P3-like signatures of temporal predictions: a computational EEG study

Visalli, A.[1,2,3*], Capizzi, M.[4], Ambrosini, E.[1,2,5], Kopp, B.[6], Vallesi A.[1,2*]

[1]Department of Neuroscience, University of Padova, Padova, Italy

[2]Padova Neuroscience Center, University of Padova, Padova, Italy

[3]IRCCS San Camillo Hospital, Venice, Italy

[4]Mind, Brain and Behavior Research Center (CIMCYC), University of Granada; Department of Experimental Psychology, University of Granada, Spain

[5]Department of General Psychology, University of Padova, Padova, Italy

[6]Department of Neurology, Hannover Medical School, 30625 Hannover, Germany

***Corresponding authors:**

Antonino Visalli,

IRCCS San Camillo Hospital

Venice, Italy, 30126

email: antonino.visalli@unipd.it or antonino.visalli.av@gmail.com

or

Antonino Vallesi,

Department of Neuroscience, University of Padova

Padova, Italy, 35121

email: antonino.vallesi@unipd.it

ABSTRACT

Many cognitive processes, ranging from perception to action, depend on the ability to predict the timing of forthcoming events. Yet, how the brain uses predictive models in the temporal domain is still an unsolved question. In previous work, we began to explore the neural correlates of temporal predictions by using a computational approach in which an ideal Bayesian observer learned the temporal probabilities of target onsets in a simple reaction time task. Because the task was specifically designed to disambiguate updating of predictive models and surprise, changes in temporal probabilities were explicitly cued. However, in the real world, we are usually incidentally exposed to changes in the statistics of the environment. Here, we thus aimed to further investigate the electroencephalographic (EEG) correlates of Bayesian belief updating and surprise associated with incidental learning of temporal probabilities. In line with our previous EEG study, results showed distinct P3-like modulations for updating and surprise. While surprise was indexed by an early fronto-central P3-like modulation, updating was associated with a later and more posterior P3 modulation. Moreover, updating was associated with a P2-like potential at centro-parietal electrodes, likely capturing integration processes between prior beliefs and likelihood of the observed event. These findings support previous evidence of trial-by-trial variability of P3 amplitudes as an index of dissociable inferential processes. Coupled with our previous findings, the present study strongly bolsters the view of the P3 as a key brain signature of temporal Bayesian inference. Data and scripts are shared on OSF: osf.io/sdy8j/

Keywords: temporal predictions; Bayesian brain; P3; Kullback-Leibler divergence, Shannon information; hazard rate

1. Introduction

Predicting the timing of forthcoming events is crucial to optimize brain processes that range from perception to action. Temporal expectations can be computationally formalized in terms of the hazard function, which is defined as the conditional probability that an event will occur given that it has not yet occurred (Niemi & Näätänen, 1981; Nobre & van Ede, 2018). It follows that temporal expectations are encoded on the basis of both prior beliefs about event timing and the information inherent in elapsed time. Previous studies employing multiple probability distributions of foreperiods (FP; i.e., the time intervals between a ready signal and the presentation of the target stimulus) showed that the brain tracks the probability of target onset as time elapses (Bueti et al., 2010; Herbst et al., 2018; Meindertsma et al., 2018). Further extending this line of research, we started to investigate how FP distributions are inferred by the brain, an issue that has been less explored.

To fill that gap, we combined a Bayesian computational approach with functional magnetic resonance imaging (fMRI) (Visalli et al., 2019) and electroencephalography (EEG) (Visalli et al., 2021) to elucidate the neural correlates of temporal belief updating. Crucially, since measures of belief updating (e.g., Kullback-Leibler divergence, D_{KL}) are often confounded with measures of surprise (e.g., Shannon information, I_S), our studies adapted a task manipulation from O'Reilly and colleagues (2013) that allowed us to differentiate updating and surprise correlates (see Baldi & Itti, 2010; Itti & Baldi, 2009 for a deeper discussion on the difference between updating and surprise measures). Briefly, our previous fMRI study identified two networks in part distinctly associated with updating and surprise: the fronto-parietal network, which was mainly involved in updating prior beliefs about target timing, and the cingulo-opercular network, which was related to the surprise elicited by target onset (Visalli et al., 2019). Moreover, our previous EEG study successfully isolated two dissociable P3-like modulations specifically indexing belief updating and surprise about the timing of events (Visalli et al., 2021).

To disentangle updating and surprise, we capitalized on a color manipulation (O'Reilly et al., 2013), in which changes in the current FP distribution were explicitly signaled. More in detail, in both of our previous studies, participants had to respond to target onsets that were separated from an uninformative ready signal by variable FPs. In the majority of trials (*predictable* trials), FP durations were sampled from a Gaussian distribution with constant mean and standard deviation (SD) during a block of trials. Means and SDs changed abruptly between blocks of trials. The key manipulation was that blocks were not temporally separated, but participants were explicitly instructed that a change in target color signaled the beginning of a new block (*updating* trials). On a few trials (*uniform* trials), interspersed with the other

trials, FP durations were sampled from a uniform distribution. Importantly, uniform targets were always white, signaling that the current trial was not igniting a new block. This color manipulation allowed us to create two types of highly surprising events, updating and uniform trials, with only the former ones requiring updating.

While in our previous studies the need of updating was explicitly signaled to experimentally tease apart updating and surprise, belief updating in a Bayesian sense is usually incidentally accomplished in everyday life. In this regard, it has been shown that cued and incidental inferential processes are differently encoded by the brain (e.g., Pearson et al., 2011; Yu & Dayan, 2005). This point raises the possibility that our results might not truly reflect Bayesian inference but, rather, a sort of *all-or-none* shift in beliefs (Nassar, Wilson, Heasly, & Gold, 2010) caused by the color manipulation. Addressing this issue is, thus, critical to substantiate (or, alternatively, circumscribe) our previous conclusions.

The present study aimed to investigate electrophysiological correlates underlying incidental Bayesian updating of temporal expectations. To this aim, we employed the same task as in our previous EEG study (Visalli et al., 2021), but without the color manipulation. Moreover, as detailed below, participants were not informed about the probabilistic structure of the task.

Our previous EEG study showed that surprise and updating acted on both early and late processing stages. The modulation of later components was in line with the literature (Kolossa et al., 2013, 2015; Mars et al., 2008) confirming that the P3 is sensitive to inferential processes (Kopp, 2008). In contrast, we could not draw firm conclusions on the specific role of earlier components. This was due to the fact that, since processing of color cues necessarily affected early components, it was not clear whether such event-related potentials (ERPs) truly encoded probabilistic information or they were just involved in color processing. Testing this issue is crucial to pinpoint the difference between cued and incidental inferential processes and to better interpret the results of our previous study.

To sum up, here we tried to gain further insight into the inferential mechanisms underlying temporal expectations when they are incidentally accomplished, thus, extending our previous work on cued temporal inferential processes.

2. Methods

2.1 Participants

This study included a sample of 28 participants [17 females; mean age: 24.5 years (SD = 4 years), range: 19-33 years; mean years of education: 16.1 (SD = 2.4)]. All of them were right-handed [Edinburgh Handedness Inventory (Oldfield, 1971) average score: 82.7 (SD = 19), range: 40-100], had normal color vision and normal or corrected-to-normal visual acuity, and reported no history of neurological or psychiatric disorders. 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 (Figure 1) was the same variable FP task as in Visalli and colleagues (2021), except for the target color that was always black. Participants were uninformed about the temporal structure or the probabilistic nature of the task. They were only instructed to respond as fast as possible to the onset of the target (a black circle displayed centrally against a gray background for 1000 ms). The target was separated from the onset of a neutral warning signal (a centrally displayed black fixation cross) by a variable FP duration. The inter-trial interval was randomly jittered between 1250 and 1750 ms (rectangular distribution). 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 sequence of events was the same as in our previous study (Visalli et al., 2021). The task was divided in 73 blocks (the first block was used in computational modeling but excluded from the inferential analyses), each one including a number of trials ranging between 8 and 13 (mean: 10.5). Blocks were not temporally separated and transition between them was not signaled. In most trials (80% of 790 trials) the FP duration was drawn from a Gaussian distribution whose mean and SD was constant during a block of trials. Gaussian parameters, which changed between blocks, were derived from an orthogonal combination of 9 means (500, 700, 900, 1100, 1300, 1500, 1700, 1900, and 2100 ms) and 8 SDs (25, 50, 75, 100, 125, 150, 175 and 200 ms). 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. In the remaining 20% of trials, FP durations were drawn from a uniform distribution ranging from 250 to 2500 ms (half of them had a FP below the mean of the Gaussian distribution of the current block and the other half above). The task was divided in four runs with self-paced breaks in between. Each new run started with either 8 or 9 trials with the same FP distribution as that used in the last block of the previous run (these trials were discarded from statistical analyses).

The task was implemented in MATLAB (The MathWorks, Inc., Natick, Massachusetts, United States) using the PSYCHOPHYSICS TOOLBOX 3 (Brainard, 1997; Kleiner et al., 2007).

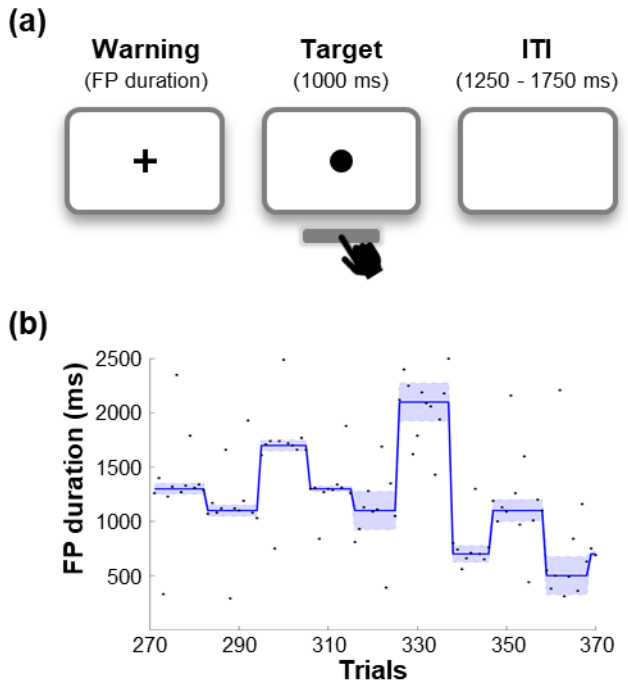

Figure 1 | (A) Schematic of a typical trial. Participants had to respond to target onsets that were separated from a neutral warning signal by variable foreperiod (FP) durations. Trials were separated by an inter-trial interval (ITI) drawn from a uniform distribution ranging from 1250 to 1750 ms. **(B) FP durations in one hundred exemplary trials.** On 80% of the trials, target onsets occurred after a FP (black dots) drawn from a Gaussian distribution, whose mean (solid blue line) and standard deviation (blue bands) were kept constant within a block. Transitions between blocks were not signaled. On 20% of the trials, targets occurred after a FP extracted from a uniform distribution.

2.3 Ideal Bayesian learner

Measures of updating and surprise were computed on the basis of a computational model that estimates, on a trial-by-trial basis, the posterior probability of putative pairs of μ (i.e., mean) and σ (i.e., SD) parameters of the Gaussian distribution from which FPs are drawn. The model is a modified version of the one used in our previous studies (Visalli et al., 2019, 2021; see O'Reilly et al., 2013), which takes into account not only differences in the paradigm and the participants' knowledge about the task (i.e., different task instructions), but also the long-term trial-wise accumulation of evidence and memory constraints. Namely, in the previous paradigm, after the change in the target color, participants could discard the information carried out by the previous trials since this was no longer useful to infer the new, current FP

distribution. Moreover, participants were instructed not to update beliefs after an FP drawn from the uniform distribution. Conversely, in the present study, participants were exposed to a long series of FPs that was experienced as a unique sequence throughout the task and without explicit differences being signaled between updating, uniform and predictable trials.

The absence of a cued transition between blocks implied that, at a given trial n, all the observations from 1 to n were available for estimating the posterior, thus, raising the question of to what extent previous observations are used for estimation. For example, it could be possible that all past events concur in the estimation, with an equal weight for distant and recent observations. However, several studies have shown that the brain integrates past experience according to a temporal gradient over past observations (Harrison et al., 2011). For instance, in the classical oddball paradigm (N. K. Squires et al., 1975), the recent trial history appears to influence the P3b amplitude to a greater extent than the remote history (K. C. Squires et al., 1976). More recently, Kolossa and colleagues (2013) implemented a model of trial-by-trial P3 fluctuations in a simple two-choice response time (RT) task, which takes into account both short-term and long-term memory decay processes. Since this model provided a superior account of P3b amplitude with respect to other previous P3b models, we supplied our model with those two short-term and long-term memory forgetting factors.

Let us call

$$P_{FP}(n) = P(FP \sim \aleph(\mu, \sigma) | FP_{1:n}) \quad (1)$$

the estimated FP posterior probability over the parameter space $\aleph(\mu, \sigma)$ on trial *n* (the parameter space had a size of 150×50, that is, the combination of all the means from 20 to 3000 ms and SDs from 20 to 1000 ms in steps of 20 ms). Based on Kolossa and colleagues (2013), the posterior can be expressed as:

$$P_{FP}(n) = \alpha_L \cdot P_{L,FP}(n) + \alpha_S \cdot P_{S,FP}(n) \quad (2)$$

that is, the weighted sum of posteriors estimated according to long-term ($P_{L, FP}$) and short-term ($P_{S, FP}$) decaying memories about prior observations. The weighting parameters α_L and α_S represent the relative contributions of $P_{L,FP}$ and $P_{S,FP}$ in forming the posterior, and they hold $\alpha_S = 1 - \alpha_L$ and $0 \leq \alpha_L \leq 1$. Both posteriors were computed as:

$$P_{i,FP}(n) = P_{i,FP}(n-1)^{\gamma_i} \cdot L_{FP}(n), i \in [L, S] \quad (3)$$

where $L_{FP}(n)$ is the likelihood $P(FP_n | FP \sim \aleph(\mu, \sigma))$ at trial *n* (see Visalli et al., 2021, for further details), and $P_{i,FP}(n-1)^{\gamma_i}$ is the power prior (Ibrahim et al., 2015) $P(FP \sim \aleph(\mu, \sigma) | FP_{1:n-1})$ (see Visalli et al., 2021, for further details) raised to the power γ_i, which

represents the exponential forgetting factor of the long-term (γ_L) or short-term (γ_S) decaying memories. From Kolossa and colleagues (2013), the short-term and long-term decay factors were respectively defined as:

$$\gamma_S = e^{-1/\beta_S} \quad \& \quad \gamma_L = e^{-1/\beta_{L,n}} \tag{4}$$

Both forgetting factors are exponential, and only the long-term memory depends on the trial number *n* as follows:

$$\beta_{L,n} = e^{-(n \cdot 1/\tau_1 + 1/\tau_2)} \tag{5}$$

This aspect determines that forgetting becomes much sharper when the number of trials increases (see Kolossa et al., 2013).

Posteriors were, then, translated from the parameter space over time (see Visalli et al., 2021), as follows:

$$P_{i,FP}(n) = \Sigma_{j,k} P(FP_n | \, FP_n \sim \aleph(\mu_j, \sigma_k)) \times P(FP_n \sim \aleph(\mu_j, \sigma_k) | FP_{1:n}) \tag{6}$$

The values for the model parameters, α_L, τ_1, τ_2, and β_S, were identified by finding the parameter combination that better explained participants' RTs (see Section 2.5) in terms of Akaike information criterion (AIC). This parameter optimization was performed using the MATLAB *fmincon* function (Table 1 provides an overview of the searched parameter space; see Table 1 in Kolossa et al., 2013).

Since Meyniel and colleagues (2016) observed that a single forgetting parameter had a better fitting as compared to the forgetting parameters by Kolossa and colleagues (2013), we implemented a second model, in which the posterior probability over the parameter space on trial *n* was computed as follows:

$$P_{FP}(n) = P_{i,FP}(n-1)^{e^{-(n-1)/\omega}} \cdot L_{FP}(n) \tag{7}$$

In (7), we used a single forgetting parameter γ = e^(-(n-1)/ω) (see, Meyniel et al., 2016), with ω optimized as described above (searched parameter space went from 1 to the number of trials in the task, i.e., 790). The first model with two forgetting factors (Equations 3-5) and the second model variant with a single forgetting factor (Equation 7) were then compared in terms of AIC in explaining RTs as in Equation 12 (see Section 2.5). Since the first model was considerably better (AIC of model 1 = -344.88; AIC of model 2 = -232.44; please note that lower AIC scores indicate a better fit), all the presented analyses refer to first model.

The above-described Bayesian learner was chosen to be as similar as possible to the one employed in our previous studies (Visalli et al., 2019, 2021), in order to facilitate between-study comparisons. Nonetheless, it has been often shown that in volatile environments (i.e., with abrupt changes at random points in time), the learning rate is influenced by the estimated volatility of the environment (Behrens et al., 2007; Liakoni et al., 2021; Nassar et al., 2010). For this reason, in the present study, we also applied the Hierarchical Gaussian filter (HGF; Mathys et al., 2011), a generic hierarchical Bayesian model frequently used to model beliefs in learning tasks under volatility. Details on the HGF implementation are presented as Supplementary Information along with HGF-based EEG results.

2.4 Model-based measures of updating and surprise

Trial-wise measures of updating and surprise were calculated with reference to the described Bayesian learner (see supplementary Figure S3). As mentioned in the Introduction, updating was quantified in terms of the Kullback-Leibler divergence (D_{KL}) between prior and posterior on each trial (Itti & Baldi, 2009):

$$D_{KL}(n) = \sum_{FP=10}^{3000} P_{FP}(n-1) \cdot log \frac{P_{FP}(n-1)}{P_{FP}(n)} \quad (8)$$

Concerning surprise, it was quantified in terms of the Shannon information (I_S; Itti & Baldi, 2009), that is, the negative log value of the probability of the observed event. Specifically, since during a trial of a variable FP task the prior probability of FP duration changes as a function of the elapse of time (i.e., hazard function), surprise on each trial was computed as:

$$I_S(n) = -log\ h(FP_n) \quad (9)$$

where, $h(FP_n)$ is the hazard rate of the observed FP duration at trial n, which is specified as:

$$h(FP_n) = \frac{P(FP_n|FP_{1:n-1})}{1-P(FP<FP_n|FP_{1:n-1})} \quad (10)$$

that is, the prior probability of target occurrence after the FP duration at trial *n*, divided by the probability that the target had not yet occurred after the FP at trial *n*.

2.5 RT data analysis

To find the best parameter combination for the Bayesian learner, log-transformed RTs (log-RTs) were analyzed by means of Linear Mixed Models (LMM) using the *fitlme* function in

MATLAB. Data from the initial familiarization block and the beginning of each new run (see Section 2.2), along with data from error trials (i.e., anticipated and/or missing response to the target; mean percentage: 2.3%, SD = 2.2%, range: 0.3-8.9%) and post-error trials were excluded from analyses. To find the best combination of model parameters for the Bayesian Learner (see above), log-RTs were analyzed using a minimal model (in order to speed-up computations) specified as the following Wilkinson-notation formula:

$$log(RT) \sim D_{KL} * I_S + (1|ID) \quad (11)$$

In detail, the fixed part included D_{KL} and I_S regressors (and their interaction) obtained from the Bayesian learner with a specific combination of model parameters. These regressors were standardized using Z-score in order to facilitate model convergence. The random part included by-participant (ID) random intercepts. Using the iterative procedure described above, we looked for the parameter combination associated with the pair of D_{KL} and I_S regressors, which better explained log-RTs (i.e., which led to the LMM with the lowest AIC).

The D_{KL} and I_S associated with the best parameter combination were used to reanalyze RT data as in our two previous studies (Visalli et al., 2019, 2021), as well as for the EEG data analysis (see below). First, a full LMM was specified using the *lme4* library (Bates et al., 2015) in R (http://www.R-project.org/) as follows:

$$log(RT) \sim Trial + PreRT + D_{KL} * I_S + (Trial + PreRT + D_{KL} + I_S|ID) \quad (12)$$

In detail, in addition to D_{KL} and I_S and their interaction, we entered the rank-order of each trial (Trial) and the log-RT at the preceding trial (PreRT) as fixed-effect terms to control for the temporal dependencies between successive trials that generally need to be addressed when modeling RTs (Baayen & Milin, 2010). These continuous predictors were standardized using Z-score in order to facilitate model convergence. The random structure included correlated by-subject random intercepts and slopes for Trial, PreRT, I_S and D_{KL}. Model selection procedure from the full model to a minimal adequate model was conducted through the *step* function 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 and pre-processing

The EEG was recorded using BrainAmp amplifiers (Brain Products, Munich, Germany) from 64 Ag/AgCl electrodes 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 to be lower than 10 kΩ 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.

Pre-processing of the EEG signal was the same as in Visalli and colleagues (2021) and it was performed using custom MATLAB scripts including functions from the EEGLAB environment (version 14.1.2b; Delorme & Makeig, 2004). To perform independent component (IC) decomposition, a low-pass (cut-off frequency = 40 Hz, transition bandwidth = 20 Hz) and a high-pass (cut-off frequency = 2 Hz, transition bandwidth = 4 Hz) non-causal zero-phase Kaiser-windowed sinc FIR filters (beta = 5.6526) were applied to continuous data. The high cut-off for the high-pass filter was applied to remove low-frequency drifts in order to improve the ICA solution (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 4 channels per subject (mean = 0.75 channels, SD = 1.27) were removed. Then, the *FastICA* algorithm (Hyvärinen & Oja, 2000) was used for IC decomposition, and the obtained ICA solution was applied to EEG data filtered with a high-pass cut-off of 0.1 Hz (transition bandwidth = 0.2 Hz), since the previously applied 2-Hz cut-off may attenuate ERP effects and introduce distortions (Tanner et al., 2015). Subsequently, the EEGLAB plugin *SASICA* (Chaumon et al., 2015) guided the identification and exclusion of artifactual ICA components (e.g., blinks, eye movements, muscle activity, noisy channels). Finally, removed channels were interpolated using spherical splines 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 EEG inferential analyses

As in Visalli and colleagues (2021), first-level (subject-specific) analysis was performed using the *Unfold* toolbox (Ehinger & Dimigen, 2019). This toolbox allowed us to perform regression-based analyses of single-trial target-evoked EEG responses by (at least partially) disentangling overlapping responses from subsequent events (i.e., warning signal onset, target onset and response) through deconvolution. Deconvolution is very helpful in our paradigm to analyze the neural response associated with target onset. Indeed, targets appeared at different stages of FP processes due to the high variability of FP intervals. In other terms, ERP response to the warning signal differently affected target-related ERPs on the basis of FP durations and their predictability. Deconvolution helped us to face this problem by

(at least partially) isolating specific ERPs associated with target onset. Moreover, baseline correction as an 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). Deconvolution methods are a way to face baseline correction in paradigms with overlapping responses to different events (Lütkenhöner, 2010).

Events from error and post-error trials were discarded. Events of no interest (i.e., warning signal onsets and responses) were modeled using a simple intercept term. Target onsets were modeled according to the following Wilkinson-notation formula: $y \sim 1 + D_{KL} + I_S$. 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, residuals 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 estimated regression coefficients (betas) for 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 consisted of one-sample t-tests performed on estimated betas of D_{KL} and I_S parameters in the data space channels × epoch time points (0-1000 ms). Results for each parameter were corrected for multiple comparisons using the *ept-TFCE* toolbox (Mensen & Khatami, 2013) in MATLAB, which combines threshold-free cluster enhancement (TFCE) technique in conjunction with permutation-based statistics (5,000 permutations, $\alpha = 0.05$).

3. Results

3.1 Behavioral results

Table 1 shows the results of the optimization of the Bayesian learner parameters. The optimized parameters were used to obtain the D_{KL} and I_S regressors for both behavioral and EEG inferential analyses. The correlation between the two regressors was 0.33.

Concerning RT inferential analyses, no term of the full model was removed by backward selection. Visual inspection of the full model 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 (1.5% of the trials). After outlier trials removal, the model was refitted and achieved reasonable closeness to normality.

A summary of the LMM results is presented in Table 2. The analysis revealed a significant effect of I_S. Figure 2 shows that 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. As depicted in Figure 2, the increase in RTs with increasing surprise (I_S) was moderately augmented with high D_{KL} in a multiplicative way. These results are highly consistent with our previous findings (see Table 1 and Figure 1 in Visalli et al., 2019, and Table 1 and Figure 1 in Visalli et al., 2021).

Table 1 | Ranges of free model parameters and optimized parameters from the RTs best-fitting model.

	Min	Max	Optimized parameters
α_L	.1	.9	**.2643**
τ_1	10	100	**84.7715**
τ_2	0.1	1	**0.3683**
β_S	1	10	**1**

Table 2 | Summary output of the final LMM model

Predictors	Estimates	S.E.	β	*p*-value
(Intercept)	5.69	0.019	-0.070	<0.001
Trial	0.007	0.005	0.035	0.106
PreRT	0.259	0.060	0.455	**<0.001**
D_{KL}	-0.003	0.003	-0.014	0.191
I_S	0.074	0.004	0.345	**<0.001**
$D_{KL} \times I_S$	0.003	0.001	0.014	**0.001**
Marginal R^2 / Conditional R^2		0.230 / 0.598		

Notes. PreRT: Response time at the preceding trial; D_{KL}: Kullback-Leibler divergence; I_S: Shannon information; S.E.: Standard Error; β: standardized coefficient.

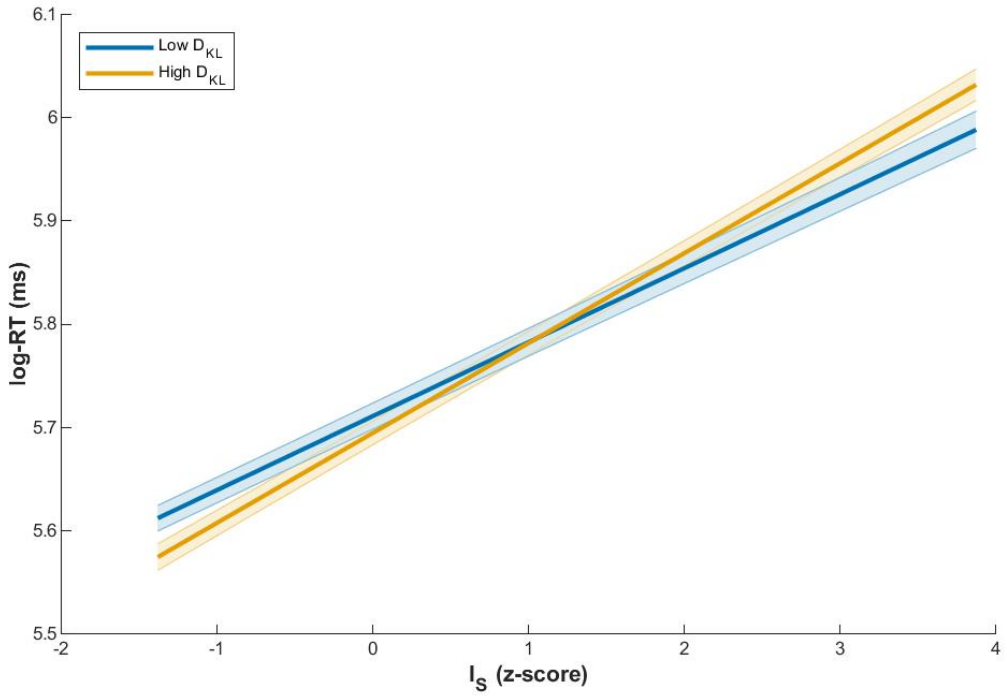

Figure 2 | Interaction effect between I_S and D_{KL} on log-transformed RTs. The figure shows the conditional effect of surprise (z-scored I_S) on log-RTs at the lowest (orange line) and highest (blue line) levels of updating (D_{KL}). Shaded error bars indicate standard errors of estimated marginal means.

3.2 EEG results

Results of the TFCE analyses on updating and surprise are shown in Figures 2 and 3, respectively.

Concerning updating (D_{KL}, Figure 3), we mainly observed two significant modulations. A first positive modulation emerged over centro-parietal scalp electrodes (CP1, P1, P3, Cz, CPz, Pz, POz, C2, CP2, P2) in a time window ranging from 150 to 300 ms. The second significant effect consisted of a late and sustained positive modulation extending approximately between 500 to 1000 ms over posterior electrodes (CP1, P1, P3, P5, P7, PO3, PO7, O1, CPz, Pz, POz, Oz, C4, C2, CP2, P2, P4, P6, PO4, PO8, O2). Correspondingly, a negative modulation was also found over frontal electrodes (Fp1, AF7, AF3, F7, F5, F3, F1, FT9, FT7, FC5, T7, Fpz, Fz, FCz, Fp2, AF8, AF4, F8, F6, F4, F2, FT10, FT8, FC6, FC4, FC2, T8).

Concerning surprise (I_S, Figure 4), we observed a significant positive modulation approximately between 260 and 520 ms over fronto-central electrodes (F5, F3, F1, FC5, FC3, FC1, C3, C1, C5, CP1, CP3, P1, Fz, FCz, Cz, CPz, F4, F2, FC6, FC4, FC2, C4, C2, CP2),

which was accompanied by a negativity mainly located at lateral and posterior channels (FT9, TP7, TP9, P3, P5, P7, PO3, PO7, O1, POz, Oz, FT10, T8, TP8, TP10, P4, P6, P8, PO4, PO8, O2). There was also an early significant positive modulation emerging at target onset and lasting for about 150 ms over central electrodes (FC3, FC1, C3, C1, CP1, CP3, P1, P3, FCz, Cz, CPz, Pz, F2, FC4, FC2, C4, C2, CP2, CP4, P2). Correspondingly, a negative modulation was observed over posterior channels (O1, Oz, FT10, TP8, TP10, P8, PO8, O2). Figure 4b shows that this early effect preceded target onset, thus suggesting an association with preparatory processes. However, the large range and variability of FPs employed in the present task (from 250 to 2500 ms) make it very difficult to analyze EEG signals during the FP. Therefore, we cannot draw any firm conclusion on the possible role of this effect, which will not be discussed further.

Results of the TFCE analyses for updating and surprise computed on the basis of the HGF model (Figure S1 and Figure S2, respectively) were remarkably similar as the above-reported ones.

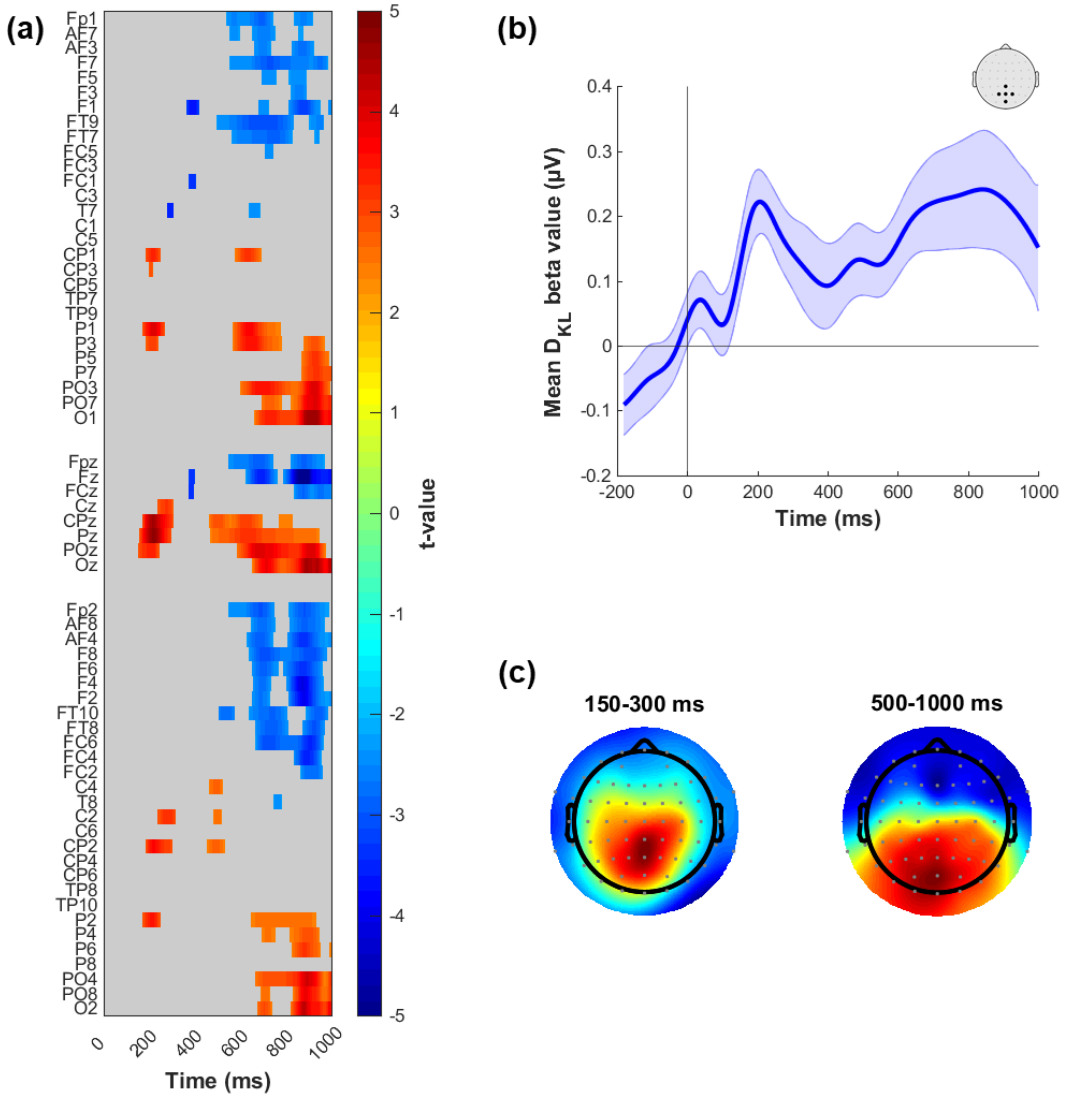

Figure 3 | Significant effects elicited by updating (D_{KL}) on target-locked EEG. (a) Raster diagram showing significant effects elicited by updating. Rectangles in warm and cold colors indicate electrodes/time points significantly modulated by the D_{KL} regressor. The colorbar indicates *t* values. Gray rectangles indicate electrodes/time points for which no significant modulations were observed. 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. **(b)** Trace-plot depicting 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 topoplot (i.e., CPz, P1, Pz, P2, POz). The shaded error bar indicates standard errors of the mean. **(c)** Topoplots showing the t values (same color scale of the raster diagram) averaged in the indicated time windows.

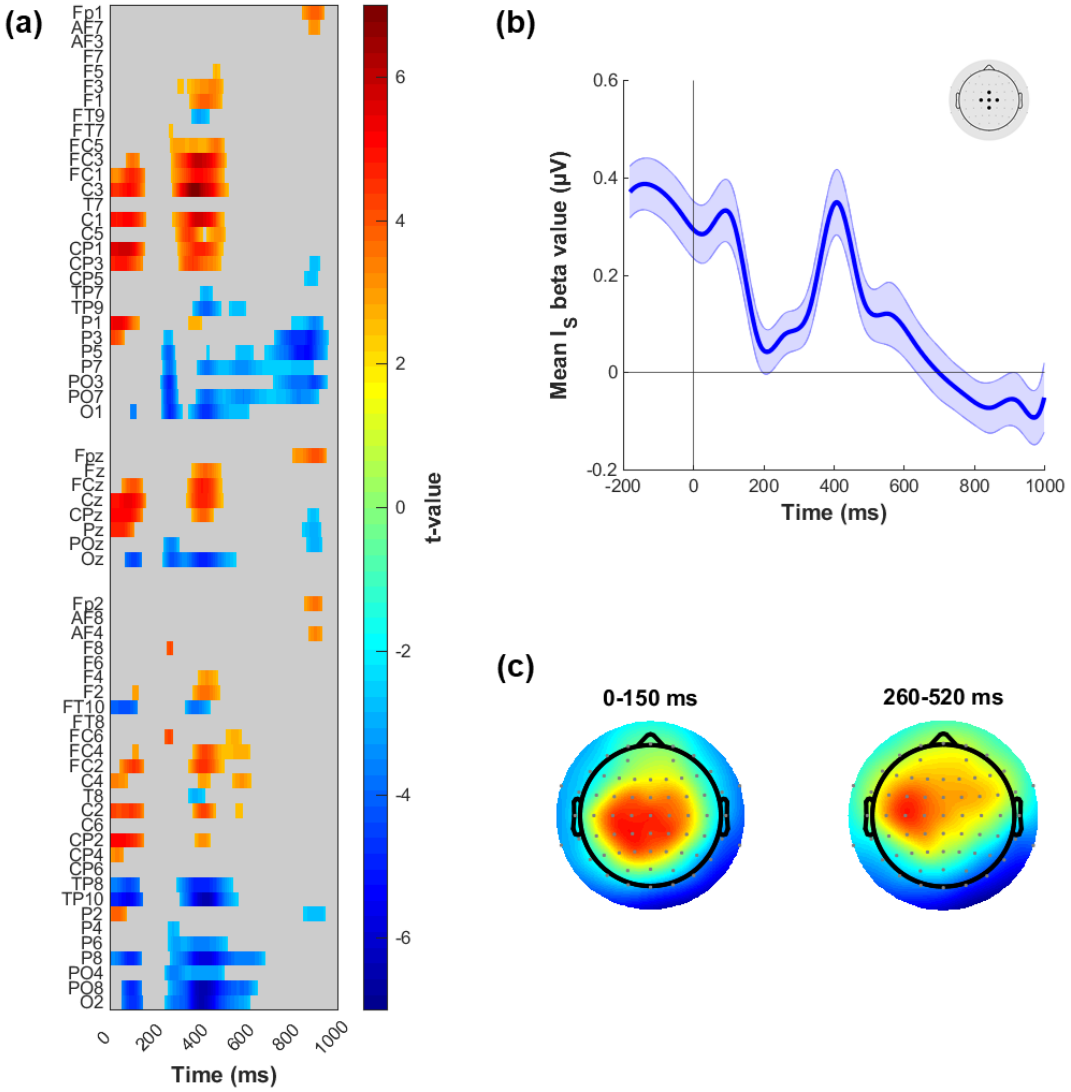

Figure 4 | Significant effects elicited by surprise (I_S) on target-locked EEG. (a) Raster diagram showing electrodes/time points significantly modulated by the I_S regressors. **(b)** Trace-plot depicting the mean I_S beta values (grand-average rERP) averaged over the electrodes indicated as black dots in the corresponding topoplot (i.e., FCz, C1, Cz, C2, CPz). **(c)** Topoplots showing the t values (same color scale of the raster diagram) averaged in the indicated time windows. For conventions, refer to Figure 3.

4. Discussion

The main aim of the present EEG study was to investigate how surprise and updating are encoded when changes in the temporal probabilities of the events are incidentally inferred. To summarize the main findings, while surprise modulated a fronto-central positivity, updating was instead associated with an earlier parietal P2-like component followed by a more posterior P3-like modulation. Extending our previous findings (Visalli et al., 2021), these results

confirmed that surprise and updating might be still distinguished at the electrophysiological level even when changes in temporal probabilities were incidentally learned.

Beginning with our surprise-related findings, the analyses showed a significant positive modulation peaking at around 400 ms over fronto-central electrodes. According to its timing and topographical scalp distribution, such a component could be viewed as belonging to the P3 family waveforms (Polich, 2007). This finding is in line with previous studies showing that trial-by-trial variability in P3 amplitudes reflects predictive surprise (Kolossa et al., 2013, 2015; Kopp et al., 2016; Mars et al., 2008; Modirshanechi et al., 2019). Moreover, once again we showed that a computational account of the P3 as a signature of surprise also holds in the temporal domain where the predictability of an event is modeled by the hazard function (Herbst et al., 2018).

Concerning updating, we corroborated our previous findings of a later P3 modulation by the updating of temporal expectations also in an incidental learning task. If, as mentioned above, several studies found that P3 amplitudes reflect surprise, other studies reported an association between P3 modulations and Bayesian updating (Bennett et al., 2015; Kolossa et al., 2015). Of note, the close relationship between updating and surprise motivated our previous cued paradigm that experimentally attempted to disentangle the two measures. In both this and our previous study (Visalli et al., 2021), we found that P3 amplitude variations are related to updating and that this modulation appeared later and was more sustained compared to the surprise-related P3. Overall, it seems that distinct P3 modulations encompass dissociable information-theory measures of updating and surprise and that these modulations are not contingent on the presence or absence of updating cues.

Following with updating-related findings, the P3 was preceded by a significant P2-like modulation peaking at around 200 ms over parietal sites. The precise functional meaning of such a component is still unclear (Luck, 2011). It has been suggested that the posterior P2 might be considered a signature of comparison mechanisms between information stored in memory and sensory input (Freunberger et al., 2007). In the literature on time perception, it has been also shown that the P2 amplitude reflected the distance between standard and comparison intervals in a temporal discrimination task, whereas it was not affected by the hazard rate of elapsed time (Kononowicz & van Rijn, 2014). Overall, these findings support the view of our P2-like component as an index of some comparison mechanisms involved in updating. Of note, this component was much more pronounced here as compared to our previous study (see Supplementary Figure S4). As one possibility, the P2 might be related to posterior computation. Indeed, in the previous study, participants were explicitly instructed about the transition between blocks (O'Reilly et al, 2013). This implies that, after update trials,

the posterior was computed starting from a uniform distribution as participants could disregard the prior to derive the posterior. Conversely, in the current study, the posterior was computed by integrating prior and likelihood. It thus makes sense that the enhanced P2 modulation observed here might reflect an increased computational complexity.

A relevant note concerns the characteristics of our experimental paradigm in the investigation of the neural underpinnings of updating and surprise. Indeed, even without an explicit manipulation, we were still able to differentiate between updating and surprise. This is because in temporal preparation tasks (like the FP task) updating and surprise rely on two different types of temporal information carried by the target. Briefly, surprise derives from the hazard rate of target onset that, as already explained, considers the passage of time during the FP. The information conveyed by the passage of time, however, is irrelevant for the updating of beliefs about target timing. Indeed, even a low surprising event such as a late target onset can induce a strong revision of our expectations. This peculiarity of temporal preparation tasks could be exploited in future research to improve our understanding of the computational meaning of several neural mechanisms. For example, relying on the P3 literature (see Nieuwenhuis, 2011), it might help to further characterize the distinct computational roles of neuromodulatory systems in encoding inferential information.

It should be acknowledged that there are multiple ways to formally define surprise (see Modirshanechi et al., 2022 for a systematic review). In recent years, an increasing number of studies have shown that different formalizations of surprise can be related to distinct EEG signatures (e.g., Gijsen et al., 2021; Kolossa et al., 2015; Visalli et al., 2021; Xu et al., 2021). Although computational approaches allow tracing terms or concepts (e.g., surprise, Bayesian surprise, postdictive surprise, belief updating, novelty) back to their formalization, it is helpful to elaborate more on the relation of our surprise measures with definitions used in previous studies. Here, the term “surprise” refers to “prediction surprise” (Modirshanechi et al., 2022), a measure also known as Shannon information (or also Shannon surprise or surprisal; Barto et al., 2013). Critically, surprise about event timing (Meindertsma et al., 2018) differs from surprise about other types of events since, as explained above, it does not represent a mismatch between the event and prior belief, but it is computed on the hazard rate of the observation. Concerning “updating”, we refer to an “information gain surprise” measure (Modirshanechi et al., 2022), also known as Bayesian surprise (Baldi & Itti, 2010; Itti & Baldi, 2009). Specifically, in this study, belief updating was not evaluated in the space of distributions over parameters, but over the space of observations (i.e., FP duration). To capture this difference, Kolossa and colleagues (2015), introduced the label “postdictive surprise” to refer to divergence (updating) between prior and posterior distributions over observation, while they used the term Bayesian surprise to describe belief updating in the parameter space. A further

difference between our work and previous studies investigating brain correlates of distinct surprise measures is that we employed continuous rather than multinomial stimuli. This is a relevant aspect, since links between information-theory measures and psychological constructs have been mainly proposed in relation to discrete states. To make an example, consider the debated distinction between surprise and novelty (Barto et al., 2013; Xu et al., 2021). Although it is easier to narratively conceive such a distinction with discrete variables and translate it into computational definitions, its formalization with continuous variables appears harder. Overall, the development of a comprehensive theory of surprise in the brain requires the well-considered integration of computational modeling and construct validation (Grahek et al., 2021).

Some aspects of the computational model used in the present study deserve attention. First, the model variant that considered learning at different time scales (Kolossa et al., 2013) provided a better fit of behavioral data as compared to the variant using a single temporal forgetting. This result is in line with recent findings showing that different brain signatures correlate with surprise values computed according to different integration time scales (Maheu et al., 2019). Taken together, these findings provide promising evidence for different neural mechanisms supporting Bayesian inference at different time scales. Second, the ideal Bayesian learner was adapted from our previous works to facilitate comparison between the present and previous findings. Although such a model was designed to capture all the peculiarities of the task (see O'Reilly et al., 2013), it represents "just a description of optimal behavior: it does not prescribe how Bayes optimal perception, sensorimotor integration or decision-making under uncertainty emerges" (Friston, 2012). A full understanding of how the brain updates temporal expectations needs more sophisticated models that consider the constraints deriving from the anatomy and the physiology of the brain, such as its functional hierarchical architecture or neuronal dynamics (e.g., Barry & Gerstner, 2022; Iigaya, 2016). In this respect, the fact that the use of an alternative and more biologically plausible model (i.e., the HGF; Mathys et al., 2011) led to very similar results corroborates our conclusions. Overall, our studies should be considered as a pioneering attempt towards the implementation of future models to understand how the brain refines temporal expectations.

To conclude, the present study isolated the electrophysiological correlates of surprise and updating of temporal expectations, highlighting both commonalities and differences between cued and incidental learning paradigms. Remarkably, the P3 family was proven again to be a valid index of inferential processes in the temporal domain even in the absence of explicit information, which reinforces the pivotal role of the P3 in understanding the Bayesian brain.

Funding: This study was in part supported by the European Research Council (ERC starting grant LEX-MEA, no. 313692, to A.Va), the “Department of Excellence 2018-2022” initiative of the Italian Ministry of University and Research (MIUR), awarded to the Department of Neuroscience – University of Padua, “Progetto giovani ricercatori” grants from the Italian Ministry of Health (project code: GR-2018-12367927 – FINAGE, to A.Va.; project code: GR-2019-12371166, to E.A.), the Spanish Ministry of Science and Innovation (MCIN/AEI/ 10.13039/501100011033) and “ERDF A way of making Europe” (grant: PID2021-128696NA-I00) to M.C., and by the PRIN 2020 grant (protocol 2020529PCP) from the Italian Ministry of University and Research (MUR) to E.A.. M.C. also acknowledges support of a María Zambrano Fellowship at the University of Granada from the Spanish Ministry of Universities and the European Union NextGeneration.

Ethics approval: All procedures performed were in accordance with the ethical standards of the 2013 Declaration of Helsinki for human studies of the World Medical Association. The procedures involved in this study were approved by the Bioethical Committee of the Azienda Ospedaliera di Padova.

Consent to participate: Participants gave their written informed consent to participate in the study.

Data and code availability: Task, model, and analysis codes and (raw) behavioral and EEG datasets are available at the Open Science Framework: osf.io/sdy8j/

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. *International Journal of Psychological Research*, *3*(2), 12–28.

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks: The Official Journal of the International Neural Network Society*, *23*(5), 649–666. doi:10.1016/j.neunet.2009.12.007

Barry, M., & Gerstner, W. (2022). *Fast Adaptation to Rule Switching using Neuronal Surprise* (p. 2022.09.13.507727). bioRxiv. doi:10.1101/2022.09.13.507727

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or Surprise? *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00907

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), Article 1. 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. *Nature Neuroscience, 10*(9), 1214–1221. doi:10.1038/nn1954

Bennett, D., Murawski, C., & Bode, S. (2015). Single-Trial Event-Related Potential Correlates of Belief Updating(1,2,3). *ENeuro*, *2*(5). doi:10.1523/ENEURO.0076-15.2015

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. doi:10.1163/156856897X00357

Bueti, D., Bahrami, B., Walsh, V., & Rees, G. (2010). Encoding of temporal probabilities in the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(12), 4343–4352. doi:10.1523/JNEUROSCI.2254-09.2010

Chaumon, M., Bishop, D. V. M., & Busch, N. A. (2015). A practical guide to the selection of

independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63. doi:10.1016/j.jneumeth.2015.02.025

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi:10.1016/j.jneumeth.2003.10.009

Ehinger, B. V., & Dimigen, O. (2019). Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ*, *7*. doi:10.7717/peerj.7838

Freunberger, R., Klimesch, W., Doppelmayr, M., & Höller, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience Letters*, *426*(3), 181–186. doi:10.1016/j.neulet.2007.08.062

Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage*, *62*(2), 1230–1233. doi:10.1016/j.neuroimage.2011.10.004

Gijsen, S., Grundei, M., Lange, R. T., Ostwald, D., & Blankenburg, F. (2021). Neural surprise in somatosensory Bayesian learning. *PLOS Computational Biology, 17*(2), e1008068. doi:10.1371/journal.pcbi.1008068

Grahek, I., Schaller, M., & Tackett, J. L. (2021). Anatomy of a Psychological Theory: Integrating Construct-Validation and Computational-Modeling Methods to Advance Theorizing. *Perspectives on Psychological Science, 16*(4), 803–815. doi:10.1177/1745691620966794

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

Harrison, L. M., Bestmann, S., Rosa, M. J., Penny, W., & Green, G. G. R. (2011). Time Scales of Representation in the Human Brain: Weighing Past Information to Predict Future Events. *Frontiers in Human Neuroscience*, *5*, 37. doi:10.3389/fnhum.2011.00037

Herbst, S. K., Fiedler, L., & Obleser, J. (2018). Tracking Temporal Hazard in the Human

Electroencephalogram Using a Forward Encoding Model. *ENeuro*, *5*(2). doi:10.1523/ENEURO.0017-18.2018

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, *13*(4–5), 411–430.

Ibrahim, J. G., Chen, M., Gwon, Y., & Chen, F. (2015). The power prior: Theory and applications. *Statistics in Medicine*, *34*(28), 3724–3749.

Iigaya, K. (2016). Adaptive learning and decision-making under uncertainty by metaplastic synapses guided by a surprise detection system. *ELife, 5*, e18073. doi:10.7554/eLife.18073

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. doi:10.1016/j.visres.2008.09.007

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?* Perception, 36 ECVP Abstract Supplement.

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*. 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. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 2931–2939. doi:10.1523/JNEUROSCI.2523-13.2014

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). Nova Science.

Kopp, B., Seer, C., Lange, F., Kluytmans, A., Kolossa, A., Fingscheidt, T., & Hoijtink, H. (2016). P300 amplitude variations, prior probabilities, and likelihoods: A Bayesian

ERP study. *Cognitive, Affective & Behavioral Neuroscience*, *16*(5), 911–928. doi:10.3758/s13415-016-0442-3

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), Article 1. doi:10.18637/jss.v082.i13

Kuznetsova, A., Christensen, R. H. B., Bavay, C., & Brockhoff, P. B. (2015). Automated mixed ANOVA modeling of sensory and consumer data. *Food Quality and Preference*, *40*, 31–38. doi:10.1016/j.foodqual.2014.08.004

Liakoni, V., Modirshanechi, A., Gerstner, W., & Brea, J. (2021). Learning in Volatile Environments With the Bayes Factor Surprise. *Neural Computation, 33*(2), 269–340. doi:10.1162/neco_a_01352

Luck, S. J. (2011). The Oxford Handbook of Event-Related Potential Components. In *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press. doi:10.1093/oxfordhb/9780195374148.001.0001

Lütkenhöner, B. (2010). Baseline correction of overlapping event-related responses using a linear deconvolution technique. *NeuroImage*, *52*(1), 86–96. doi:10.1016/j.neuroimage.2010.03.053

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.” *Journal of Neuroscience Methods*, *266*, 164–165. doi:10.1016/j.jneumeth.2015.12.003

Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *ELife, 8*, e41541. doi:10.7554/eLife.41541

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. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*,

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. *Frontiers in Human Neuroscience, 5*, 39.

Meindertsma, T., Kloosterman, N. A., Engel, A. K., Wagenmakers, E.-J., & Donner, T. H. (2018). Surprise about sensory event timing drives cortical transients in the beta frequency band. *Journal of Neuroscience*, *38*(35). doi:10.1523/JNEUROSCI.0307-18.2018

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

Meyniel, F., Maheu, M., & Dehaene, S. (2016). Human Inferences about Sequences: A Minimal Transition Probability Model. *PLOS Computational Biology, 12*(12), e1005260. doi:10.1371/journal.pcbi.1005260

Modirshanechi, A., Kiani, M. M., & Aghajan, H. (2019). Trial-by-trial surprise-decoding model for visual and auditory binary oddball tasks. *NeuroImage*, *196*, 302–317. doi:10.1016/j.neuroimage.2019.04.028

Modirshanechi, A., Brea, J., & Gerstner, W. (2022). A taxonomy of surprise definitions. *Journal of Mathematical Psychology, 110*, 102712. doi:10.1016/j.jmp.2022.102712

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 30*(37), 12366–12378. doi:10.1523/JNEUROSCI.0822-10.2010

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162.

Nieuwenhuis, S. (2011). Learning, the P3, and the locus coeruleus-norepinephrine system. *Neural Basis of Motivational and Cognitive Control*, 209–222.

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews. Neuroscience*, *19*(1), 34–48. doi:10.1038/nrn.2017.141

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

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), E3660-3669. doi:10.1073/pnas.1305373110

Pearson, J. M., Heilbronner, S. R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior Cingulate Cortex: Adapting Behavior to a Changing World. *Trends in Cognitive Sciences*, *15*(4), 143–151. doi:10.1016/j.tics.2011.02.002

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. doi:10.1016/j.clinph.2007.04.019

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 (New York, N.Y.)*, *193*(4258), 1142–1146. doi:10.1126/science.959831

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401.

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

Visalli, A., Capizzi, M., Ambrosini, E., Kopp, B., & Vallesi, A. (2021). Electroencephalographic correlates of temporal Bayesian belief updating and surprise. *NeuroImage*, *231*, 117867. doi:10.1016/j.neuroimage.2021.117867

Visalli, A., Capizzi, M., Ambrosini, E., Mazzonetto, I., & Vallesi, A. (2019). Bayesian modeling of temporal expectations in the human brain. *NeuroImage*, *202*, 116097. doi:10.1016/j.neuroimage.2019.116097

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. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society (EBMC)*, *2015*, 4101–4105. doi:10.1109/EMBC.2015.7319296

Xu, H. A., Modirshanechi, A., Lehmann, M. P., Gerstner, W., & Herzog, M. H. (2021). Novelty is not surprise: Human exploratory and adaptive behavior in sequential decision-making. *PLOS Computational Biology, 17*(6), e1009070. doi:10.1371/journal.pcbi.1009070

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692. doi:10.1016/j.neuron.2005.04.026