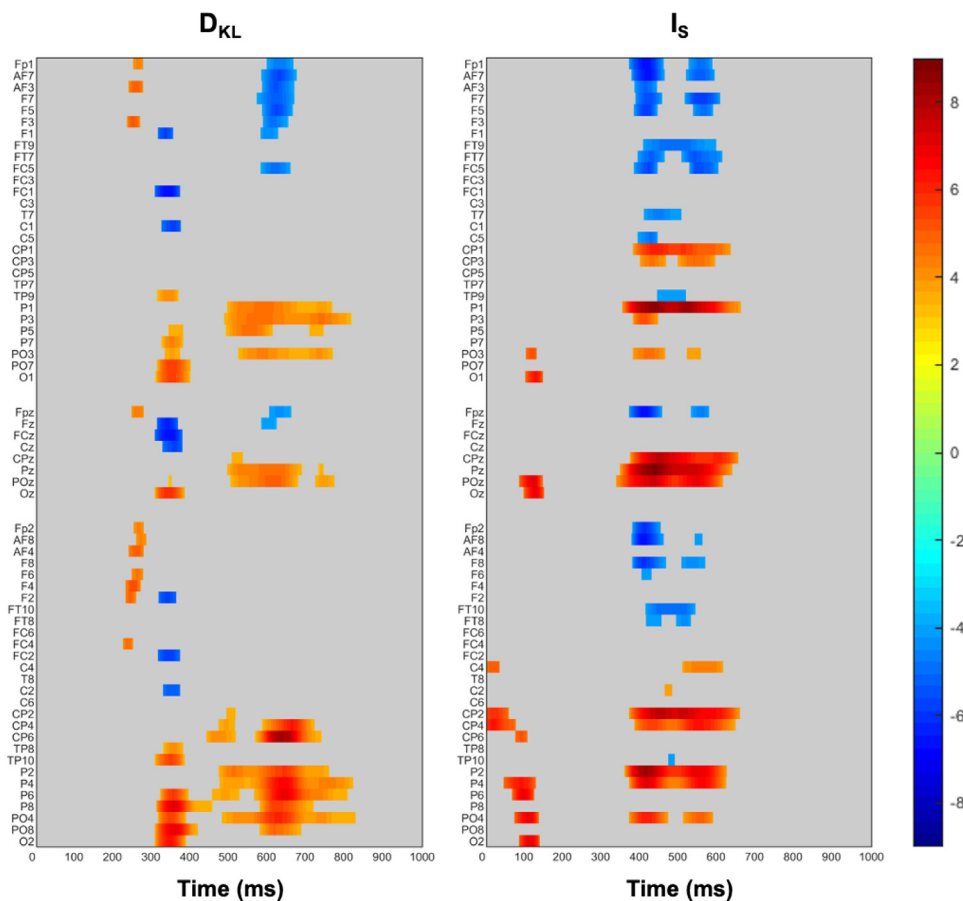


Fig. 3. Raster diagram showing significant effects elicited by updating (D_{KL}) and surprise (I_S) according to TFCE analysis. Rectangles in warm and cold colors indicate electrodes/time points significantly modulated by model-based regressors. The colorbar indicates t values. Gray rectangles indicate electrodes/time points that were not significantly modulated. Note that electrodes are organized along the y-axis somewhat topographically (Groppe et al., 2011). Electrodes on the left side of the scalp are grouped on the top of the diagram, midline electrodes are shown in the middle, and right electrodes are grouped at the bottom. Within those three groupings, the y-axis top-to-bottom corresponds to anterior-to-posterior electrodes

peaked earlier and was less sustained compared to the P3-like modulation associated with D_{KL} .

3.3. Source EEG data results

Results of the TFCE analysis are shown in Fig. 6. Concerning updating, analyses revealed no significant effect in any of the ROIs selected on the basis of our previous findings (Visalli et al., 2019). Concerning surprise, we found significant modulations in all cingulo-opercular network (CON) ROIs (dorsal anterior cingulate cortex, dACC: 226-356 ms; left anterior insula, lAI: 408-458 ms; right anterior insula, rAI: 346-474 ms) and in two parietal ROIs (IPPC2: 262-366 and 470-584 ms; posterior cingulate cortex, PCC: 380-436 ms).

4. Discussion

The present study investigated electrophysiological correlates of Bayesian inference about temporal beliefs, and it targeted separating decomposable event-related potential measures of surprise and belief updating. In order to achieve these goals, we dissociated surprise from updating through a FP paradigm that had been utilized in a previous fMRI study (Visalli et al., 2019). In that particular FP paradigm, the request for belief updating was purposefully decoupled from presenting surprising events through corresponding experimental manipulations. Corroborating previous research on the Bayesian brain and P3-like ERPs – and extending this line of research to the temporal domain – we found that surprise and requests for belief updating selectively modulated late positive ERP waveforms, as detailed below.

Before elaborating on the ERP findings, we shall briefly dwell on the present behavioral results, which replicated those from our previous fMRI study (Visalli et al., 2019). To begin with, we found that in-

creases in surprise were associated with increased RTs. With regard to the previous FP literature, this result is consistent with the well-observed modulation of RT by the hazard rate (Nobre and van Ede, 2018). Beyond FP studies, RTs are considered to reliably index, at least in part, the effort related to information encoding (Hyman, 1953; Teichner and Krebs, 1974; Zénon et al., 2019). This finding is hence in line with information-theory accounts of cognition, which posit that surprising events are effortful to encode (see Zénon et al., 2019, for an overview). Furthermore, we also replicated our previously reported finding of an interaction between surprise and updating, indicating that belief updating incurred higher RT costs at higher levels of surprise compared to lower levels of surprise. Overall, the replication of these RT costs provides novel insights into how RTs relate to information encoding effort. Prior to the present study, Zénon and colleagues (2019) suggested that these RT costs are equivalent to the amount of information that is required for belief updating. However, the present RT findings suggest that, at least in the temporal domain, it is the stimulus surprise that mainly affects RTs, and that updating-related contributions to RT costs can be appreciated only in interaction with surprise. We suggest that the distinction between surprise and updating in modulating RT should be more carefully sorted out in future research.

Turning to the confirmatory part of the present ERP analyses, we found a first significant late positive waveform being modulated in amplitude by surprise. Due to its timing (i.e., 360–640 ms following target onset) and its parietal scalp distribution, this ERP deflection may be best described as representing a modulation of the P3b component (Luck and Kappenman, 2011). The finding of surprise-related P3b amplitude modulation is consistent with analogous findings from some previous studies (Kolossa et al., 2013; 2015; Mars et al., 2008).

Previous work did not provide more precise definitions of those cognitive processes that should be captured by the information theoretic

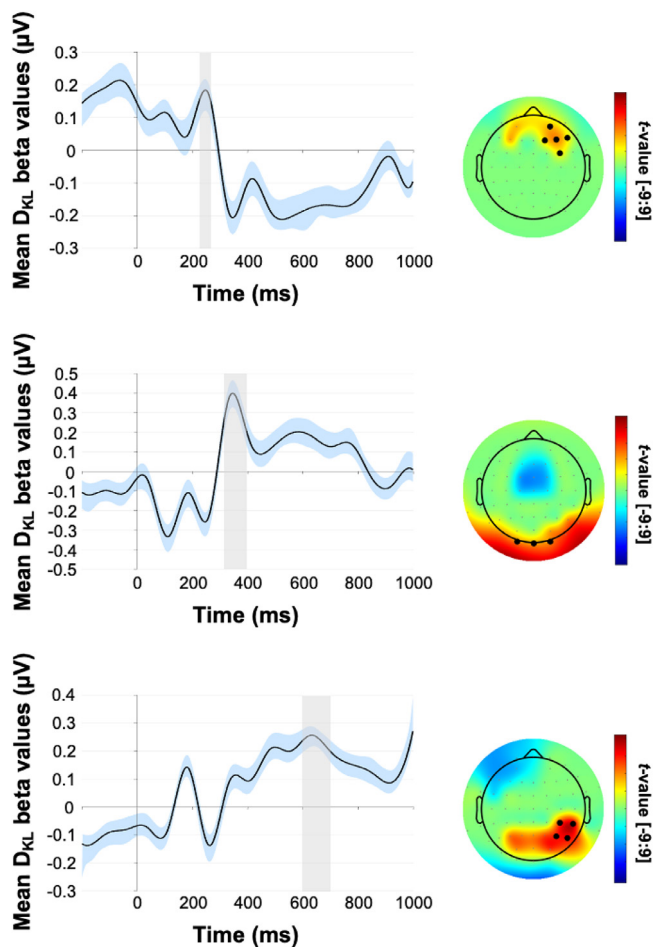


Fig. 4. Effects significantly modulated by Updating. The trace-plots depict the mean beta values (grand-average rERP) of the D_{KL} parameter (UNFOLD analysis) averaged over the electrodes indicated as black dots in the corresponding topoplots. Mean beta values are surrounded by standard error bounds. The topoplots show the t values averaged in the time windows grey shaded in the trace-plot (top: 230–270 ms; middle: 320–400 ms; bottom: 600–700 ms), which reflect the time windows of significance overlap across the electrodes indicated as black dots.

measure of surprise, I_S . Some features of the present FP paradigm may shed some light on this issue, given that we purposefully decoupled surprising and updating-inducing events in our experimental design. The color manipulation rendered it possible to present surprising events without any request for belief updating (see the Methods, Section 2.2). In addition to that, surprise was computed in the present study on the hazard rate of target onset rather than on its prior probability. To clarify this distinction, consider the example that we made in the Introduction. Imagine a task with three possible FP durations. When the participant's prior beliefs are [$FP_{short} = .8$, $FP_{medium} = .1$, $FP_{long} = .1$], if s/he sees the target occurring at FP_{long} , surprise equals 0 since the target probability has reached 1 according to the hazard function. However, despite surprise equaling 0, belief updating will still occur in response to the target, rendering FP_{long} more probable. Therefore, the information derived from the hazard function may be very useful during the current trial for facilitating behavioral responses to the target. In contrast, hazard rate at target onset is not utilized for updating temporal beliefs about FP durations on forthcoming trials. It is thus reasonable to consider the reported surprise-related P3b amplitude modulation as likely indexing neurocognitive processes that primarily subserve efficient immediate responses to targets.

A related issue is that previous studies typically used paradigms with two or more types of targets, which were associated with distinct mo-

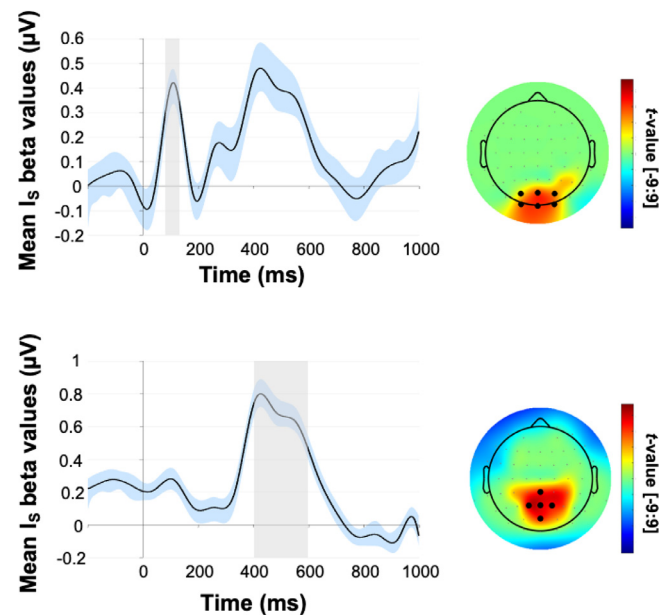


Fig. 5. Effects significantly modulated by Surprise. The trace-plots depict the mean beta values (grand-average rERP) of the I_S parameter (UNFOLD analysis) averaged over the electrodes indicated as black dots in the corresponding topoplots. Mean beta values are surrounded by standard error bounds. The topoplots show the t values averaged in the time windows grey shaded in the trace-plot (top: 100–150 ms; bottom: 400–600 ms), which reflect the time windows of significance overlap across the electrodes indicated as black dots.

tor responses (i.e., n -choice tasks, with $n \geq 2$). This aspect inevitably introduces another serious confound, as the surprise-related P3b amplitude modulation may either reflect stimulus-related sensory surprise or, not necessarily mutually exclusive, surprise related to the selection of the corresponding (more or less surprising) motor responses in n -choice RT tasks (Barceló et al., 2008; O'Connell et al., 2012). The fact that the present FP paradigm requires only one motor response (simple response task, i.e., $n = 1$) suggests that violations of sensory expectations per se (more precisely, temporal expectations here) are sufficient for eliciting the surprise-related P3b amplitude modulation. Nevertheless, since interactions between temporal preparation and task-demands have been shown to be captured by P3-like components (Barceló and Cooper, 2018a, 2018b), it would be interesting for future research to compare our P3 modulation with that elicited in tasks manipulating sensorimotor uncertainty (e.g., a two-forced choice paradigm).

Belief updating likewise modulated a positive P3b-like potential distributed over parietal electrodes. This amplitude modulation occurred at somewhat later latencies and as a more sustained modulation (480–820 ms) compared to the surprise-related P3b amplitude modulation. At first sight, this updating-related P3b amplitude modulation seems to lie in disagreement with results that were obtained in previous Bayesian studies. These studies reported associations between belief updating and waveform modulations of the relatively earlier and more anteriorly distributed scalp-recorded P3a component (Bennett et al., 2015; Kolossa et al., 2015). However, the discrepancy between present and previous results can be easily accounted for by differences in the way updating was computed, as detailed in the following paragraph.

As an example, Kolossa and colleagues (2015) used a modified version of the urn-ball paradigm (Phillips and Edwards, 1966), and they decomposed belief updating into two distinct processes referred to as “Bayesian surprise” and “postdictive surprise”. In their study, Bayesian surprise represents the updating of beliefs over latent variables (the urns from which balls are drawn), whereas postdictive surprise represents the change in beliefs over the observable stimuli (the balls). To better understand the difference between these two measures, we will make

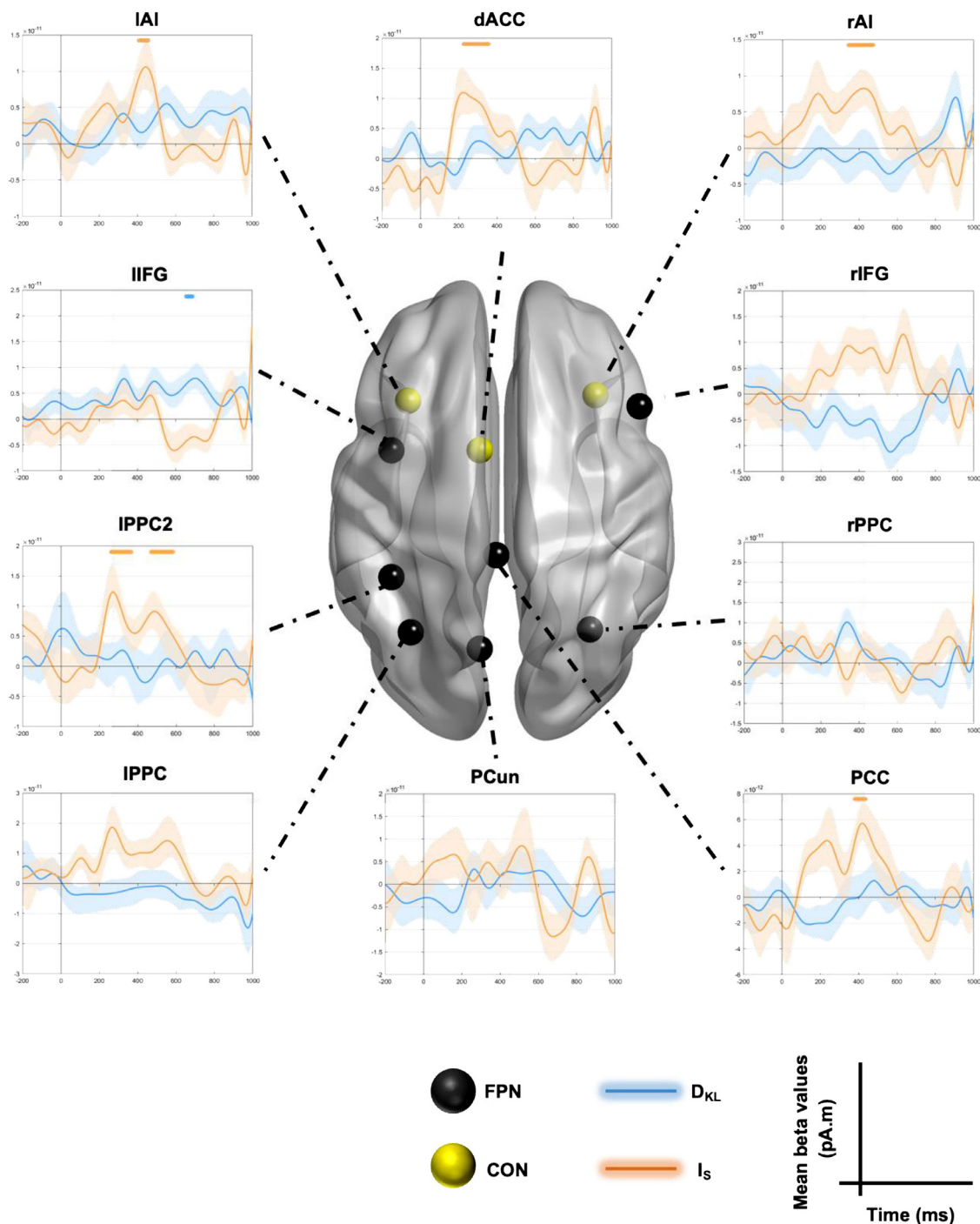


Fig. 6. Effects modulated by Updating and Surprise at the source space. The traceplots depict the mean beta values (grand-average rERP) of the D_{KL} (blue line) and I_S (orange line) parameters for each ROI. Mean beta values are surrounded by standard error bounds. The blue and orange straight lines above the traceplots indicate time points identified as significant by the TFCE analysis. MNI-space location of FPN (dark grey spheres) and CON (yellow spheres) ROIs is shown on an inflated surface rendering of the human brain (3D rendering was plotted with BrainNet Viewer, Xia et al., 2013). Abbreviations: IFG: inferior frontal gyrus; PPC: posterior parietal cortex; PCun: precuneus; PCC: posterior cingulate cortex; AI: anterior insula; dACC: dorsal anterior cingulate cortex; pre-SMA: pre-supplementary motor area; l and r indicate left and right.

a parallel with our FP paradigm. In the named study (Kolossa et al., 2015), latent variables would correspond to Gaussian distributions from which normal FPs were drawn in the present FP paradigm (i.e., see the Methods Section 2.3, specifically the description of the parameter space). Kolossa and colleagues' (2015) notion of Bayesian surprise therefore corresponds to changes in the probability over the parameter space in the present study. Note that this measure of Bayesian surprise was not considered in our study. In contrast, Kolossa and colleagues'

(2015) notion of postdictive surprise corresponds to the D_{KL} that we utilized to quantify belief updating in the present study. Kolossa and colleagues (2015) found that Bayesian surprise modulated P3a waveforms, whereas postdictive surprise better explained the modulation of positive slow wave (SW) amplitudes, which took place somewhat later compared to the proper P3b waveforms. As already mentioned above, I_S was related to the P3b in the named study (Kolossa et al., 2015). Considering all these issues, it seems reasonable to assume that our surprise- and

updating-related P3b amplitude modulations may correspond very well to (surprise-related) P3b and (postdictive surprise-related) SW waveform modulations, respectively, which had been reported by [Kolossa and colleagues' \(2015\)](#). On a side note, the mathematical equivalence between postdictive surprise in the named previous study and temporal belief updating in the present study exemplifies an important advantage of computational approaches because they allow tracing terms or concepts (e.g., surprise, Bayesian surprise, postdictive surprise, belief updating) back to their computational core.

As an interim conclusion, our data provide additional hints that classically observed P3b waveforms may result from spatio-temporal overlap of updating- and surprise-related ERP modulations. The present results further reconcile more traditional, yet loosely defined views of the P3b as an index of “context updating” ([Donchin and Coles, 1988](#); [Polich, 2007](#)) with previous Bayesian ERP studies that found P3b modulated by surprise. It is worth noting that none of our P3-like responses resembled the anterior P3a sub-component. More traditional views have associated the P3a with involuntary attentional responses to deviant/novel non-target stimuli ([Polich, 2007](#)), whereas some previous Bayesian ERP studies have reported a modulation of this component by belief updating ([Bennett et al., 2015](#); [Kolossa et al., 2015](#)). Although discussing about the absence of evidence is risky, focusing on differences between previous paradigms and ours (e.g., temporal modulation of attention, absence of non-target stimuli, use of non-discrete continuous stimuli) might inform future studies for a deeper (computational) comprehension of the P3a. Overall, our findings, in line with those from previous Bayesian ERP studies ([Bennett et al., 2015](#); [Kolossa et al., 2013](#); [Kolossa et al., 2015](#); [Mars et al., 2008](#)), reinforce the view according to which the P3 complex represents a valuable tool for decomposing neural correlates of Bayesian computations – such as surprise and belief updating – across a variety of experimental paradigms and cognitive domains. Overall, our findings, in line with those from previous Bayesian ERP studies ([Bennett et al., 2015](#); [Kolossa et al., 2013](#); [Kolossa et al., 2015](#); [Mars et al., 2008](#)), reinforce the view according to which the P3 complex represents a valuable tool for decomposing neural correlates of Bayesian computations – such as surprise and belief updating - across a variety of experimental paradigms and cognitive domains.

Turning to the more exploratory part of the present ERP analyses, we found associations between model-based regressors for updating and surprise and earlier ERP waveform modulations. Regarding surprise, we observed the modulation of a positive waveform that peaked at around 100 ms with occipital scalp distribution, resembling a P1 ERP component ([Luck, 2014](#)). Previous findings from temporal attention studies, which utilized symbolic cues predicting the likely timing of target onset, showed that early visual ERP components such as the P1 may be enhanced in response to validly cued temporal target onset under certain conditions (for a discussion on this issue, see [Correa et al., 2006](#)). This phenomenon has been explained in terms of pre-stimulus alpha synchronization ([Gruber et al., 2014](#)). At first glance, our finding of enhanced P1 amplitudes for surprising targets seems to be at odds with previous research. However, the majority of our surprise trials comprised “white” uniform targets. P1 amplitudes are known to be enhanced in response to more intense (i.e., brighter) stimuli ([Johannes et al., 1995](#)). As a result of such a limit in our task design, the observed P1 effect may be more parsimoniously explainable as being related to changes in physical stimulus features rather than being related to changes in their information properties.

Regarding belief updating, we found two other significant ERP waveform modulations: a positive deflection at fronto-central sites around 250 ms, and a somewhat later P2-like posterior positivity around 350 ms (most pronounced at occipital electrodes). Together, these two ERP waveform modulations suggest that belief updating may involve additional top-down perceptual processes. This possibility rests on the argument that the frontally discernible ERP waveform modulation resembled ERP components that were recently observed in perceptual decision-making studies in terms of their scalp topography and timing. Notably,

similar ERP waveform modulations have been considered as markers of prefrontal neural activities that reflect top-down perceptual processing in Go-NoGo paradigms ([Di Russo et al., 2017](#)). Even more specifically, the prefrontal P2 (pP2) was suggested to reflect the accumulation of sensory evidence to reach “Go” decisions ([Berchicci et al., 2016](#)). By analogy, the modulation of the fronto-central positive waveform in the present study might represent decisions for belief updating in response to color changes.

Concerning the updating-related posterior P2 amplitude modulation, previous studies have related P2 amplitudes to cognitive processes such as encoding of novel stimuli, target validity/incongruency effects and memory processes (see [Freunberger et al., 2007](#), for an overview). The present updating-related P2 findings may possibly index the recruitment of additional cognitive processes involved in belief updating. Of course, all these post-hoc interpretations should be treated with caution, and our computational approach was not meant to target these cognitive processes. These exploratory ERP findings suggest that surprise and updating affected also early stages of target processing differently, a topic that should be considered in more detail in future research.

A final point concerning our results pertains to the source reconstruction of the EEG signal, an analysis that included some cortical ROIs selected on the basis of our previous study ([Visalli et al., 2019](#)). Considering the parietal scalp topography of both surprise-related and updating-related P3-like effects, and since our fMRI study indeed showed some overlapped modulation by surprise and updating in the posterior parietal cortex (PPC), we expected to find similar P3 cortical generators in this area (a hypothesis supported by several fMRI-constrained EEG source analyses; see, [Bledowski et al., 2004](#); [Crottaz-Herbette and Menon, 2006](#); [Horowitz et al., 2002](#)). In contrast to this hypothesis, however, our analyses revealed only significant effects associated with surprise, thereby only partially corroborating our previous fMRI findings. Although these results do not add to our previous fMRI findings on updating, they are still important to get some valuable hints on the temporal course of neural activation associated with surprise. More in detail, like our previous fMRI study, we replicated that surprise elicited responses in the cingulo-opercular network (CON). This modulation was first observed in the dorsal anterior cingulate cortex (dACC) and then in bilateral anterior insulae (AI). As regards the fronto-parietal network (FPN), we found that surprise elicited a response in a portion of the left posterior parietal cortex (IPPC2) which was also modulated by surprise only in the fMRI study. However, we also observed a surprise-related significant modulation in the posterior cingulate cortex (PCC), an area found in our previous fMRI study to be associated only with belief updating. Overall, the present findings provide some evidence favoring our prior interpretation of the CON as a “surprise” network ([Visalli et al., 2019](#)). Our perspective represents an information-theoretic formalization of those theories that see the CON as a “salience network” involved in the transient detection of relevant stimuli in order to guide behavior ([Menon and Uddin, 2010](#); [Seeley et al., 2007](#); but see [Dosenbach, 2008](#)). According to their model, AI and ACC are involved in the rapid identification of deviant stimuli and in the generation of top-down control signals to neocortical regions, such as primary sensory areas, premotor areas, as well as lateral FPN areas generating the scalp-recorded P3b, including the PPC ([Menon and Uddin, 2010](#); [Sridharan, et al., 2008](#)), which are involved in the phasic adjustment of top-down control in response to identified salient stimuli ([Dosenbach et al., 2008](#)). Of course, these considerations should be taken with caution, since source reconstruction was not computed on high density EEG data or using the individual MRI of the participants.

The present study isolated ERP correlates of temporal belief updating by separating them from ERP correlates elicited by surprise. For such a purpose, we utilized the experimental design of our previous fMRI study ([Visalli et al., 2019](#)). However, a potential concern might be that the ecological validity of our study design was rather limited. For example, it might be difficult to generalize our results to contexts in which changes in the statistical properties of the environment are not explicitly sig-

naled (Nassar et al., 2010; Yu and Dayan, 2005). The ecological validity of our FP paradigm, and of Bayesian inference paradigms in general, remains a potentially important issue that should be investigated in future research.

To conclude, our main results can be summarized as follows: (i) In line with our previous fMRI study (Visalli et al., 2019), we were able to identify behavioral and neural mechanisms associated with updating of temporal belief and facing of temporally surprising events, thus confirming that Bayesian inference can represent a promising framework to model dynamic adaptations of temporal beliefs; (ii) Temporally surprising events elicited updating- and surprise-related P3 waveform modulations similar to those observed in non-timing tasks. These results, thus, suggest that the P3 is a valid index of inferential processes also in the temporal domain; (iii) Analyses in the source space revealed the involvement of areas belonging to control networks in the encoding of surprise, especially areas within the CON (suggesting a functional definition of the CON as a surprise network); (iv) Finally, the methodological approach adopted here (combining a task design thought to disentangle belief updating and surprise plus regression-based massive EEG analysis) allowed a partial reconciliation of long-held accounts of P3 amplitude fluctuations with more recent findings from Bayesian ERP studies. This progress was achieved by showing that belief updating and surprise elicited similar, yet clearly distinguishable P3b amplitude modulations in their time-course, which likely summed together in traditional P3b studies. Overall, the P3 has once again proved to be a valid indicator of information processing at the neural level and in the temporal domain. The present contribution for the functional interpretation of the P3 represents, therefore, a relevant step forward in the computational modeling of brain information processing.

5. Ethics statement

All procedures performed in this study involving human participants were approved by the Bioethical Committee of the Azienda Ospedaliera di Padova and were in accordance with the 1964 Helsinki declaration. Informed consent was obtained from all individual participants included in the study. Participants were reimbursed 20 Euros for their time.

6. Data and code availability statement

Task, model, and analysis codes and (raw) behavioral and EEG datasets are available at the Open Science Framework (<https://osf.io/ckqa5/>).

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Credit authorship contribution statement

Antonino Visalli: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Funding acquisition. **Mariagrazia Capizzi:** Conceptualization, Investigation, Writing - review & editing. **Ettore Ambrosini:** Methodology, Writing - review & editing. **Bruno Kopp:** Writing - review & editing, Supervision, Resources. **Antonino Vallesi:** Writing - review & editing, Supervision, Resources, Funding acquisition.

Acknowledgments

This work was partly funded by the European Research Council (ERC starting grant LEX-MEA, no. 313692, to A.Va). The first author acknowledges the support of the Boehringer Ingelheim Fonds in the form of a Travel Grant for his research period in the Lab of Prof. Kopp.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2021.117867](https://doi.org/10.1016/j.neuroimage.2021.117867).

References

- Alday, P.M., 2019. How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology* 56 (12), e13451. doi:10.1111/psyp.13451.
- Baayen, R.H., Milin, P., 2010. Analyzing reaction times. *Int. J. Psychol. Res.* 3 (2), 12–28.
- Baldi, P., Itti, L., 2010. Of bits and wows: a Bayesian theory of surprise with applications to attention. *Neural Netw.* 23 (5), 649–666. doi:10.1016/j.neunet.2009.12.007.
- Baillet, S., Riera, J.J., Marin, G., Mangin, J.F., Aubert, J., Garnero, L., 2001. Evaluation of inverse methods and head models for EEG source localization using a human skull phantom. *Phys. Med. Biol.* 46 (1), 77. doi:10.1088/0031-9155/46/1/306.
- Barceló, F., Cooper, P.S., 2018a. Quantifying contextual information for cognitive control. *Front. Psychol.* 9. doi:10.3389/fpsyg.2018.01693.
- Barceló, F., Cooper, P.S., 2018b. An information theory account of late frontoparietal ERP positivities in cognitive control. *Psychophysiology* 55 (3). doi:10.1111/psyp.12814.
- Barceló, F., Periáñez, J.A., Nyhus, E., 2008. An information theoretical approach to task-switching: evidence from cognitive brain potentials in humans. *Front. Hum. Neurosci.* 2, 13. doi:10.3389/neuro.09.013.2007.
- Barto, A., Mirolli, M., Baldassarre, G., 2013. Novelty or Surprise? *Front. Psychol.* 4. doi:10.3389/fpsyg.2013.00907.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi:10.18637/jss.v067.i01.
- Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10 (9), 1214–1221. doi:10.1038/nrn1954.
- Bennett, D., Murawski, C., Bode, S., 2015. Single-trial event-related potential correlates of belief updating. *eNeuro* 2 (5). doi:10.1523/ENEURO.0076-15.2015, ENEURO.0076-15.2015.
- Berchicci, M., Spinelli, D., Di Russo, F., 2016. New insights into old waves. Matching stimulus-and response-locked ERPs on the same time-window. *Biol. Psychol.* 117, 202–215. doi:10.1016/j.biopsycho.2016.04.007.
- Bledowski, C., Prvulovic, D., Hoehstetter, K., Scherg, M., Wibral, M., Goebel, R., Linden, D.E., 2004. Localizing P300 generators in visual target and distractor processing: a combined event-related potential and functional magnetic resonance imaging study. *J. Neurosci.* 24 (42), 9353–9360. doi:10.1523/JNEUROSCI.1897-04.2004.
- Brainard, D.H., 1997. The psychophysics toolbox. *Spatial Vis.* 10 (4), 433–436. doi:10.1163/156856897X00357.
- Buetti, D., Bahrami, B., Walsh, V., Rees, G., 2010. Encoding of temporal probabilities in the human brain. *J. Neurosci.* 30 (12), 4343–4352. doi:10.1523/JNEUROSCI.2254-09.2010.
- Chater, N., Oaksford, M., Hahn, U., Heit, E., 2010. Bayesian models of cognition. *Wiley Interdiscipl. Rev.: Cognit. Sci.* 1 (6), 811–823. doi:10.1002/wcs.79.
- Chaumon, M., Bishop, D.V., Busch, N.A., 2015. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J. Neurosci. Methods* 250, 47–63. doi:10.1016/j.jneumeth.2015.02.025.
- Cocchi, L., Zalesky, A., Formito, A., Mattingley, J.B., 2013. Dynamic cooperation and competition between brain systems during cognitive control. *Trends Cognit. Sci.* 17 (10), 493–501. doi:10.1016/j.tics.2013.08.006.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ.
- Correa, A., Lupiáñez, J., Madrid, E., Tudela, P., 2006. Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Res.* 1076 (1), 116–128. doi:10.1016/j.brainres.2005.11.074.
- Correa, A., Nobre, A.C., 2008. Neural modulation by regularity and passage of time. *J. Neurophysiol.* 100 (3), 1649–1655. doi:10.1152/jn.90656.2008.
- Coull, J.T., 2009. Neural substrates of mounting temporal expectation. *PLoS Biol.* 7 (8). doi:10.1371/journal.pbio.1000166.
- Courville, A.C., Daw, N.D., Touretzky, D.S., 2006. Bayesian theories of conditioning in a changing world. *Trends Cognit. Sci.* 10 (7), 294–300. doi:10.1016/j.tics.2006.05.004.
- Crottaz-Herbette, S., Menon, V., 2006. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. *J. Cognit. Neurosci.* 18 (5), 766–780. doi:10.1162/jocn.2006.18.5.766.
- De Lange, F.P., Heilbron, M., Kok, P., 2018. How do expectations shape perception? *Trends Cognit. Sci.* 22 (9), 764–779. doi:10.1016/j.tics.2018.06.002.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21. doi:10.1016/j.jneumeth.2003.10.009.
- Di Russo, F., Berchicci, M., Bozzacchi, C., Perri, R.L., Pitzalis, S., Spinelli, D., 2017. Beyond the “Bereitschaftspotential”: action preparation behind cognitive functions. *Neurosci. Biobehav. Rev.* 78, 57–81. doi:10.1016/j.neubiorev.2017.04.019.
- Donchin, E., 1981. Surprise!... surprise? *Psychophysiology* 18 (5), 493–513. doi:10.1111/j.1469-8986.1981.tb01815.x.
- Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11 (3), 357–374. doi:10.1017/S0140525X00058027.
- Dosenbach, N.U., Fair, D.A., Cohen, A.L., Schlaggar, B.L., Petersen, S.E., 2008. A dual-networks architecture of top-down control. *Trends Cognit. Sci.* 12 (3), 99–105. doi:10.1016/j.tics.2008.01.001.
- Doya, K., Ishii, S., Pouget, A., Rao, R.P., 2007. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. MIT press, Cambridge.

- Ehinger, B.V., Dimigen, O., 2019. Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ* 7, e7838. doi:10.7717/peerj.7838.
- Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39 (2), 175–191. doi:10.3758/bf03193146.
- Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8 (4), 272–284. doi:10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4.
- Freunberger, R., Klimesch, W., Doppelmayr, M., Höller, Y., 2007. Visual P2 component is related to theta phase-locking. *Neurosci. Lett.* 426 (3), 181–186. doi:10.1016/j.neulet.2007.08.062.
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G., 2017. Active inference: a process theory. *Neural Comput.* 29 (1), 1–49. doi:10.1162/NECO_a.00912.
- Friston, K., 2012. The history of the future of the Bayesian brain. *Neuroimage* 62 (248), 1230–1233. doi:10.1016/j.neuroimage.2011.10.004.
- Friston, K., 2005. A theory of cortical responses. *Philos. Trans. R. Soc. B: Biol. Sci.* 360 (1456), 815–836. doi:10.1098/rstb.2005.1622.
- Gramfort, A., Papadopoulos, T., Olivi, E., Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* 9 (1), 45. doi:10.1186/1475-925X-9-45.
- Groppe, D.M., Urbach, T.P., Kutas, M., 2011. Mass univariate analysis of event related brain potentials/fields I: a critical tutorial review. *Psychophysiology* 48 (12), 1711–1725. doi:10.1111/j.1469-8986.2011.01273.x.
- Gruber, W.R., Zauner, A., Lechinger, J., Schabus, M., Kutil, R., Klimesch, W., 2014. Alpha phase, temporal attention, and the generation of early event related potentials. *Neuroimage* 103, 119–129. doi:10.1016/j.neuroimage.2014.08.055.
- Herbst, S.K., Fiedler, L., Obleser, J., 2018. Tracking temporal hazard in the human electroencephalogram using a forward encoding model. *eNeuro* 5. doi:10.1523/ENEURO.0017-18.2018, ENEURO-0017.
- Horowitz, S.G., Skudlarski, P., Gore, J.C., 2002. Correlations and dissociations between BOLD signal and P300 amplitude in an auditory oddball task: a parametric approach to combining fMRI and ERP. *Magn. Reson. Imaging* 20 (4), 319–325. doi:10.1016/s0730-725x(02)00496-4.
- Hyman, R., 1953. Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* 45 (3), 188–196. doi:10.1037/h0056940.
- Hyvarinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. *Neural Netw.* 13 (4-5), 411–430. doi:10.1016/s0893-6080(00)00026-5.
- Itti, L., Baldi, P., 2009. Bayesian surprise attracts human attention. *Vis. Res.* 49, 1295–1306. doi:10.1016/j.visres.2008.09.007.
- Janssen, P., Shadlen, M.N., 2005. A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* 8, 234–241. doi:10.1038/nn1386.
- Johannes, S., Münte, T.F., Heinze, H.J., Mangun, G.R., 1995. Luminance and spatial attention effects on early visual processing. *Cognit. Brain Res.* 2 (3), 189–205. doi:10.1016/0926-6410(95)90008-x.
- Kersten, D., Mamassian, P., Yuille, A., 2004. Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271–304. doi:10.1146/annurev.psych.55.090902.142005.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., 2007. What's new in psychtoolbox-3. *Perception* 36 (14), 1–16.
- Knill, D.C., Pouget, A., 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27 (12), 712–719. doi:10.1016/j.tins.2004.10.007.
- Kobayashi, K., Hsu, M., 2017. Neural mechanisms of updating under reducible and irreducible uncertainty. *J. Neurosci.* 37 (29), 6972–6982. doi:10.1523/JNEUROSCI.0535-17.2017.
- Kolossa, A., Fingscheidt, T., Wessel, K., Kopp, B., 2013. A model-based approach to trial-by-trial p300 amplitude fluctuations. *Front. Hum. Neurosci.* 6, 359. doi: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.
- Kononowicz, T.W., van Rijn, H., 2014. Decoupling interval timing and climbing neural activity: a dissociation between CNV and N1P2 amplitudes. *J. Neurosci.* 34 (8), 2931–2939. doi:10.1523/JNEUROSCI.2523-13.2014.
- Kopp, B., Seer, C., Lange, F., Kluytmans, A., Kolossa, A., Fingscheidt, T., Hooijink, H., 2016. P300 amplitude variations, prior probabilities, and likelihoods: a Bayesian ERP study. *Cognit., Affect. Behav. Neurosci.* 16 (5), 911–928. doi:10.3758/s13415-016-0442-3.
- Ed. Kopp, B., 2008. The P300 component of the event-related brain potential and Bayes' theorem. In: Sun, M.K. (Ed.), *In Cognitive Sciences at the Leading Edge*. Nova Science, New York, NY, pp. 87–96 Ed..
- Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest: tests in linear mixed effects models. *J. Stat. Softw.* 82 (13), 1–26. doi:10.18637/jss.v082.i13.
- Kuznetsova, A., Christensen, R.H.B., Bavy, C., Brockhoff, P.B., 2015. Automated mixed ANOVA modeling of sensory and consumer data. *Food Qual. Pref.* 40, 31–38. doi:10.1016/j.foodqual.2014.08.004, A.
- Kybic, J., Clerc, M., Abboud, T., Faugeras, O., Keriven, R., Papadopoulos, T., 2005. A common formalism for the integral formulations of the forward EEG problem. *IEEE Trans. Med. Imaging* 24 (1), 12–28. doi:10.1109/tmi.2004.837363.
- Luck, S.J., 2014. *An Introduction to the Event-Related Potential Technique*. MIT press, Cambridge.
- Luck, S.J., Kappenman, E.S., 2011. *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press, New York.
- Maess, B., Schröger, E., Widmann, A., 2016. High-pass filters and baseline correction in M/EEG analysis. Commentary on: "How inappropriate high-pass filters can produce artefacts and incorrect conclusions in ERP studies of language and cognition. *J. Neurosci. Methods* 266, 164–165. doi:10.1016/j.jneumeth.2015.12.003.
- 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. *J. Neurosci.* 28 (47), 12539–12545. doi:10.1523/JNEUROSCI.2925-08.2008.
- Mathys, C., Daunizeau, J., Friston, K.J., Stephan, K.E., 2011. A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5, 39. doi:10.3389/fnhum.2011.00039.
- McGuire, J.T., Nassar, M.R., Gold, J.I., Kable, J.W., 2014. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 84 (4), 870–881. doi:10.1016/j.neuron.2014.10.013.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214 (5-6), 655–667. doi:10.1007/s00429-010-0262-0.
- Mensen, A., Khatami, R., 2013. Advanced EEG analysis using threshold-free cluster-enhancement and non-parametric statistics. *Neuroimage* 67, 111–118. doi:10.1016/j.neuroimage.2012.10.027.
- Nassar, M.R., Wilson, R.C., Heasley, B., Gold, J.I., 2010. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J. Neurosci.* 30 (37), 12366–12378. doi:10.1523/JNEUROSCI.0822-10.2010.
- Niemi, P., Näätänen, R., 1981. Foreperiod and simple reaction time. *Psychol. Bull.* 89 (1), 133.
- Nobre, A.C., Van Ede, F., 2018. Anticipated moments: temporal structure in attention. *Nat. Rev. Neurosci.* 19 (1), 34. doi:10.1038/nrn.2017.141.
- O'Connell, R.G., Dockree, P.M., Kelly, S.P., 2012. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* 15 (12), 1729. doi:10.1038/nrn.3248.
- O'Reilly, J.X., Schuffelgen, 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. *Proc. Natl. Acad. Sci., USA* 110 (38), E3660–E3669. doi:10.1073/pnas.1305373110.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. doi:10.1016/0028-3932(71)90067-4.
- Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72 (2), 184–187. doi:10.1016/0013-4694(89)90180-6.
- Pezzulo, G., 2012. An active inference view of cognitive control. *Front. Psychol.* 3, 478. doi:10.3389/fpsyg.2012.00478.
- Phillips, L.D., Edwards, W., 1966. Conservatism in a simple probability inference task. *J. Exp. Psychol.* 72 (3), 346. doi:10.1037/h0023653.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi:10.1016/j.clinph.2007.04.019.
- 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.
- ... & Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27 (9), 2349–2356. doi:10.1523/JNEUROSCI.5587-06.2007.
- Smith, N.J., Kutas, M., 2015. Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology* 52 (2), 157–168. doi:10.1111/psyp.12317.
- 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.
- Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci.* 105 (34), 12569–12574. doi:10.1073/pnas.080005105.
- 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.
- Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011. doi:10.1155/2011/879716.
- Tanner, D., Morgan-Short, K., Luck, S.J., 2015. How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology* 52 (8), 997–1009. doi:10.1111/psyp.12437.
- Teichner, W.H., Krebs, M.J., 1974. Laws of visual choice reaction time. *Psychol. Rev.* 81 (1), 75. doi:10.1037/h0035867.
- Vallesi, A., 2010. Neuro-anatomical substrates of foreperiod effects. In: Coull, J.F., Nobre, A.C. (Eds.), *In Attention and Time*. Oxford University Press, New York, pp. 303–316.
- Verleger, R., 1988. Event-related potentials and cognition: a critique of the context updating hypothesis and an alternative interpretation of P3. *Behav. Brain Sci.* 11 (3), 343–356. doi:10.1017/S0140525X00058015.
- Visalli, A., Capizzi, M., Ambrosini, E., Mazzone, I., Vallesi, A., 2019. Bayesian modeling of temporal expectations in the human brain. *Neuroimage* 202, 116097. doi:10.1016/j.neuroimage.2019.116097.
- Wilkinson, G., Rogers, C., 1973. Symbolic description of factorial models for analysis of variance. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* 22 (3), 392–399. doi:10.2307/2346786.
- Winkler, I., Debener, S., Müller, K.R., Tangermann, M., 2015. On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP, pp. 4101–4105 IEEE.
- Xia, M., Wang, J., He, Y., 2013. BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8 (7). doi:10.1371/journal.pone.0068910.
- Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. *Neuron* 46 (4), 681–692. doi:10.1016/j.neuron.2005.04.026.
- Zénon, A., Solopchuk, O., Pezzulo, G., 2019. An information-theoretic perspective on the costs of cognition. *Neuropsychologia* 123, 5–18. doi:10.1016/j.neuropsychologia.2018.09.013.