

Research Articles | Behavioral/Cognitive

Beta oscillations and pupil-linked arousal regulate the updating of cognitive models

<https://doi.org/10.1523/JNEUROSCI.0415-26.2026>

Received: 9 March 2026

Revised: 16 July 2026

Accepted: 14 August 2026

Copyright © 2026 Marshall et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

This Early Release article has been peer reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.jneurosci.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

Beta oscillations and pupil-linked arousal regulate the updating of cognitive models

Charlotte R. Marshall^{1,2*}, Mengxi Wang¹, Emma L. Lawrance^{1,3}, Nils Kolling^{1,4,5}, Jill X. O'Reilly¹

¹ Department of Experimental Psychology, University of Oxford, United Kingdom, OX1 3EL;

² Centre for Human Brain Health, School of Psychology, University of Birmingham, United Kingdom, B15 2TT;

³ Faculty of Medicine, Department of Surgery & Cancer, Imperial College London, United Kingdom, SW7 2AZ;

⁴ Université Lyon 1, INSERM, Stem Cell and Brain Research Institute, Lyon, France, 69500.

⁵ Centre Hospitalier Le Vinatier, Pôle EST, Bron, France

* Correspondence: jill.oreilly@psy.ox.ac.uk

Abbreviated title (32 characters): Tonic beta tracks model updating

Word count: Abstract (208), Introduction (647), Discussion (1429)

Conflict of Interest Statement: The authors declare no competing financial interests.

Acknowledgements: The authors would like to thank Sammi Chekroud for technical assistance with this project.

Funding: Medical Research Council Career Development Fellowship, MR/L019639/1 to JXOR, BBSRC sLOLA grant sLOLA Grant, BB/W003392/1 to JXOR, European Research Council (ERC to NK, grant FORAGINGCORTEX, project number 101076247

Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Code accessibility:

All code available on request by emailing the corresponding author

Abstract

The beta rhythm (neural oscillations in the range 13-30 Hz) is associated with the control of movement. Discrete movements (such as a button press) are bounded by pre-movement dip in beta power and post-movement beta rebound, which are thought to release and terminate the action, respectively. However, an additional or alternative role has been suggested for the beta rebound, in regulating the extent to which sensory feedback updates the task model. Using EEG and pupillometry in human volunteers of both sexes, we developed a task that dissociates updating of a cognitive model from motor control. We found that beta oscillations were modulated in two independent modes – phasic and tonic – with distinct computational correlates. When abstract task models are updated by integrating the outcomes of several actions, there is a tonic suppression of the beta rhythm, which originates at the time of the post movement rebound, spans multiple trials and is measurable at both the pre- and post- movement stages. Independently, phasic trial-to-trial fluctuations in beta power, primarily measurable in the pre-movement phase, were associated with motor readiness. The decrease in tonic beta power during model updating was mediated by an increase in tonic pupil dilation, a known correlate of model updating often considered a proxy for neuromodulatory tone.

Significance Statement

Beta oscillations have traditionally been linked to movement, but are increasingly implicated in learning and cognitive control. We show that these functions are supported by two computationally distinct modes of beta activity. A transient, phasic component reflects motor readiness immediately surrounding voluntary actions. In contrast, a sustained, tonic suppression of beta develops over multiple trials as evidence accumulates that an internal model of the environment should be updated. This tonic signal emerges from the post-movement beta rebound, tracks Bayesian estimates of state uncertainty, and parallels tonic pupil-linked arousal. These findings unify motor and cognitive accounts of beta oscillations by demonstrating that beta supports distinct computations on separate timescales.

Introduction

Sensorimotor beta oscillations (13-30Hz), arise over the frontal cortex and are reliably observed during voluntary movements (Jasper 1949). Whilst classic models suggest that beta oscillations gate the release of movement, more recent research suggest a role, specifically for post-movement beta, in controlling the learning rate for higher level sensorimotor models (the degree to which sensory feedback updates models), and theory suggests a more extensive role in controlling the update of cognitive models (Engel and Fries 2010).

During punctate, voluntary movements such as button presses, beta power follows a characteristic temporal motif: a pre-movement dip (Pfurtscheller 1981), followed by a post-movement rebound (Pfurtscheller, Stancak et al. 1996), both showing high intra-individual reliability (Espenhahn, de Berker et al. 2017). Beta oscillations are inhibitory of movement, for example in Parkinson's disease (Kuhn, Williams et al. 2004). Therefore an interpretation of the beta motif is that pre-movement beta dip 'releases' actions (by suppressing the holding beta rhythm), whilst a post-movement beta rebound terminates each action.

More recently, evidence suggests that beta oscillations not only gate the release of movement, but control the rate of updating of high-level motor plans by feedback (Gilbertson, Lalo et al. 2005), that is, the local learning rate. In particular, the post-movement beta rebound is well-timed to block sensory feedback, resulting from one's actions, from updating the sensorimotor model for future occasions. In motor learning, the beta rebound is reduced in conditions requiring update of a behavioural model, such as following large-magnitude movement errors in a visuomotor adaptation task (Tan, Jenkinson et al. 2014), especially when confidence in the world-model is low because the environment is unpredictable (Tan, Wade et al. 2016), and when sensory feedback is more informative because it is less noisy (Palmer, Auksztulewicz et al. 2019). Similarly, occipital beta power decreases when visual feedback is to be used, and increases when it was to be ignored (Limanowski, Litvak et al. 2020). Taken together, these results suggest that the phasic beta rebound acts at the time of feedback to block (or permit) the updating of a motor plan or model, thus acting at the level of motor planning rather than the immediate control of movement.

96 Despite this evidence, two important questions remain. First, although Engel and Fries
97 (2010) proposed a general role for beta oscillations in maintaining or updating cognitive
98 models, most existing work has focused on sensorimotor adaptation. Second, it remains
99 unclear whether update-related beta changes are necessarily confined to the phasic post-
100 movement rebound. In many cognitive tasks, model updating depends on integrating
101 evidence across multiple trials rather than responding to a single feedback event (Wilson,
102 Nassar et al. 2013)). If learning rate remains elevated across successive trials, beta activity
103 might also show sustained, tonic changes. Indeed, Engel and Fries proposed that tonic beta
104 activity promotes internally driven processing, whereas reduced tonic beta favours stimulus-
105 driven updating.

106
107 Here, we used a task that dissociates updating beliefs about the external environment from
108 motor execution and perceptual errors. This allowed us to distinguish the multiple factors
109 associated with the beta rhythm, across multiple timescales: motor readiness, feedback
110 processing relating to the current action and sensory stimulus, and higher-level uncertainty
111 about a model of the state of the world. This is distinct from previous research which has
112 generally used visuo-motor tasks, in which model feedback, model updating and uncertainty
113 tend to be highly correlated. We used a hierarchical perceptual decision-making task, in
114 which uncertainty varied independently at the level of individual decisions, and at the
115 (higher) level of the participant's beliefs about the current state of the world. We recorded
116 the beta rhythm using whole-head EEG and report distinct computational effects on the beta
117 rhythm in the phasic and tonic timescales. These effects were mirrored in parallel
118 modulation of pupil dilation on the phasic- and tonic timescales, linking tonic suppression of
119 beta oscillations to an existing literature in which pupil-linked arousal increases around
120 change points in the environment.

Methods

Participants

Thirty-six healthy right handed participants (aged 18-40 years), both male and female, took part in the study, which was approved by the University Research Ethics Committee. Participants were recruited from a student population at the University of Oxford and all gave written consent using a form approved by the ethics committee. All participants completed the total 360 trials of the experiment. Nine participants were excluded from analyses due to frequent eye movements (in over 25% trials their gaze deviated from 200 pixels on the x-axis from the screen centre), resulting in a final sample size of twenty-seven. Participants received reimbursement for their time upon completion of the experiment.

Experimental design and statistical analysis

Task

To dissociate between immediate feedback on a proximal action and higher-level model updating, we designed a hierarchical perceptual decision-making task in which negative feedback on a given trial could be attributed either to (one's own) poor performance, or to a mismatch between the expected and actual generative state of the environment.

<Figure 1 here>

Participants performed a hierarchical orientation judgement task (figure 1a). On each trial, participants made a **perceptual decision**: they judged the direction of rotation between two (sequential) grating stimuli; the difficulty of this judgement (angle of rotation) varied from trial to trial (figure 1b).

State inference task. The interpretation of feedback on each trial was complicated by a further factor: out of a choice of two stimulus-pairs (left side, right side), only one gave performance-contingent feedback (figure 1c). On each trial participants chose to respond to the left or right stimulus-pair. If the selected pair was the contingent pair, correct responses

were always met with positive feedback. If the selected pair was the non-contingent pair, positive or negative feedback was given at random. Which side (left/right) was contingent was stable for 5-15 trials at a time and then reversed without warning (**state reversal**). We refer to the task of inferring which stimulus-pair was contingent as the **state-inference task**. Note that participants were instructed of this manipulation before the task and given explicit practice with feedback to ensure they understood that only one stimulus gave contingent feedback at a given time.

Interpreting feedback. Because there was uncertainty at the perceptual level and the state level, when negative feedback was received, participants needed to decide whether this indicated a perceptual error (incorrect rotation reporting), or that they were working on the non-contingent stimuli, in other words that their model of the environment was inconsistent with its current state. A single instance of negative feedback was generally insufficient for a participant to conclude that the contingent side had changed, since negative feedback could also be attributed to a perceptual error. Therefore, maximising reward necessitated *tracking* feedback history over several trials, building up evidence for a decision to switch sides in the case of repeated negative feedback ('behavioural state reversal', figure 1d). This manipulation allowed us to investigate changes in beta power during extended periods of state uncertainty or updating, as distinct from individual feedback events.

Task details: Trial structure

Each trial began with the presentation of central fixation dot for 0.5s, followed by two Gabor patches oriented randomly between 30° and 60° to left and right of fixation on the horizontal midline subtending 18.3° visual angle for 0.5s. After a variable working memory interval (also called Inter-Stimulus interval or ISI) of 3-5s, mean 4s, a second fixation dot (0.5s) preceded re-presentation of the two gratings (0.5s), both rotated either clockwise or counter-clockwise from their initial positions by 0° (i.e., unchanged), 3°, 6°, 9° or 12°. The magnitude of rotation was the same for both gratings, but the direction of rotation was random. Participants chose a response from four buttons, indicating both the direction of rotation, and the stimulus (left or right) to which they were responding; this allowed us to monitor their belief about the contingent side as well as perceptual accuracy. There was a variable jitter of 300-800 ms, then feedback was indicated by a colour-change at the central fixation dot (1s). There was an ITI of 4s giving an average total trial duration of 12s.

Each participant completed 3 runs of the orientation judgement task, consisting of 120 trials, for a total of 360 trials.

Behavioural data analysis

Single-trial predictors of perceptual accuracy and reaction time

We expected that participants would be more likely to give a 'perceptually correct' answer (one in which the direction of rotation was reported correctly), and would show faster reaction times, when the angle of rotation between the first and second stimulus was larger and the working memory delay was shorter. We used multiple logistic- and linear regression, respectively, with two explanatory variables (angle of rotation α and duration of working memory interval, *ISI*) to test these hypotheses.

$$p(\text{choose 'clockwise'}) \sim [\text{constant} + \alpha + \text{signed ISI}]$$

Eq M1: Logistic regression for probability of perceptually correct response

Note that the response variable here was whether the answer the participant indicated clockwise rotation regardless of whether the feedback-contingent side was correctly chosen, which we have termed a 'perceptually correct' response. The explanatory variables were angle of rotation (here and only here, signed as positive for clockwise rotations) and the working memory interval or ISI, here and only here modified by a sign so that the prediction is positive (more likely to respond 'clockwise') for long intervals *when the correct response is clockwise*, and the converse (more likely to respond 'counterclockwise') for long intervals, *when the correct response is counter clockwise*. This is to allow a logistic curve to be fit to the binary response clockwise/counter clockwise.

For the analysis of reaction times we used a linear regression:

$$RT \sim [\text{constant} + \alpha + ISI]$$

Eq M2: Linear regression predicting reaction time based on task factors

In this case the unsigned angle of rotation (ie, difficulty) and working memory interval were used as predictors.

Single- and cross-trial predictors of switching

We expected that participants would (as instructed) attempt to choose the feedback-contingent side on each trial, but that there would be some delay after the contingent side reversed before participants switched responding to the new contingent side due to the need to accumulate evidence.

We modelled the probability of a behavioural state reversal (that is, the participant switching which side they respond to) after feedback on trial t using a lagged logistic regression in with binary explanatory variables that captured the presence of negative feedback on previous trials as follows:

$$\text{EV1: } x'_t = 1 \text{ if } (x_t = 0)$$

$$\text{EV2: } x'_{t-1} = 1 \text{ if } (x_t = 0 \text{ and } x_{t-1} = 0)$$

$$\text{EV3: } x'_{t-2} = 1 \text{ if } (x_t = 0 \text{ and } x_{t-1} = 0 \text{ and } x_{t-2} = 0)$$

$$\text{EV4: } x'_{t-3} = 1 \text{ if } (x_t = 0 \text{ and } x_{t-1} = 0 \text{ and } x_{t-2} = 0 \text{ and } x_{t-3} = 0)$$

Eq M3: Explanatory variables for lagged regression predicting side switch

The sequence was concluded if there was a behavioural state reversal (i.e., the participant switched which side they selected) at any point. In addition, we included EVs representing the perceptual difficulty (denoted as α , the angle of rotation between stimulus 1 and 2) on each negative feedback trial, as we expected that negative feedback would have a greater influence on state reversals when the perceptual judgement was easy (and the observer should have high confidence in their prediction). Difficulty was only modelled for trials with negative feedback, i.e.

$$\alpha'_t = \begin{cases} \alpha_t & \text{if } x'_t = 1 \\ 0 & \text{otherwise} \end{cases}$$

Eq. M4 Parametric modulation by difficulty

Thus the complete regression model was:

$$p(\text{switch}) \sim [\text{constant} + x'_{t-1} + x'_{t-2} + x'_{t-3} + \alpha_t + \alpha_{t-1} + \alpha_{t-2} + \alpha_{t-3}]$$

Eq M5: logistic lagged regression predicting side switch

Bayesian model

To illustrate how an ideal observer would accumulate evidence for a state reversal, we modelled the evidence in favour of each side being the deterministic side using a Bayesian learner model. The model gives the ‘evidence for/against side’ which will be plotted in Figures 5d and 6d.

The hypothesis space of the model was a simple binary: [L, R], representing the belief that the left or right stimulus-pair is feedback contingent, respectively. Because exactly one stimulus-pair was feedback contingent on each trial, the participant’s prior belief $p(L_t|x_{0:t-1})$ and posterior belief $p(L_t|x_{0:t})$ were scalars.

On a given trial t , the likelihood $p(L_t|x_t)$ was equivalent to the probability of obtaining the given feedback, given the angle of rotation and working memory delay on that trial, and assuming that L was in fact true. This was obtained for each individual using logistic regression (as described above) :

$$p(L_t|x_t) = p(x_t|L_t) \sim (\alpha + \text{ISI})$$

Eq M6: Likelihood function of Bayesian model

.. where α is the angle of rotation between first and second stimuli, determining task difficulty, and ISI is the duration of the working memory interval.

The posterior belief $p(L_t|x_{0:t})$ was then obtained using Bayes’ rule:

$$p(L_t|x_{0:t}) \propto p(L_t|x_t) \cdot p(L_t|x_{0:t-1})$$

Eq M7: Bayesian update

And the prior on the following trial was obtained from the posterior, using a simple leak that reflected the true frequency of reversals, $H = \frac{1}{10}$

$$p(L_{t+1}|x_{0:t}) = (p(L_t|x_{0:t}) \cdot (1 - H)) + ((1 - p(L_t|x_{0:t})) \cdot H)$$

Eq M8: Transition function from Bayesian model

The prior on the first trial of the run was set to 0.5.

The “evidence for side” plotted in figure 6d is the probability that the chosen side is indeed the correct side (noting that the chosen side changes at the point of the switch). The “evidence against side” in figure 7d is the complement of “evidence for side” – shown this way to match the polarity of the pupil effect and therefore aid visual comparison.

EEG acquisition

During the orientation judgement task EEG was recorded using a 61-channel 10-20 Easycap active EEG cap (Easycap GmbH, Wörthsee, Germany) and SynAmps RT Amplifier (Compumedics GmbH, Hamburg, Germany), at 1KHz, online referenced to the right mastoid.

EEG analysis

All EEG analysis were conducted in FieldTrip (Oostenveld, Fries et al. 2011). Data were imported, band-passed at 1-40Hz, and re-referenced to linked mastoids.

ICA cleaning was first performed on continuous data. For every participant, data from the three EEG runs were divided into 1-second epochs, concatenated, and then manual inspection was used to identify ‘bad’ EEG segments (function `ft_rejectvisual.m`, manual rejection based on epoch variance). Independent component analysis (method ‘runica’) (Bell and Sejnowski 1995) was then performed on all data epochs not marked as ‘bad’. Data were downsampled to 250Hz and PCA was used to reduce the rank of the data to 40 to speed up ICA. The first 20 ICA component time courses were then correlated with the horizontal and vertical EOG time courses to identify components containing primarily ocular artifacts. Components with an overall correlation of above 0.25 were marked for removal. This led to between 1 and 4 ICA components being marked for rejection per participant (mean = 2.2).

The original raw data were then re-epoching into 5-second trials around the button-press events (from 2.5 seconds pre-press, to 2.5 seconds post-press). These trial-based data were downsampled to 500Hz and the ICA unmixing weights from the continuous data were used to remove identified ICA components. Scalp current density (surface Laplacian (Oostendorp and van Oosterom 1996)) was estimated to increase spatial specificity of the EEG data. Time-frequency analysis was then performed for integer frequencies from 1-40Hz

using a 400ms sliding window multiplied with a Hanning taper moving across the entire 5-second epoch in 50ms increments.

We then averaged over a broad beta-band of 16-24Hz over sensors CP 3,1,Z,2,4 and P 3,1,Z,2,4. Although beta power is sometimes defined as 13-30 Hz, we focussed on the slightly narrower 16-24Hz band to minimize contamination from adjacent bands (in particular the alpha band, see Figure 2), and because inspection of the all-trials grand average suggested that this range contained the bulk of peri-movement beta modulation.

Whilst beta power is often studied as a lateralized phenomenon, in the current task, by design, the response hand could not be predicted on a given trial in advance of the actual response. Therefore, we did not expect to see lateralized effects of state inference. For completeness, we show here topographies for lateralized and non-lateralized beta band power. It can be seen that when all trials are averaged, the peak peri-movement beta dip occurs in the range 16-24 Hz at the CP and P sensors. In contrast, when beta power is plotted as a difference between sensors contra- and ipsilateral to the hand of response, the beta motif is seen at more anterior sensors (the C and CP sensors) and slightly later in time. The analysis presented hereafter focusses on non-lateralized beta power – to preview the results, no task variables of interest were found to significantly predict lateralized beta power.

<Figure 2 here>

Pupil data acquisition

Pupil dilation was tracked in a single eye at 1000Hz, using an Eyelink 1000 optical eye tracker (SR Research, Canada). Pupil data were pre-processed using custom software implemented in MATLAB. Data were smoothed with a Gaussian kernel (FWHM 50ms). Blinks were detected as rapid downward deflections in the pupil size trace; blinks plus a 75ms buffer either side were coded as missing data and excluded from further analysis.

Results

We first present behavioural analysis confirming that our manipulation of perceptual difficulty was effective, and that participants integrated feedback history with perceptual difficulty when detecting state reversals, before moving on to examine neural signals.

Both accuracy and reaction time depend on perceptual difficulty

<Figure 3 here>

We first confirmed that the perceptual difficulty manipulation was effective. Decreasing the magnitude of the angle between the first and second stimuli produced significant performance deficits; it both decreased the probability of correct response (fig2a) and increased mean reaction time (fig2b). In a logistic regression (Eq M1 in Methods) the effect of the signed angle of rotation on the reported direction of rotation was significant: across individuals, the mean effect was large: a 27% increase in the odds ratio for choosing the correct side, for each additional degree of rotation. This was statistically significant (sign-rank test for beta values across the group, $p=5.6 \times 10^{-6}$). We included the duration of the working memory interval (inter-stimulus interval; ISI) as a covariate as we reasoned longer intervals could lead to lower accuracy due to forgetting, however, the effect of ISI was near zero (0.6% change in the odds ratio for each additional second of ISI) and not statistically significant ($p = 0.67$, sign-rank test). Both angle magnitude and ISI significantly predicted reaction time (linear regression, Eq M2 in Methods): mean effect size across individuals for angle magnitude is -7.5 (7.5ms decrease in reaction time for each additional degree of rotation), which was statistically significant. For ISI the mean effect size was -8.9×10^{-3} , that is just under 1ms decrease in reaction time for each 1s increase in ISI; although small, representing less than 5ms difference in RT between the longest and shortest ISIs, this effect was statistically significant $t(26) = -2.30$, $p = 0.029$). The effects of decreasing angle magnitude on accuracy and reaction time were correlated across participants ($r(26) = -0.79$, $p = 7.9 \times 10^{-7}$). No such correlation was observed for the effects of ISI on accuracy and reaction time ($r(26) = 0.028$, $p = 0.89$). Since the above implies that the difficulty manipulation was effective, we expected participants to behave differently on easy vs difficult trials, when weighing up decisions to switch sides.

Participants integrated feedback history, but not perceptual difficulty history, into decisions to switch

Participants were significantly more likely to switch following a run of negative feedback, and were more likely to switch following negative feedback on an easy than a difficult trial, indicating that participants weighed both factors when evaluating whether to switch sides.

Statistically, we tested for these effects using a lagged logistic regression (Eq M3-M5 in *Methods*), predicting the probability of a side-switch based on negative feedback on the current and previous three trials, and the interaction of feedback with angle of rotation (inverse difficulty) (i.e., the angle of rotation was modelled for negative-feedback trials only). There was a significant effect of negative feedback on the current and previous trials. Median odds ratios for significant effects were as follows: participants ($n=27$) were 56.1 times more likely to switch following negative- than positive feedback on the current trial (p-value from sign-rank test on group beta values = 5.6×10^{-6}); indeed 95.4% of switches were on negative-feedback trials. Furthermore, participants were on average 2.26 times more likely to switch if there was, additionally, negative feedback on the previous trial (p-value from sign-rank test on group beta values = 4.42×10^{-5}). The average number of trials completed between a ground-truth reversal in the relevant side, and a behavioural switch is 7.4 (s.d. of participants means: 1.54), or 7.7 (s.d. 1.72), if limiting the analysis to 'correct' switches in which the participant really switches from the relevant to irrelevant side).

The greater the rotation between stimuli (on the current trial), the more likely participants were to switch (Figure 3c) - they were 1.11 times more likely to switch for each additional degree of rotation (p-value from sign-rank test on group beta values = 1.37×10^{-5}). The effect of negative feedback on trials t-2 and earlier, and of angle of rotation on all previous trials (t-1 and earlier) was non-significant ($p>0.1$).

This interplay between switching, feedback, and perceptual difficulty is logical; it suggests participants viewed the negative feedback on a difficult trial as ambiguous (possibly due to either a perceptual error or a side error) but were more likely to attribute negative feedback on an easy trial to a side error. The results suggest that, whilst feedback on the current trial is weighted by difficulty, feedback from previous trials may be maintained simply as binary outcomes (positive or negative) – although it should be noted that the lack of a historical effect of trial difficulty cannot be regarded as conclusive, as it could reflect a lack of sensitivity .

Reaction times decrease before a switch, but accuracy is maintained

Figure 4 shows reaction times, accuracy and frequency of positive feedback around the time of a switch. Reaction times and accuracy are shown both in raw form, and corrected for potential confounding factors (trial difficulty, ISI, and feedback valence on the previous trial) are accounted for (the black line in Figure 4a shows residual RT and accuracy after the above confounds have been modelled out using linear and logistic regression, respectively). Reaction time was significantly reduced on the trial preceding the switch, compared to the grand average over all trials ($t=-2.05$, $p=0.050$ for corrected reaction times); there was no significant effect on any other trial. There were no significant changes in accuracy.

Note that this analysis was limited to 'correct' switches, that is cases on which participants were working on the irrelevant side (perhaps following a reversal in the ground-truth relevant side) and switched to the relevant side. In the converse case ('incorrect switches') accuracy was necessarily low on the trial preceding the switch, presumably because in this case only inaccurate trials would generate negative feedback, which in turn triggered switching.

<Figure 4 here>

Physiological measures

Having established that the decision to switch was both affected by properties of the current stimulus, and evolved over trials due to reward history, we then sought to link these two phenomena to beta-band activity in the EEG and pupil dilation.

Beta oscillations

We defined two time windows for statistical analysis of beta-band activity – a pre-movement window, from 1000-500ms before movement onset, and a post-movement phase, from 500-1000ms after movement onset.

Throughout the analysis of EEG data we used the log of beta power as the dependent variable as $\log(\text{beta})$ was approximately normally distributed.

Pupil dilation

We log transformed our measure of pupil size (pupil area, expressed as a percentage of the session average pupil area). We also down-sampled the pupil data, which were recorded at 1000Hz, by averaging in 100ms bins, to increase SNR.

Pre- but not post- movement beta, and phasic but not tonic pupil dilation, reflect motor readiness

The phasic motif of pre-movement beta dip and post-movement beta rebound has been posited to control action release and termination. Quantitatively, the amplitude of the pre-movement beta power reduction predicts motor-readiness (reaction time and ability to cancel actions), but such evidence is lacking for the post-movement beta rebound. Therefore we first investigated whether response-readiness predicted beta band power both pre- and post-movement.

Predicted and residual RT

Before relating reaction time to beta power and pupil dilation, we first partitioned reaction time into a component that could be predicted based on task (perceptual difficulty and the duration of the working memory interval – Eq. M2 in *Methods*), and the (substantial) residual variation that could be predicted using task variables; we therefore defined ‘residual RT’ as the difference between the actual reaction time and that predicted based on perceptual difficulty and working memory interval.

$$RT = \beta_0 + \beta_1\alpha + \beta_2ISI + \textbf{Residual RT}$$

Eq 1. Definition of Residual RT

Residual RT likely reflects endogenous factors such as motor readiness or level of alertness; notably residual RT was higher on trials when participants made perceptual errors (paired t-test - $t_{26} = -7.99, p < 0.0001$), perhaps because participants were in a less attentive state on these trials. This effect was highest when errors were made on easy trials (interaction of perceptual correctness and difficulty: $t_{26} = -3.78, p < 0.0001$) - presumably, errors on easy trials occurred only in the most inattentive states.

Pre-movement beta depends on predicted and residual RT

Pre-movement beta power reduction is associated with greater motor-readiness (faster reaction times and reduced ability to cancel actions (He, Everest-Phillips et al. 2020, Wessel 2020)). Task factors affected reaction time, namely the angle of rotation (determining

difficulty) and the duration of the working memory interval (ISI). We expected the speedier responses associated with easy perceptual decisions and long ISIs to be associated with greater pre-movement reduction in beta power. However, reaction time varies over and above the expected effects dependent on task variables; we asked whether this residual RT, which putatively reflects endogenous variation in response-readiness, also affects the depth of the pre-movement beta dip.

We performed a linear regression predicting beta power in the pre-movement period, based on the elements of predicted RT (the angle of rotation and the working memory interval, termed ISI), and the measured residual RT:

$$\log(\beta_{pre}) \sim [\text{constant} + \alpha + \text{ISI} + \text{residual RT}]$$

Eq 2: linear regression predicting beta power by response readiness

... where α is the angle of rotation between stimulus 1 and 2 (equivalent to perceptual difficulty), ISI is the duration of the working memory interval, and residual RT is the difference between actual reaction time and the value predicted based on α and ISI in Eq 1 above.

We observed that factors which increased response-readiness (longer working memory interval and lower perceptual difficulty) were reflected in increased beta power reduction in the pre-movement phase, as was our measure of unpredicted variation in response-readiness (residual RT). Statistically, we carried out group-level t-tests on the coefficients from the participants-level GLMs (as per Eq. 2); time points where $p < 0.05$ with- and without Bonferroni-correction for 81 timepoints are indicated by heavy- and light lines in Figure 4, respectively.

Post-movement beta does not depend on predicted and residual RT

We then asked whether the converse effect was present in the post-movement period. We might have expected greater post-movement beta on trials with greater response-readiness, reflecting a movement termination process matched to the vigorous movement initiation. In fact, the factors affecting pre-movement beta had little effect in the post-movement phase, with only a briefly-significant effect of working memory interval evident in this time period (Figure 4 below).

The phasic pupil response decreases when 'response readiness' is higher

Mirroring the analysis of beta power, we used a time-resolved GLM to investigate whether pupil dilation varied with the 'response readiness factors' that drove reductions in pre-movement beta power.

$$\log(pupil) \sim [\text{constant} + \alpha + \text{ISI} + \text{residual RT}]$$

Eq 3: GLM response readiness - pupil

Pupil size was increased on trials with high residual RT (4) – this effect was significant ($p < 0.05$ two-tailed, Bonferroni corrected for 100 timepoints) for a window from 550 ms before movement onset to 2050 ms afterwards. This represents a modulation of the phasic component of the pupil response (Aston-Jones and Cohen 2005).

Neither of the task factors driving 'predicted RT' (difficulty and working memory interval or ISI) had a statistically significant effect on pupil size at any time point, suggesting that endogenous variations in arousal and motor readiness, rather than task factors were the key factor associated with pupil size in this phase.

<Figure 5 here>

State inference is reflected in post movement beta, and tonic beta power

Next we asked whether the detection of state reversals (i.e. the changes in which stimulus gave deterministic feedback that participants had to track) was reflected in changes in beta power or pupil dilation.

Previous research (Tan, Wade et al. 2016) suggests that the post movement beta rebound reflects a mechanism by which feedback is permitted or blocked from updating a motor model. However, in the current task, model updating or *state inference* involved integrating over multiple trials (because the meaning of feedback was modified by perceptual uncertainty) and the model to be updated was more abstract.

Behavioural analysis (lagged regression, described above) confirmed that participants accumulate evidence of a state reversal over at least two trials. We therefore reasoned that a component associated with state inference might be sustained over multiple trials in the present task.

To test this hypothesis, we ran a time-resolved GLM in which feedback valence on the current and previous trials were included as an explanatory variable (in addition to the response-readiness factors included in Eq 2 above).

$$\log(\beta_t) \sim [\text{constant} + x_t + x_{t-1} + \alpha + \text{ISI} + \text{residual RT}]$$

Eq. 4: Linear regression predicting beta power based on feedback history

... where, as in previous equations, α is the angle of rotation between stimulus 1 and 2 (equivalent to perceptual difficulty), ISI is the duration of the working memory interval, and residual RT is the difference between actual reaction time and the value predicted based on α and ISI as defined in Eq 1. x_t denotes feedback on trial t , coded as 1 if feedback is positive and zero if feedback is negative.

We found that indeed, there was a post-feedback decrease in beta power when the feedback was negative.

Additionally, there was a sustained decrease in beta power throughout the trial when the previous trial has received negative feedback (making the observer less confident/ more willing to update) rather than positive feedback (making the observer more confident in their model) - Figure 6 a,b. We refer to this sustained decrease in beta power, which builds up over several trials, as a 'tonic beta' effect.

<Figure 6 here>

Post-movement beta power decreases around the time of a state reversal

We next asked whether the feedback-related effects built up systematically around state reversals – state reversal was defined from participant behaviour, as the point at which participants switch from responding on one side, to the other. Visualisation of beta power on the trials preceding and following a switch (Figure 6 c,d) shows a large decrease in post-feedback beta on the final trial before the side switch. Post-feedback beta power remains low after the side-switch, and returns to baseline gradually over several trials. The reduction in beta power following the side-switch is sustained across the whole trial (figure 6e,f – summarized in figure 6d as the average beta power in the time window 1 to 0.5s before the movement on each trial). Trials on which beta power in each time window differed

significantly from the whole-task average (for that time window) were identified using a general linear model in which the effects of feedback valance, previous-trial feedback valance, task difficulty and residual RT were included as confound regressors (the red and black traces in figures 6c reflect beta power corrected for these confounds). Results were Bonferroni-corrected for 8 comparisons (four trials before and four trials after the state switch).

In summary, after a negative feedback event there was an immediate reduction in beta power (seen in the post-movement phase in Figure 6a). The reduction in beta power was sustained as a tonic effect throughout the following trial (seen in Figure 6b). In the run-up to a side switch, each negative feedback resulted in a further reduction in tonic beta power (red trace in figure 6c), which was sustained and compounded over subsequent trials (black and red traces in figure 6c). Following a side switch tonic beta power gradually returned to a higher level, apparently 'stepping up' each time a positive feedback event occurred (figures 6c).

Tonic beta power tracks the evidence for a state reversal

To provide a more mechanistic characterization of the changes in beta power, we constructed an optimal Bayesian observer model (see *Methods*) that inferred, given the feedback and sensory uncertainty on each trial, the probability that the currently selected side was the deterministic side. The prior- and posterior- belief that the currently selected side is the correct (deterministic) one is plotted in figure 6d. This belief falls gradually over a number of trials but falls dramatically between the prior- and posterior on the final trial before a (behavioural) reversal. The strength of belief that the currently selected side is indeed correct ('state belief') therefore strongly resembles the time-course of the tonic beta effect.

No significant effects of task variables on lateralized beta power

Whilst beta power is often studied as a lateralized phenomenon, in the current task, by design, the response hand could not be predicted on a given trial in advance of the actual response. Therefore, we did not expect to see lateralized effects of state inference. For completeness, we tested for effects of predicted RT, residual RT and trials until switch, using lateralized beta power as the response variable (left-right reversing topographies for right-hand response trials to allow for averaging across all trials). As expected, on average across all trials, we observed a lateralized version of the peri-movement beta motif (pre-movement

decrease in beta power followed by a post movement beta rebound, contralateral to the response hand), which was located anterior to the observed non-lateralized effects, being most prominent over central electrodes (C1/2, C3/4, CP1/2 and CP3/4). Lateralized beta power around the time of the button-press was not affected by state inference (contra-ipsi beta power did not depend on the number of trials until a switch). There was a trend towards an effect of motor readiness (residual RT) on lateralized beta power, in the rising phase between the peri-movement minimum and post-movement rebound; this effect lay outside of pre-planned time windows and was not significant with multiple comparisons correction for all timepoints.

No post-error slowing effects in behaviour or beta power

Following a negative feedback trial, we might expect decreased response-readiness (post error slowing) and a concomitant increase in pre-movement beta power. In fact, there was no significant increase in reaction time following negative feedback on either perceptually correct trials ($t_{26} = -0.95$, $p = 0.35$, two tailed, when negative feedback on the previous trial was added as a co-regressor to Eq M2) or perceptually incorrect trials ($t_{26} = 1.00$, $p = 0.33$, two tailed). If there had been a behavioural post-error slowing effect, we would have expected to see an increase in pre-movement beta power (reflecting increased response inhibition); since there was no behavioural post-error slowing effect this was not necessarily expected. However, pre-movement beta power actually showed the opposite effect, that is it *decreased* following negative feedback (effect of previous trial feedback, x_{t-1} in the model updating GLM, eq 4 – see Figure 6b). Rather than a change in response caution, this sustained decrease in beta power following negative feedback may reflect that the state model is susceptible to updating.

Tonic pupil changes also track state uncertainty, mirroring changes in tonic beta

The effects seen for tonic beta power are reminiscent of pupil dilation effects in which the baseline (tonic) pupil size increases around the time of state reversals ((Nassar, Rumsey et al. 2012, Browning, Behrens et al. 2015, Muller, Mars et al. 2019).

Therefore, mirroring the analyses performed on beta power, we asked whether feedback on the current and previous trials affected pupil size:

$$\log(pupil_t) \sim [\text{constant} + x_t + x_{t-1} + \alpha + \text{ISI} + \text{residual RT}]$$

Eq. 4: Linear regression predicting pupil size based on feedback history

633

634 The results mirrored the effects seen for tonic beta power. After a negative feedback event
635 there was an immediate increase in tonic pupil size (seen in the post-feedback phase in
636 Figure 7a). The increase in pupil size was sustained as a tonic effect throughout the
637 following trial (seen in Figure 7b). In the run-up to a side switch, each negative feedback
638 event resulted in a further increase in tonic pupil size (red trace in figure 7c), which was
639 sustained and compounded over subsequent trials (black and red traces in figure 7c).
640 Following a side switch tonic pupil size gradually returned to a lower level, apparently
641 'stepping down' each time a positive feedback event occurred (figure 7c). As for tonic beta
642 power (above), the changes in tonic pupil size reflect closely the strength of belief that the
643 alternative side is indeed correct ('state belief') from an optimal Bayesian model (Figure 7d).
644 Note that here we plot the strength of belief *against* the currently selected side (which is
645 exactly the complement of the parameter plotted for beta power in Figure 6) to match the
646 polarity of the pupil dilation effect and therefore aid visual comparison.

647

648 <Figure 7 here>

649

650 **Post-feedback beta power mediates the effects of task on pupil dilation**

651 The parallel effects observed for beta power and pupil dilation hint at a common neural
652 mechanism driving both EEG and pupil effects. To investigate this hypothesis, we conducted
653 two formal mediation analyses.

654 Firstly, we asked whether the effect of feedback valence on pupil dilation in the post-
655 feedback period (2-4s post feedback) was mediated by beta power in the post-feedback
656 window (500-1000ms after feedback) – this links the parallel EEG and pupillometry results
657 shown in figure 6a,b and 7a,b. Secondly, we asked whether the effect of response-readiness
658 (residual-RT) on the phasic pupil response (0-2s after the button-press) was mediated by
659 pre-movement beta power reduction (in the window 1000-500ms prior to the button press) –
660 this links the parallel EEG and pupillometry results shown in Figure 5.

661 *Feedback effects*

662 Following the procedure of Barron and Kenny (1996) we first confirmed the effect of our
663 independent variable, feedback valence, in the context of all other task factors, on the
664 dependent variable which was pupil size in the post-feedback window (2-4s post feedback,
665 shaded region on figure 7 a,b,d,f): this is the analysis in Eq. 4 and the effect of the

independent on dependent variable was significant in the chosen time window– the pupil is larger after negative feedback $t_{26} = 6.69, p < 0.005$.

We then verified that the independent variable, in the context of the other task variables, predicts the putative mediator, post-feedback beta power (in the window 500-1000ms post-feedback). This is the analysis in Eq 3 and the effect of the independent on mediator variable was significant – the beta power is lower after negative feedback $t_{26} = 7.41, p < 0.005$.

We then modelled the effect of feedback valence on the pupil size in the context of the mediator (beta power),

$$\log(pupil_t) \sim [\text{constant} + x_t + x_{t-1} + \alpha + \text{ISI} + \text{residual RT} + \log(\beta_t)]$$

Eq. 5 Mediation of feedback effect on pupil, by beta power

Using a paired t-test, we asked whether the coefficient for the effect of feedback valence on pupil size was significantly reduced in the presence of the mediator. This was indeed the case $t_{26} = 2.96, p = 0.0067$, indicating that beta power significantly mediates the effect of feedback valence on pupil dilation.

Response-readiness effects

Next, we asked whether the effect of response-readiness (residual-RT) on the phasic pupil response (0-2s after the button-press) was mediated by pre-movement beta power reduction (in the window 1000-500ms prior to the button press). Following the Barron-Kenny procedure as above, we verified that the effect of residual RT on pupil size was significant (Eq 3 above) – it was, $t_{26} = 6.53, p < 0.005$. We also verified that residual RT predicted beta power in the pre-movement interval (Eq 2 above) – it was, $t_{26} = 6.33, p < 0.005$. We then asked whether there was a significant mediation effect, using a paired t-test to compare the coefficients for the effect of residual RT on pupil size with and without the inclusion of pre-movement beta power

$$\log(pupil_t) \sim [\text{constant} + \alpha + \text{ISI} + \text{residual RT} + \log(\beta_t)]$$

Eq. 6 Mediation of response-readiness effect on pupil, by beta power

However there was no reduction $t_{26} = -0.42, p = 0.68$ indicating that the effects of response-readiness (residual RT) on the pre-stimulus beta power and the phasic pupil response are independent.

Time-resolved relationship between beta power and pupil dilation

Finally, to visualize the relationship between beta power and pupil size, we used a two dimensional time-resolved regression in which we predicted pupil size at each timepoint i from beta power at each timepoint j .

$$\log(pupil_i) \sim [\text{constant} + \log(\beta_j)]$$

Eq 7: predicting pupil size from raw beta power

It can be seen (Figure 8a) that increased beta power in the post-feedback period (in the time period in which updating effects were observed) was associated with increased pupil dilation throughout the post-feedback epoch, perhaps reflecting a state in which model updating is enabled.

Additionally, decreased beta power in the pre-movement phase (the same time window in which beta power tracks residual RT) was associated with decreased pupil size across most of the trial epoch, perhaps reflecting an overall 'on task' or focussed state (Aston-Jones and Cohen 2005).

<Figure 8 here>

Discussion

We examined distinct computational influences on beta power: motor readiness, sensory feedback and model updating, and their parallels in pupil dilation.

The results suggest that beta oscillations, and pupil dilation, operate on two distinct timescales, with different computational drivers and consequences. Firstly (as expected), a 'phasic' effect is seen: in the few hundred milliseconds around movement release, there is a transient suppression of the beta rhythm (before the movements) and rebound (after the movement) which are associated with movement release and termination. This effect is paralleled by changes in the size of the phasic pupil response. Secondly, slower 'tonic' variations in beta power and pupil size occur spanning multiple movements over tens of seconds. These tonic changes appear to track the strength of evidence for a state reversal (as captured in a Bayesian ideal observer model). The tonic beta- and pupil effects seem to reflect a common underlying neural mechanism, as the effects of task on pupil size are statistically mediated by tonic beta power. This tonic effect is much less well known, to date, than links between phasic beta and response-readiness.

It is established in the literature, and reconfirmed here, that motor readiness and arousal primarily affect the phasic beta time course, specifically, pre-movement beta power reduction is strongly associated with motor readiness. When the pre-movement beta dip is large, reaction times are faster and participants are less able to 'cancel' planned actions (as in the stop signal task (Swann, Tandon et al. 2009, Wessel 2020)). Movements initiated during the beta-rebound caused by a previous movement (i.e., when beta activity is high) are slower (Muralidharan and Aron 2021). Brain stimulation studies confirm a causal role for pre-movement beta desynchronisation in movement speed: Volitional decrease of movement-related beta via neurofeedback produces faster movements (He, Everest-Phillips et al. 2020), and transcranial entrainment of beta oscillations via beta-frequency alternating current stimulation (TACS) slows movement (Pogosyan, Gaynor et al. 2009). In the present study, as expected, pre-movement beta power correlated with motor readiness, operationalized as residual variation in reaction time that could not be explained by task factors.

Interestingly, motor readiness also correlated with the size of the phasic pupil response, perhaps reflecting a relationship between pupil size and arousal (Aston-Jones and Cohen 2005). However, statistical mediation analysis suggested that phasic beta and phasic pupil correlates of response readiness may have independent mechanisms, as there was no evidence that beta mediates the link between response readiness (as measured by reaction time) and pupil size.

The key finding of this study is that a second component of the beta rhythm reflects a more abstract or higher-order feature of the environment – beta power declines tonically over the course of several trials as participants accumulate that their (abstract) model of the environment should be updated, and rises gradually over several trials after such an update. This is mirrored by changes in tonic pupil dilation, and these two biomarkers may reflect a common underlying mechanism, as formal mediation analysis shows that changes in tonic beta mediate the effect of task on tonic pupil size. The changes in tonic beta power and pupil dilation closely resemble the strength of evidence for a state reversal as captured using a Bayesian ideal observer model.

We describe the model-update effects on beta power and pupil dilation as tonic, because they are long lasting, spanning multiple trials, and is sustained independently of the time-course of individual actions or trials; this is distinct from the phasic pre-movement beta desynchronization and post-movement beta rebound which form a characteristic movement-locked time course, and the phasic pupil effect, a characteristic sharp dilation and contraction of the pupil that occurs after movement release. However, the model-update component in tonic beta does appear to have a close link with the phasic beta rebound: When feedback is received that conflicts with the current model of the environment, the beta rebound is suppressed. This suppression is carried over into a tonic effect.

The results are concordant with previous findings that phasic post-movement beta power is linked to model updating. It has previously been reported that the post-movement beta rebound is decreased when feedback should be permitted to update an internal model, and increased when feedback should be blocked. In a sensorimotor context, the beta rebound is reduced 1) when there is a prediction error (Tan, Jenkinson et al. 2014) , 2) when the internal model is uncertain because of noise (Tan, Wade et al. 2016), and 3) when the sensory feedback is relatively certain (low noise) (Palmer, Aukstulewicz et al. 2019). Beyond the sensorimotor framework, beta power in occipital cortex is reduced when visual feedback should be used to update a model and increased when visual feedback should be ignored (Limanowski, Litvak et al. 2020). In decision-making, the beta rebound is reduced when exploratory (rather than exploitative) choices are made, and especially when the feedback confirms that switching preference to the explored option would be beneficial (Chernyshev, Pultsina et al. 2023).

The tonic beta effect may be more evident in the present study than in previous work because of difference in the task design. Most previous reports of update-related modulations of the beta rhythm come from visuomotor adaptation paradigms, in which model update tends to be strongly correlated with feedback valence (and magnitude), and the model is continuously updated. In contrast, we used a hierarchical perceptual decision

making task designed to partially dissociate model updating from both feedback valence and error magnitude. The two key features of the task were that negative feedback should be given a variable weight depending on the difficulty of the perceptual decision, and that the detection of state switches (need for model update) depended in the integration of several feedback events, culminating in rare 'complete switches' of the model rather than incremental updating.

Two studies have reported effects in the pre-movement 'beta desynchronisation' period, consistent with our observed tonic effect. 1) In a sensorimotor adaptation study, Palmer et al (2019) showed that the pre-movement beta desynchronisation was increased when model uncertainty was high (following a sensorimotor perturbation). 2) in a decision-making study, Chernyshev et al (2020) found that (phasic) pre-decision beta power was reduced when people made exploratory decisions. In both cases, because trial-wise baseline correction (or mean correction) was employed, tonic effects as reported in the present manuscript would be less apparent, and might appear only as modulations of phasic effects.

In the pre-movement beta phase, we observe a juxtaposition of two apparently independent effects on beta power. The magnitude of the phasic pre-movement beta dip was most closely linked to trial-by-trial variations in motor readiness, independently of model certainty; however, we also recorded effects in the pre-movement beta period of model updating, as a part of a long-timescale, tonic effect on beta power. This effect was independent of motor-readiness effects; indeed, tonic beta power was suppressed following a model update (figure 5c), but in this same time period reaction times are not affected (Figure 4). This juxtaposition suggests two functionally distinct components to the beta rhythm, even within the same (pre-stimulus) time period – as the tonic beta suppression associated with model updating does not result in faster reaction times, even though beta power is in fact lower in the pre-stimulus window associated with motor readiness

There are two possible interpretations of this two-timescale beta phenomenology. One is that a single beta generator tracks both readiness to respond to and update the model underlying our responses; as these processes naturally happen on different timescales, modulation of beta on two timescales is seen. An alternative explanation is that beta power as measured in the scalp EEG is an overlay of contributions from two sources, associated with motor readiness/gating, and model updating, respectively. MEG studies using beamforming have previously suggested that pre-movement beta desynchronization, and post-movement beta rebound may be topographically dissociable: Beta rebound tends to originate from a relatively anterior source, and beta desynchronisation relatively posterior, in sensorimotor cortex (Jurkiewicz, Gaetz et al. 2006, Fry, Mullinger et al. 2016, Pakenham, Quinn et al. 2020). Outside the field of beta oscillations, model updating is associated with

825 activity in the medial prefrontal cortex and anterior cingulate cortex (Holroyd and Coles
826 2002, Alexander and Brown 2011, O'Reilly, Schuffelgen et al. 2013) which lie anterior to the
827 motor cortex. The present study, using scalp EEG which has relatively poor ability to
828 dissociate neural sources, cannot distinguish between sources of the tonic and phasic
829 components; however the independent influences of phasic (pre-movement) beta and tonic
830 beta on reaction time, and the differential relationship of tonic- and phasic beta power to
831 pupil dilation, argue in favour of a two-source model.

JNeurosci Accepted Manuscript

References

- Alexander, W. H. and J. W. Brown (2011). "Medial prefrontal cortex as an action-outcome predictor." Nat Neurosci **14**(10): 1338-1344.
- Aston-Jones, G. and J. D. Cohen (2005). "An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance." Annu Rev Neurosci **28**: 403-450.
- Bell, A. J. and T. J. Sejnowski (1995). "An information-maximization approach to blind separation and blind deconvolution." Neural Comput **7**(6): 1129-1159.
- Browning, M., T. E. Behrens, G. Jocham, J. X. O'Reilly and S. J. Bishop (2015). "Anxious individuals have difficulty learning the causal statistics of aversive environments." Nat Neurosci **18**(4): 590-596.
- Chernyshev, B. V., K. I. Pultsina, V. D. Tretyakova, A. S. Miasnikova, A. O. Prokofyev, G. L. Kozunova and T. A. Stroganova (2023). "Losses resulting from deliberate exploration trigger beta oscillations in frontal cortex." Front Neurosci **17**: 1152926.
- Engel, A. K. and P. Fries (2010). "Beta-band oscillations--signalling the status quo?" Curr Opin Neurobiol **20**(2): 156-165.
- Espenhahn, S., A. O. de Berker, B. C. M. van Wijk, H. E. Rossiter and N. S. Ward (2017). "Movement-related beta oscillations show high intra-individual reliability." Neuroimage **147**: 175-185.
- Fry, A., K. J. Mullinger, G. C. O'Neill, E. L. Barratt, P. G. Morris, M. Bauer, J. P. Folland and M. J. Brookes (2016). "Modulation of post-movement beta rebound by contraction force and rate of force development." Hum Brain Mapp **37**(7): 2493-2511.
- Gilbertson, T., E. Lalo, L. Doyle, V. Di Lazzaro, B. Cioni and P. Brown (2005). "Existing motor state is favored at the expense of new movement during 13-35 Hz oscillatory synchrony in the human corticospinal system." J Neurosci **25**(34): 7771-7779.
- He, S., C. Everest-Phillips, A. Clouter, P. Brown and H. Tan (2020). "Neurofeedback-Linked Suppression of Cortical beta Bursts Speeds Up Movement Initiation in Healthy Motor Control: A Double-Blind Sham-Controlled Study." J Neurosci **40**(20): 4021-4032.
- Holroyd, C. B. and M. G. H. Coles (2002). "The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity." Psychol Rev **109**(4): 679-709.
- Jasper, H., Penfield, W. (1949). "Zur Deutung des normalen Elektrencephalogramms und selner Veriinderungen . Electroencephalograms in man : Effect of voluntary movement upon the electrical activity of the precentral gyrus." Arch. Psychiatr. Nervenkr. **174**: 163-174.
- Jurkiewicz, M. T., W. C. Gaetz, A. C. Bostan and D. Cheyne (2006). "Post-movement beta rebound is generated in motor cortex: evidence from neuromagnetic recordings." Neuroimage **32**(3): 1281-1289.

870 Kuhn, A. A., D. Williams, A. Kupsch, P. Limousin, M. Hariz, G. H. Schneider, K. Yarrow and
871 P. Brown (2004). "Event-related beta desynchronization in human subthalamic nucleus
872 correlates with motor performance." Brain **127**(Pt 4): 735-746.

873 Limanowski, J., V. Litvak and K. Friston (2020). "Cortical beta oscillations reflect the
874 contextual gating of visual action feedback." Neuroimage **222**: 117267.

875 Muller, T. H., R. B. Mars, T. E. Behrens and J. X. O'Reilly (2019). "Control of entropy in
876 neural models of environmental state." Elife **8**.

877 Muralidharan, V. and A. R. Aron (2021). "Behavioral Induction of a High Beta State in
878 Sensorimotor Cortex Leads to Movement Slowing." J Cogn Neurosci **33**(7): 1311-1328.

879 Nassar, M. R., K. M. Rumsey, R. C. Wilson, K. Parikh, B. Heasly and J. I. Gold (2012).
880 "Rational regulation of learning dynamics by pupil-linked arousal systems." Nat Neurosci
881 **15**(7): 1040-1046.

882 O'Reilly, J. X., U. Schuffelgen, S. F. Cuell, T. E. Behrens, R. B. Mars and M. F. Rushworth
883 (2013). "Dissociable effects of surprise and model update in parietal and anterior cingulate
884 cortex." Proc Natl Acad Sci U S A **110**(38): E3660-3669.

885 Oostendorp, T. F. and A. van Oosterom (1996). "The surface Laplacian of the potential:
886 theory and application." IEEE Trans Biomed Eng **43**(4): 394-405.

887 Oostenveld, R., P. Fries, E. Maris and J. M. Schoffelen (2011). "FieldTrip: Open source
888 software for advanced analysis of MEG, EEG, and invasive electrophysiological data."
889 Comput Intell Neurosci **2011**: 156869.

890 Pakenham, D. O., A. J. Quinn, A. Fry, S. T. Francis, M. W. Woolrich, M. J. Brookes and K. J.
891 Mullinger (2020). "Post-stimulus beta responses are modulated by task duration."
892 Neuroimage **206**: 116288.

893 Palmer, C. E., R. Aukstulewicz, S. Ondobaka and J. M. Kilner (2019). "Sensorimotor beta
894 power reflects the precision-weighting afforded to sensory prediction errors." Neuroimage
895 **200**: 59-71.

896 Pfurtscheller, G. (1981). "Central beta rhythm during sensorimotor activities in man."
897 Electroencephalogr Clin Neurophysiol **51**(3): 253-264.

898 Pfurtscheller, G., A. Stancak, Jr. and C. Neuper (1996). "Post-movement beta
899 synchronization. A correlate of an idling motor area?" Electroencephalogr Clin Neurophysiol
900 **98**(4): 281-293.

901 Pogosyan, A., L. D. Gaynor, A. Eusebio and P. Brown (2009). "Boosting cortical activity at
902 Beta-band frequencies slows movement in humans." Curr Biol **19**(19): 1637-1641.

903 Swann, N., N. Tandon, R. Canolty, T. M. Ellmore, L. K. McEvoy, S. Dreyer, M. DiSano and
904 A. R. Aron (2009). "Intracranial EEG reveals a time- and frequency-specific role for the right
905 inferior frontal gyrus and primary motor cortex in stopping initiated responses." J Neurosci
906 **29**(40): 12675-12685.

- Tan, H., N. Jenkinson and P. Brown (2014). "Dynamic neural correlates of motor error monitoring and adaptation during trial-to-trial learning." J Neurosci **34**(16): 5678-5688.
- Tan, H., C. Wade and P. Brown (2016). "Post-Movement Beta Activity in Sensorimotor Cortex Indexes Confidence in the Estimations from Internal Models." J Neurosci **36**(5): 1516-1528.
- Wessel, J. R. (2020). "beta-Bursts Reveal the Trial-to-Trial Dynamics of Movement Initiation and Cancellation." J Neurosci **40**(2): 411-423.
- Wilson, R. C., M. R. Nassar and J. I. Gold (2013). "A mixture of delta-rules approximation to bayesian inference in change-point problems." PLoS Comput Biol **9**(7): e1003150.

Figure Legends

Figure 1: Behavioural task. A) Structure of a single trial; the mean total trial duration is 12s, including an ITI of 4s. B) Example difficulty manipulation: Second stimuli (s2) could rotate either clockwise or counter-clockwise relative to first stimuli (s1). Rotations could range from 3°-12° or remain unchanged (0° rotation). C) Outcome probabilities on a given trial. Of the four possible choices, only one guaranteed positive feedback; selecting both the correct side *and* the correct direction of rotation. Perceptual errors resulted in negative feedback. On trials with zero rotation, one answer was arbitrarily assigned as correct and participants were given feedback accordingly. Side errors resulted in positive *or* negative feedback, at chance level. D) Example sequence of reward-contingent sides (solid grey line) and side chosen by participant (red and blue dots). Feedback is indicated by dot fill; unfilled dots indicate negative feedback at the end of that trial. The magnitude (but not direction) of the rotation was matched for left- and right-hand stimuli on each trial, to preclude a strategy in which participants monitored both stimulus pairs and responded to the one with the easier perceptual decision.

Figure 2. Lateralized and non-lateralized beta power grand averages. a) average beta band power averaged over all trials, without correction for hand of response. Box shows the time window used for the topo plot. Time windows were selected based on the time-frequency diagram and are therefore illustrative only. b) Average lateralized beta band power corrected for hand of response, at ipsi- and contralateral central sensors (C 1/2,3/4, and CP1/2,3/4). Boxes show the time-frequency windows used for the topo plots; time windows were selected based on the time-frequency diagram and are therefore illustrative only. Topo plots: topographies on right-hand response trials are left-right reversed to allow for averaging. C)

Time course of beta band (16-24Hz) activity, non-lateralized as in a, and lateralized as in b. Qualitatively, the non-lateralized beta dip appears to precede the lateralized effect.

For this visualization, power was calculated as decibel change from baseline at -2250:-1750 ms before the response

Figure 3: Effects of perceptual difficulty on behaviour. A) Change in % perceptually correct responses as a function of angle difference between stimulus 1 and 2. Dots represent individual participants. Note that larger angle changes equate to easier trials, i.e., the most difficult trials are plotted on the right side. B) As A, but change in mean reaction time. Dots represent individual datapoints. C) Given that a trial t was unrewarded, the probability of switching decreases with increasing perceptual difficulty on trial t .

Figure 4. Before correct switches, reaction time decreases but accuracy is maintained

a) Mean reaction time in the four trials before- and after a switch, error bars are SEM across participants. Grey line – raw mean reaction times. Black line – mean residual reaction times after the effect of task difficulty (angle of rotation), ISI and feedback valence on the previous trial have been removed with linear regression. B) as in A, but here the measure is proportion of trials perceptually correct rather than reaction time. Red line is the proportion of positive feedback – note that switching overwhelmingly follows negative feedback.

Figure 5. Response readiness is reflected in decreased phasic pupil response and pre-movement beta power, but has no effect on post-movement beta power or tonic pupil dilation.

a,c,e: Beta-band power, time-locked to the button-press response. Pre-movement beta power was lower when a) difficulty was low, c) the working memory interval was longer and e) the reaction time was faster than expected given difficulty and working memory interval. b,d,e: Pupil dilation, time-locked to the button-press response. Units are arbitrary. The phasic pupil response to the button-press (a transient dilation lasting around 2s after an event) is larger on trials with slower reaction times than expected, given difficulty and working memory interval. There was no effect of task difficulty or ISI on pupil size.

In panels c-f, tertile splits are shown for illustration only; the statistical analysis was a time-resolved GLM using continuous measures. In all panels, horizontal lines indicate time points

at which the effect of each explanatory was statistically significant (t-test of beta values from individual participant GLMs, at each time-point against zero (26 df; $\alpha = 0.05$, two tailed; Bonferroni-corrected for 81 timepoints (thick line) or uncorrected (thin line)).

Figure 6. Beta power is tonically reduced around the point of state transition

a) Following negative feedback (red), beta power immediately decreases. **b)** beta power remains tonically lowered throughout the trial following negative feedback. **Horizontal lines** indicate the periods in which there is a statistically significant reduction in beta power associated with negative feedback on current trial (top) or previous trial (bottom); $p < 0.05$ one-tailed uncorrected (fine lines) and Bonferroni corrected for 81 timepoints (heavy lines). **c)** Around a state reversal (side switch), the phasic beta response to feedback (**red markers**) is affected; there is a large decrease in post-feedback beta on the final trial before the side switch, and beta power in the post-feedback window returns to baseline gradually over several trials following the side switch. The reduction in beta power following the side-switch is sustained across the whole trial; the presented values (**black markers**) are the average beta power in the time window 1 to 0.5s before the movement on each trial. The windows used for the red and black trace are shaded in parts a,b and e,f Time points significantly different to the overall mean are marked with star ($p < 0.05$ corrected for 8 comparisons) **d)** The tonic reduction in beta power shown in part c mirrors the strength of evidence favouring a side switch, obtained from a Bayesian ideal observer model (see *Methods*) – y axis shows the probability that the chosen side is indeed the correct (deterministic) side on each trial. Note that the posterior probability indeed falls below 0.5 (i.e. favours the non-chosen side), only on the trial directly preceding the switch – however the belief that one is choosing the correct side starts to fall ahead of the switch and rises slowly over several trials after the switch. **e,f)** Illustrative plot showing the full beta motif on the four trials after a side switch (divided into trials with positive (e) and negative (f) feedback). Note the tonic reduction in beta is evident across the whole time window.

Figure 7. Pupil size is tonically increased around the point of state transition

a) Following negative feedback (red), pupil size immediately increases. **b)** pupil size remains tonically raised throughout the trial following negative feedback. **Horizontal lines** indicate the periods in which there is a statistically significant increase in pupil size associated with negative feedback on current trial (top) or previous trial (bottom); $p < 0.05$ one-tailed uncorrected (fine lines) and Bonferroni corrected for 81 timepoints (heavy

lines). **c)** Around a state reversal (side switch), the phasic beta response to feedback (**red markers**) is affected; there is a large increase in post-feedback pupil size on the final trial before the side switch, and pupil size in the post-feedback window returns to baseline gradually over several trials following the side switch. The increase in pupil size following the side-switch is sustained across the whole trial; the presented values (**black markers**) are the average pupil size in the time window 1 to 0.5s before the movement on each trial. The windows used for the red and black trace are shaded in parts a,b and e,f. Time points significantly different to the overall mean are marked with star (large star, $p < 0.05$ corrected for 8 comparisons; small star $p < 0.05$ uncorrected) **d)** The tonic increase in pupil size shown in part c mirrors the strength of evidence favouring a side switch, obtained from a Bayesian ideal observer model (see *Methods*) – y axis shows the probability that the non-chosen side is indeed the correct (deterministic) side on each trial. Note that the posterior probability indeed rises above 0.5 (i.e. favours the non-chosen side), only the trial directly preceding the switch – however the belief that one should choose the non-chosen side starts to rise ahead of the switch and falls slowly over several trials after the switch. Note that the values plotted here are simply the complement of figure 4d, shown in this orientation to facilitate comparison with the pupil-size effect **e,f)** Illustrative plot showing the full pupil dilation motif on the four trials after a side switch (divided into trials with positive (e) and negative (f) feedback). Note the tonic increase in pupil size is evident across the whole time window.

Figure 8. t scores from a 2D time-resolved regression in which beta power is used to predict pupil size. The marginal traces are the mean beta power and pupil size over all trials (similar to the traces in figures 6 and 7), shown for context. Saturated colours are significant at $p < 0.05$ two tailed, correcting for two multiple comparisons (with the plan to look at the two predefined pairs of time windows, as defined in the mediation analysis: EEG 1000-500ms pre response against pupil 0-2000ms post response (response-readiness analysis – orange boxes) and EEG 500-1000ms post feedback against pupil 2000-4000ms post feedback (feedback valence analysis – red boxes). a) This figure, shown for illustration of the full time-resolved relationship between pupil and beta power, indicates that the response-readiness effects are phasic in nature - pre-movement beta power is related only to the phasic pupil response (note lack of saturated colours elsewhere in the vertical orange box). In contrast, feedback related effects are sustained over time - post-feedback beta power has a relationship with pupil size throughout the trial, reflecting the tonic nature of this effect. b) Analysis as in panel a, but now showing the residual relationship between beta power and pupil size when all shared task effects (as in Eq 4 and 5) are first modelled out of the pupil

and EEG data; in other words the residuals from equation 4 and 5 are used as the pupil and EEG data in this figure. Note that, in contrast to panel (a), once shared task effects are removed, the response-readiness effects (orange boxes) disappear, indicating that these are driven by shared task factors. In contrast, the feedback-related effects (red box) remain, indicating a link between tonic beta power and tonic pupil size that is not explained simply by shared task influences.

JNeurosci Accepted Manuscript















