Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments

Peter R. Murphy, Niklas Wilming, Diana C. Hernandez-Bocanegra, Genis Prat-Ortega and Tobias H. Donner

Many decisions under uncertainty entail the temporal accumulation of evidence that informs about the state of the environment. When environments are subject to hidden changes in their state, maximizing accuracy and reward requires non-linear accumulation of evidence. How this adaptive, non-linear computation is realized in the brain is unknown. We analyzed human behavior and cortical population activity (measured with magnetoencephalography) recorded during visual evidence accumulation in a changing environment. Behavior and decision-related activity in cortical regions involved in action planning exhibited hallmarks of adaptive evidence accumulation, which could also be implemented by a recurrent cortical microcircuit. Decision dynamics in action-encoding parietal and frontal regions were mirrored in a frequency-specific modulation of the state of the visual cortex that depended on pupil-linked arousal and the expected probability of change. These findings link normative decision computations to recurrent cortical circuit dynamics and highlight the adaptive nature of decision-related feedback to the sensory cortex.

Many decisions under uncertainty entail the temporal accumulation of noisy information about the state of the world. When the source of this ‘evidence’ remains constant during decision formation, maximizing decision accuracy requires perfect, linear integration (that is, equal weight given to each evidence sample)\(^2\). Progress in computational theory, behavioral modeling, and neurophysiological analysis has translated these ideas into an influential framework wherein decision-relevant evidence is encoded by neurons in sensory cortical regions and then accumulates over time along sensory–motor cortical pathways\(^6\). This computation can account for the build-up of choice-predictive pre-decisional activity that has been observed in several brain regions\(^5\).

This framework has two important limitations. First, it was developed for an artificial context: environments in which the source (that is, stimulus category) generating evidence remains constant during decision formation. In this special case, the agent’s uncertainty originates only from noise corrupting the evidence. However, natural environments can undergo hidden changes in their state, constituting another source of uncertainty\(^7\). Perfect and linear evidence accumulation is suboptimal for changing environments\(^1\). Second, the linear, feedforward integration of evidence entailed in the above framework contrasts with the recurrent activity computations to recurrent cortical circuit dynamics and highlight the adaptive nature of decision-related feedback to the sensory cortex.

This framework has two important limitations. First, it was developed for an artificial context: environments in which the source (that is, stimulus category) generating evidence remains constant during decision formation. In this special case, the agent’s uncertainty originates only from noise corrupting the evidence. However, natural environments can undergo hidden changes in their state, constituting another source of uncertainty\(^7\). Perfect and linear evidence accumulation is suboptimal for changing environments\(^1\). Second, the linear, feedforward integration of evidence entailed in the above framework contrasts with the recurrent activity computations to recurrent cortical circuit dynamics and highlight the adaptive nature of decision-related feedback to the sensory cortex.

Results

To interrogate the mechanisms of adaptive perceptual evidence accumulation, we developed a task with hidden changes in the environment (that is, evidence source; Fig. 1a). The evidence samples were small checkerboard patches presented in a rapid sequence at different locations along a semicircle in the lower visual hemifield. Sample positions were generated from one of two noisy sources, the probability distributions shown in the upper row of Fig. 1a. The ‘active’ source was chosen randomly at the beginning of each trial and could change after each sample, with a probability (hazard rate \(H\)) of 0.08 in the main experiment. The participant’s task was to report which source was active at the end of each sequence via left- or right-handed button press.

Signatures of adaptive evidence accumulation. The normative model maximizing accuracy on this task (Fig. 1b) entails the accumulation of evidence samples that carry information (in the form of log-likelihood ratios, LLRs) about the two possible environmental states\(^9\), a process also known as ‘belief updating.’ This model serves as a benchmark against which the behavior and neural activity of human participants can be compared. The key difference between the normative model and previous accumulator models\(^1\) lies in dynamic shaping of evidence accumulation (the combination of prior belief \(\psi\) with new evidence LLR to yield posterior belief \(L\)) by

---

\(^{1}\)Section Computational Cognitive Neuroscience, Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. \(^{2}\)Trinity College Institute of Neuroscience and School of Psychology, Trinity College Dublin, Dublin, Ireland. \(^{3}\)Interactive Systems Group, Department of Computer Science and Applied Cognitive Science (INKO), University of Duisburg-Essen, Duisburg, Germany. \(^{4}\)Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain. \(^{5}\)Centre de Recerca Matematica (CRM), Campus UAB, Bellaterra, Barcelona, Spain. \(^{6}\)Bernstein Center for Computational Neuroscience, Charité Universitätsmedizin, Berlin, Germany. \(^{7}\)Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands. \(^{8}\)Amsterdam Brain and Cognition, University of Amsterdam, Amsterdam, the Netherlands. \(^{✉}\)e-mail: murphyp7@tcd.ie; t.donner@uke.de

---
the estimated rate of environmental state change \( (H) \). Specifically, the prior for the next sample \( (\psi_{n+1}) \) is determined by passing the updated belief \( (L_n) \) through a non-linear function that, depending on \( H \), can saturate (slope \( \approx 0 \)) for strong \( L_n \) and entail more moderate information loss \( (0 < \text{slope} < 1) \) for weak \( L_n \).}

In our setting, this non-linearity could be cast in terms of sensitivities of evidence accumulation to two quantities: (1) the uncertainty about the environmental state before encountering a new sample \( (\psi_{n+1}) \); and (2) the change-point probability \( (\text{CPP}) \). We investigated these sensitivities by performing a two-tailed permutation test on weights for the first six samples of the normative model for our task (Extended Data Fig. 1d), participants performed the task better than expected from two simple strategies (Fig. 2a, empty bars, and Extended Data Fig. 2): deciding based on the sum of all samples \( (73.9\% \pm 1.2\%, \text{mean} \pm \text{s.e.m.}; P < 0.0001, \text{two-tailed permutation test}; \text{Fig. 2a, magenta}) \) or on the final sample only \( (71.8\% \pm 0.6\%; P < 0.0001, \text{two-tailed permutation test}; \text{Fig. 2a, gray}) \). Participants’ performance \( (77.5\% \pm 1.9\%) \) came close to the benchmark performance of the ideal observer with perfect knowledge of \( H \) and the generative distributions and without noise or bias \( (81.8\% \pm 0.9\%; \text{mean and s.e.m. across participant-specific stimulus sequences}; \text{Fig. 2a, navy bar}) \).

Critically, our participants exhibited the same modulation of evidence accumulation by CPP and \( -|\psi| \) as the normative process depicted in Fig. 1. We quantified this modulation through logistic regression (model 2, Methods). The first set of regression weights, which quantified the time course of the leverage of LLR on choice \( \text{(also called psychophysical kernel)} \); increased over time \( (P < 0.0001, \text{two-tailed permutation test on weights for the first six versus the last six samples}; \text{Fig. 2b, left}) \). Thus, similar to the normative model for our task \( (\text{Extended Data Fig. 1d}) \), participants tended to discount early information in their final choices. The second and third sets of regression weights captured the modulation of evidence weighting by CPP and \( -|\psi| \). Each revealed strong positive modulations \( (P < 0.0001, \text{Fig. 2b, middle}; -|\psi|, P = 0.0002, \text{Fig. 2b, right}; \text{two-tailed cluster-based permutation tests}) \). The modulatory effect of CPP was also larger than that of \( -|\psi| \).
Fig. 2 | Signatures of adaptive evidence accumulation in human behavior and the cortical circuit model. a. Performance of n = 17 independent human participants (unfilled bars), an ideal observer (navy), the cortical circuit model (orange), perfect accumulation (magenta), and deciding based on the last sample only (gray). Cyan dots, accuracies of normative model fits. b. Time-resolved weight of evidence on choice (left) and its modulation by CPP (middle) and $-|\psi|$ (right). Black, mean and s.e.m. (error bars) of participants’ data; significance bars, time points for which weights differ from zero (P < 0.05, two-tailed cluster-based permutation test; evidence-weighting cluster, P < 0.0001; CPP clusters, P = 0.03 and P < 0.0001; $-|\psi|$ cluster, P = 0.0002). Orange, kernels produced by the circuit model, scaled to align with kernel magnitudes from participants. Cyan and magenta shadings, fits of normative and perfect accumulator (acc.) models, respectively. AU, arbitrary units. c. Schematic of the circuit model. $D_1$ (left choice, green) and $D_2$ (right choice, red), choice-selective populations of excitatory neurons; $I$, shared population of inhibitory interneurons. d. Response of choice-selective populations to an example trial. Top, LLRs; middle, raster plot of spiking activity from $D_1$ and $D_2$ neurons; bottom, population firing rates (sp s$^{-1}$, spikes per second). Arrows at the top and bottom highlight changes in environmental state and changes of mind, respectively.

In other words, both CPP and $-|\psi|$ ‘upweighted’ the impact of the associated evidence on choice. The same evidence-weighting signatures were produced by the normative accumulation scheme for a range of task statistics (Extended Data Fig. 1d). The signatures also replicated in a separate dataset at faster and slower sample presentation rates (Extended Data Fig. 3).

These behavioral signatures ruled out a range of alternative evidence accumulation schemes. Perfect linear accumulation (without bounds) produced flat psychophysical kernels (Fig. 2b, magenta), while perfect accumulation toward an absorbing bound produces stronger weights for early samples$^{15}$. Leaky accumulation with exponential decay of accumulated evidence$^{15}$ produces recency but not the CPP or $-|\psi|$ modulations of evidence weighting (Extended Data Fig. 4). In sum, the signatures in Fig. 2b suggested that participants’ behavior was produced by a computation that approximated the normative evidence accumulation process shown in Fig. 1b.

This conclusion was supported by fitting different models to participants’ choices (Fig. 2 and Extended Data Figs. 4 and 5). We fitted several variants of the normative model, with between three and 11 free parameters (see Methods for details). Here, we use the term ‘normative model’ as shorthand for the adaptive accumulation scheme from Fig. 1b, without implying a noise-free or unbiased process. All variants included a free subjective hazard rate parameter ($H$), a decision noise parameter and at least one parameter for assigning subjective weight (LLR) to the range of possible sample locations$^{15}$. Some variants also included up to seven additional parameters for a non-linear scaling of subjective LLR and/or a gain parameter controlling the evidence weighting depending on its consistency with current belief$^{13,20–23}$. Inclusion of these additional parameters was supported by model comparison (Extended Data Fig. 5 and Methods). The choices of the best-fitting model variant (‘normative fit’ in Fig. 2, cyan) were highly consistent with participants’ choices (88.8% ± 0.8%). The normative fit outperformed all considered versions of perfect and leaky accumulator models (mean $\Delta$Bayes information criteria (BIC) = 71.1 ± 7.6; higher BIC for all 17 participants; Extended Data Fig. 5).

One alternative to the normative model, perfect accumulation toward non-absorbing bounds, did produce the above evidence-weighting signatures (Extended Data Fig. 4) and only marginally worse fits to the data than the best-fitting version of the normative model ($\Delta$BIC = 11.1 ± 2.9; higher BIC for 16 of 17 participants; Fig. 2c). Indeed, the non-linearity imposed by the non-absorbing bounds can approximate the normative accumulation scheme for our task setting (low $H$ and high signal-to-noise ratio (SNR))$^{11}$. This indicates that saturation for strong belief states (and not leak for weak belief states) was critical in accounting for participants’ behavior.

In sum, our analyses indicate that, while limited in their performance by internal noise and biased internal representation of task variables, participants approximated the normative evidence accumulation process for the current setting: adaptive, non-linear accumulation, characterized by sensitivity during periods of uncertainty and plateaus in the evolving decision variable during periods of relative stability.

Implementation via recurrent cortical circuit dynamics. A prominent cortical circuit model for decision making$^{28}$ is made up of recurrently connected spiking neurons (Fig. 2c). The cortical model has previously been used to simulate single-neuron activity in regions implicated in decision making during standard perceptual choice tasks$^{16}$. This model precludes perfect accumulation, which limits its performance in standard perceptual choice tasks without change.
Motoneuronal activity tracks adaptive accumulation. The above behavioral analyses indicate that participants (and the circuit model) implemented an adaptive evidence accumulation scheme in line with normative principles. We next sought to identify signatures of the resulting decision variable in cortical population activity, measured via magnetoencephalography (MEG). Previous work on choice tasks without change points has identified signatures of evidence accumulation in neural signals reflecting the preparation of an action plan. To isolate this motor preparatory activity, we applied a linear (spectral and spatial) filter to the data from the main task (Methods and Fig. 3a, right). The filter was constructed using data from a separate delayed motor-response task in which participants prepared left- or right-handed button presses and which yielded sustained lateralization of alpha (8–14 Hz) and beta (16–30 Hz)-band power over motor cortices contralateral to the prepared response (Fig. 3a, left).

The selective motor preparatory activity built up gradually throughout the prolonged period of evidence accumulation (Fig. 3b). All analyses were pooled across correct and error trials because, unlike in tasks without change points, errors do not necessarily imply a failure of inference: even the ideal observer performing our task made errors on ~18% of trials, despite the absence of noise and biases and the inference process being exactly the same as for correct trials (Fig. 2a, navy).
The build-up of motor preparatory activity closely mirrored the trajectories of the decision variables from the normative and circuit models but not those from the perfect accumulator, which failed to rapidly respond to change points (Fig. 3c,d). Single-trial activity was significantly correlated with decision variable trajectories of both the normative fit and the circuit model across the group (Fig. 3d; normative, mean $r = -0.085$, $P < 0.0001$, two-tailed permutation test; circuit, mean $r = -0.081$, $P < 0.0001$, two-tailed permutation test) and individually within 16 of 17 participants (Fig. 3d, filled circles). Also, motor preparatory activity after each updating step (that is, after computation of $w_{n+1}$, the prior for the next sample) showed a sigmoidal dependence on $I_{ns}$ as predicted by both the normative and circuit models (Fig. 3e).

To quantify the adaptive nature of the dynamics of motor preparatory activity following each sample, we regressed it on a linear model composed of prior belief ($\psi$), new evidence (LLR$_n$) and the interactions of new evidence with CPP$_n$ and $-|\psi|$. (Methods, model 3). While the two interaction terms would be zero for linear (perfect or leaky) accumulation schemes (Fig. 3g, pink or gray bars), a signal reflecting the normative decision variable should have non-zero regression coefficients for all terms, with a stronger interaction term for CPP$_n$ than that for $-|\psi|$ (Fig. 3g, cyan).

These are our findings for the motor preparatory activity (Fig. 3g). The first three terms were encoded in motor preparatory activity in both beta and alpha bands ($\psi$, $P = 0.0006$; CPP$_n$, $P = 0.0002$; LLR$_n$CPP$_n$, $P = 0.0009$; two-tailed cluster-based permutation tests) with sustained encoding of the prior and delayed encoding of the evidence sample that was in turn modulated by CPP. The $-|\psi|$ interaction was weaker than the CPP$_n$ interaction, as expected (cyan bars for normative fits), showing only a trend in the expected direction ($P = 0.08$, cluster corrected; Fig. 3g). Formal comparison between versions of the above regression model containing only

---

**Fig. 4 | Frequency-specific encoding of evidence and decision variable along the visuo–motor cortical pathway.** a. Time–frequency maps of evidence (LLR) encoding across anatomical ROIs in visual, temporal, posterior parietal and motor–premotor cortices (ROI locations are plotted on the cortical surface in the top-left and top-right insets, with colors corresponding to those of ROI-specific time–frequency map titles). The color scale of time–frequency maps reflects group-level t-scores. Note that positive LLR encoding is present in the gamma band (40–65 Hz; black arrow for IPS0/1) in all visual field maps but absent in the aIPS, IPS/PCeS, PMd/v and M1, and negative LLR encoding is present in the alpha band (8–14 Hz; gray arrow for IPS0/1) in all ROIs. The aIPS lacks both the transient gamma-band response of posterior ROIs and the sustained beta-band (16–30 Hz) response of anterior ROIs and was omitted from further clustering. b. Time–frequency maps reflecting encoding of other constituent components of adaptive evidence accumulation (same as in Fig. 3g), averaged for compactness across three sets of ROIs (insets) encompassing occipital visual field maps (V1–V4), temporal and parietal visual field maps (LO1/2, MT+, VO1/2, PHC, V3A/B, IPS0–IPS3) and anterior intraparietal or motor cortices (IPS/PCeS, PMd/v, M1). Black contours, $P < 0.05$ (two-tailed cluster-based permutation test); gray contours, largest subthreshold clusters of the LLR−|$\psi$| effect ($P < 0.05$, two-tailed t-test).
that LLR was weakly positively encoded in transient low-frequency participants averaged over the highlighted clusters). We also note these considerations motivated a set of further analyses described in the following sections. These aimed to (1) elucidate the evolution of the decision variable across the sensory–motor pathway for our change-point task, from the primary visual cortex (V1) to the primary motor cortex (M1), and (2) quantify the arousal modulation of the decision computation across this pathway.

**Visual cortical encoding of evidence and decision variable.** We reconstructed the dynamics of selective activity lateralization in a set of regions of interest (ROIs; Supplementary Table 1) along the sensory–motor pathway (Methods): multiple visual field maps in occipital, parietal and temporal cortices as well as parietal and frontal regions encoding left- versus right-hand movements. In all ‘action-related’ cortical regions except the anterior intraparietal sulcus (aIPS), the lateralization of alpha- and beta-band activity exhibited sustained encoding of LLR (Fig. 4a) and prior belief ($\psi$), as well as modulation of LLR encoding by CPP (Fig. 4b). This is consistent with the sensor-level results from Fig. 3 and further supports the representation of the decision variable in the format of the associated action plan.

In the visual field maps, the encoding of computational variables exhibited different features. First, the lateralization of short-latency gamma-band (40–65 Hz) responses transiently encoded LLR (Fig. 4a) but none of the other variables (Fig. 4b), a profile consistent with pure evidence encoding. By contrast, alpha-band lateralization in the visual cortex mirrored the profile of alpha and beta lateralization in downstream action-related regions, in particular, the sustained encoding of prior belief ($\psi$) and an interaction between LLR and CPP (Fig. 4). Thus, alpha-band activity even in the early visual
cortex (V1–V4) was also consistent with encoding of the normative decision variable. Overall, alpha-band activity in the visual cortex was reminiscent of feedback of the evolving decision variable identified in earlier work on standard perceptual choice tasks (that is, without change points)\(^{25,26}\), as well as the idea that visual cortical alpha-band activity reflects cortical feedback signaling\(^{19,20,29,30}\). We further explored this idea in a series of analyses focusing on a hierarchically ordered set of dorsal visual field maps\(^{26}\) (V1, V2–V4, V3A/B, intraparietal sulcus (IPS)/0/1, IPS/2/3) as described in the next section. The areas of these and other clusters are defined in Supplementary Table 1.

**Signatures of adaptive decision feedback in the visual cortex.** We first delineated the feedforward processing of sensory evidence by means of multivariate decoding of the spatial patterns of evoked responses\(^{25,26}\) separately for each ROI (Fig. 5a–d). The LLR decoding latency of action-related regions (termed ‘motor’ in Fig. 4) lagged that of V1 (P = 0.005, two-tailed weighted permutation tests, Methods). The timescale of LLR decoding increased from V1 to extrastriate (P = 0.048) and motor areas (P = 0.007; extrastriate versus motor, P = 0.021, two-tailed weighted permutation tests, Fig. 5d). Within the visual cortical hierarchy, timescales increased monotonically (P = 0.017, two-tailed weighted permutation test on the slope of regression line across areas; Fig. 5c). The increase of evidence decoding timescales across the visual cortical hierarchy was mirrored by progressively slow intrinsic activity fluctuations during the pretrial interval (Extended Data Fig. 7), in line with studies on monkeys\(^{25}\). In sum, these results indicate feedforward processing of evidence from V1 to higher-tier visual regions to motor regions, in line with previous work\(^{25,26}\).

The feedforward scheme suggested by Fig. 5a–d and Extended Data Fig. 7 stood in sharp contrast to the signatures of LLR encoding in alpha-band lateralization (Fig. 5e–h). Alpha-band LLR encoding latency of V1 lagged that of extrastriate regions (P < 0.0001, two-tailed weighted permutation test) but not that of motor regions (P = 0.11; Fig. 5e), and the LLR encoding timescale was slower in V1 than in extrastriate regions (P = 0.002) but not motor regions (P = 0.28; Fig. 5g). Alpha-band LLR encoding timescales monotonically decreased across the visual cortical hierarchy (P = 0.0004, two-tailed weighted permutation test; Fig. 5h, compare with Fig. 5c). In sum, the dynamics (latencies and timescales) of LLR encoding in alpha-band lateralization were also consistent with a feedback signal.

Previous monkey\(^{17}\) and human\(^{27}\) physiology of perceptual choice tasks without change points inferred decision-related feedback from the covariation between intrinsic (stimulus-independent) fluctuations of neural activity and behavioral choice. Here, we computed these fluctuations as the residuals, over and above the activity explained by the components of the normative decision variable depicted in Fig. 3h (model 6, Methods). Fluctuations in the alpha- and beta-band lateralization of IPS/PcES, PMd/v and M1 were robustly predictive of choice toward the end of the trial (Fig. 6a, right) but not at the beginning of the trial (Fig. 6b, right). Choice-predictive fluctuation is expected for these downstream regions, because they encode the action plan that ultimately controls the behavioral choice, even when that choice deviates from the choice prescribed by the normative decision variable. Further, this choice-predictive, action-related activity is expected specifically for such late trial intervals because, due to possible changes of mind, decision states earlier in the trial could differ substantially from the choice-determining state toward the end of the trial (Fig. 3c).

Remarkably, the early visual cortex (V1 and weakly in V2–V4), but not higher-tier visual field maps, exhibited similarly late, choice-predictive fluctuations, specifically expressed in the lateralization of alpha-band activity (Fig. 6a). Again, there was no effect earlier in the trial (Fig. 6b), ruling out the possibility that this visual cortical alpha-band signal might reflect a biased baseline state before the start of evidence accumulation, due, for example, to slow attentional drift or choice history effects across trials\(^{7,12}\).

In sum, we found similar encoding of the evolving decision state in the alpha-band from the motor cortex and V1 and more prominent decision signatures when progressing backward along the visual cortical hierarchy, from the extrastriate cortex to V1. Because MEG source reconstruction is limited by signal leakage, a possible concern is that the alpha-band effects observed in the visual cortex reflect signal leakage from downstream motor regions. However, the absence, or even sign reversal (Fig. 4b and Fig. 6a, top middle), of beta-band effects in the visual cortex renders this
The opposite was true for the (positive) LLR encoding in the gamma band, as predicted by the normative accumulation scheme (Fig. 7d, black). Evidence-weighting time courses then switched sign from sample to sample, again as predicted by the normative evidence accumulation scheme (Fig. 7d, gray; compare with Fig. 2b and Extended Data Fig. 3). By contrast, in the high-volatility condition, evidence-weighting time courses then switched sign from sample to sample, again as predicted by the normative evidence accumulation scheme (Fig. 7d, black).

For low volatility, we also replicated the pattern of LLR encoding (positive for gamma ($P = 0.0028$ for V1, $P < 0.0001$ for extrastriate regions V2–IPS3), negative for alpha ($P < 0.0001$ for V1 and V2–IPS3; Fig. 7e) and of choice-predictive fluctuations in the lateralization of early visual cortical activity ($P = 0.036$, all two-tailed permutation tests; Fig. 7f). Critically, both signatures adapted to the statistical environment (Fig. 7e,f). The strength of the (negative) alpha-band LLR encoding was significantly reduced under high volatility, both in V1 ($P = 0.003$, two-tailed permutation test) and the extrastriate visual cortex (V2–IPS3 pooled, $P = 0.0002$; Fig. 7e). The opposite was true for the (positive) LLR encoding in the gamma band. Thus, the overall pattern of results provides strong support for the notion of feedback of the evolving decision variable to the early visual cortex during decision formation.

Other work has speculated that decision-related feedback to the sensory cortex may stabilize the evolving decision state\(^4\), akin to confirmation bias\(^1\). Such an active consolidation of the evolving decision through feedback may be beneficial in relatively stationary environments with protracted periods of stability, as was the case in our task ($H \ll 0.5$). But it would be disadvantageous in highly volatile environments that are likely to change. In another experiment, we assessed whether the feedback signatures in the early visual cortex adapt to changing levels of environmental volatility. We used the same task as in Fig. 1 but manipulated the volatility to be either similarly low, as before ($H = 0.1$), or high ($H = 0.9$) in different blocks (Fig. 7a and Methods).

All but one participant adapted their behavior to these different settings (Fig. 7c) and again performed in line with the normative evidence accumulation scheme (Fig. 7d and Extended Data Fig. 8b; the remaining participant based choices on only the last evidence sample). Please note that the contribution of $-|\psi|$ to normative accumulation was negligible for the generative settings of this experiment (Extended Data Fig. 8a). For the low-volatility condition, we again replicated the consistently positive evidence weighting on choice, stronger weighting of late evidence and upweighting of evidence associated with high CPP (Fig. 7d, gray; compare with Fig. 2b and Extended Data Fig. 8a). By contrast, in the high-volatility condition, evidence-weighting time courses then switched sign from sample to sample, again as predicted by the normative evidence accumulation scheme (Fig. 7d, black).
Fig. 8 | Arousal responses to perceived state changes upweight the associated evidence in brain and behavior. a, Average trial-related pupil response (gray, normalized via z-scoring and expressed in units of s.d. of z) and its first derivative (green, expressed as z per second). Dashed gray vertical lines represent sample onsets. b, Pupil derivatives for subsets of trials in which only single change points occurred at different latencies, relative to pupil derivatives on trials without any state change. Grayscale traces, relative CPP for the same trials from normative model fits (scaled and time shifted by the same factors across trial groups). Traces truncated from three samples before to one sample after the average time of state change. c. Encoding of CPP and uncertainty in pupil responses to individual samples (CPP, black; −|ψ|, gray). d, Multiplicative effect of intrinsic, single-trial pupil derivative fluctuations on weighting of evidence samples in choice. e, Across-participant correlation of pupil choice effect from d with a fitted model parameter reflecting additional weight ascribed to new evidence that was inconsistent with belief. Inset: time-resolved Pearson correlation coefficients spanning the significant temporal cluster from d, the highlighted time point was used to generate the scatterplot. f, Multiplicative effect of pupil response to individual samples on encoding of new evidence in the lateralized, regionally specific MEG signal. Colors reflect group-level t-scores. Contours, P < 0.05 (two-tailed cluster-based permutation test). Shaded areas in a–d indicate the s.e.m. Significance bars in c, d, P < 0.05 (two-tailed cluster-based permutation tests).

band, which was enhanced under high volatility (V1, P = 0.022; V2–IPS3, P = 0.040). Likewise, the pattern of choice-predictive fluctuations in V1 alpha-band activity also changed from low to high volatility (difference, P = 0.004), with a switch from negative to positive sign under high volatility (P = 0.011, two-tailed permutation test; Fig. 7f), possibly indicating an active destabilization of the evolving decision state in this environmental setting.

In sum, high environmental volatility seemed to boost the bottom–up sensory evidence encoding (gamma band) and reduce the feedback signaling in early visual cortical alpha-band activity, suggesting flexible adaptation of these signal features to the stability of the environment.

Phasic arousal shapes accumulation and visual cortex state. Cortical circuit dynamics are not only shaped by intracortical feedback signals but also by input from brainstem arousal systems.24,25 Transient (phasic) arousal signals, specifically from the noradrenergic line system, might dynamically adjust the impact of new evidence on belief updating, particularly when the evidence is surprising.36–38 In a final set of analyses of data from the main experiment, we used rapid, sample-evoked dilations of the pupil to assess the involvement of brainstem arousal.19,40 We quantified rapid pupil responses as the first derivative (that is, the rate of change) of the raw response to increase temporal precision (Fig. 8a) and specificity for noradrenaline release.41

Samples strongly indicative of a change in environmental state evoked strong pupil responses (Fig. 8b,c). We quantified this effect by fitting a linear model consisting of model-derived variables including the CPP and −|ψ| values associated with each sample to the corresponding pupil response (model 7, Methods). We found a robust positive contribution of CPP (P < 0.0001, two-tailed cluster-based permutation test) in addition to a weaker positive contribution of −|ψ| (P = 0.023). Thus, in our task, phasic arousal was recruited by the two computational quantities, CPP and −|ψ|, that modulate evidence weighting in the normative model.

Sample-to-sample fluctuations in the evoked arousal response predicted an ‘upweighting’ of the impact of the associated evidence sample on choice (Fig. 8d–f). Variations in pupil responses beyond those explained by the above computational variables exhibited a positive, multiplicative interaction with LLR in its impact on choice (Fig. 8d; P = 0.003, two-tailed cluster-based permutation test). The magnitude of this effect also correlated with a participant-specific gain parameter, estimated in our normative model fits, which quantified the weighting of belief-inconsistent evidence beyond the upweighting entailed in the normative model (Fig. 8e; peak r = 0.66, uncorrected P = 0.004, two-tailed).

The pupil-predicted upweighting of the behavioral impact of evidence (Fig. 8d) was accompanied by a corresponding modulation of LLR encoding across the visual cortical hierarchy (Fig. 8f). Here, we expanded our decomposition of cortical dynamics (Fig. 4) by a term that reflected the modulation of evidence encoding by the associated pupil response (model 8, Methods). This showed a transient modulatory effect of pupil responses on selective evidence encoding in the power lateralization of visual cortical areas (Fig. 8f). The selective pupil-linked effect was superimposed onto a more sustained, global (across hemispheres) effect on low-frequency power across the visuo–motor pathway (Extended Data Fig. 9). The selective effect was specifically expressed in alpha-band power lateralization (Fig. 8f) and not evident in downstream, action-related regions (all P > 0.3; Fig. 8f). Thus, phasic arousal
modulated evidence encoding in the visual cortical system in the specific frequency band that our previous analyses of the same neural data had linked to decision feedback.

Discussion

Computational work has developed abstract, normative solutions for decision-making problems or simulated the underlying synaptic interactions in cortical microcircuits. We reasoned that the possibility of change, a feature of natural environments, is critical for developing an understanding that bridges between these different levels of analysis. Previous studies of perceptual decisions in changing environments focused on behavior. Here, we identified the cortical representation of an approximately normative decision variable for changing environments. We found that (1) this decision variable is encoded in the build-up activity of cortical regions involved in action planning and (2) its adaptive computation naturally emerges from recurrent interactions in cortical microcircuits. This forges a tight link between existing circuit models and normative evidence accumulation.

In addition, we found that the state of the visual cortex, expressed in activity fluctuations in the alpha frequency band, is shaped by an adaptive, stabilizing feedback of the evolving decision variable as well as pupil-linked arousal. Here, we conceptualize ‘cortical state’ as cortical activity distinct from feedforward sensory responses. Our data are in line with the notion that the cortical state entails an interplay of selective cortical feedback and neuromodulatory signals. Thus, our findings from the visual cortex call for an extension of models for evidence accumulation by multiple processing stages incorporating feedback and neuromodulatory input. We propose that these features enable the brain to adapt evidence accumulation to a range of environmental statistics.

Our account starts from the realization that adaptive evidence accumulation in volatile environments typically entails the non-linear modulation of evidence weighting by CPP and uncertainty. The impact of these modulations varies across environments. It is negligible in contexts that preclude the formation of strong belief states: high noise and/or a rate of change that renders the environment unpredictable (\( H \approx 0.5 \)). Correspondingly, a linear, leaky accumulator model fits the behavior of rats (with strong internal noise) well in an auditory task similar to ours. Yet, these modulations make a substantial contribution to normative accumulation in a wide range of contexts (Extended Data Fig. 1).

We reiterate that CPP and uncertainty are not an explicit component of the normative model but served as a vehicle for identifying diagnostic behavioral and neural signatures of normative evidence accumulation (see Extended Data Fig. 10 for the relationship to commonly used measures of surprise). In line with other work, we found that both variables were encoded in pupil-linked arousal responses (and CPP more strongly so than other surprise metrics; Extended Data Fig. 10e). These observations further validate our dissection of the adaptive accumulation process.

Previous work on statistical learning showed that humans tune their learning rate as a function of surprise and uncertainty. The learning tasks from this work operated on timescales at least an order of magnitude slower than that for our task. Yet, our results highlight an analogy between the adjustment of learning rate by surprise, uncertainty or pupil-linked arousal and the modulation of perceptual evidence weighting. It will be interesting to assess whether this analogy originates from shared circuit mechanisms. The previous work on dynamic adjustments of learning focused on identifying neural correlates of the variables that control the regulation (that is, surprise or uncertainty). By contrast, our study illuminates how these variables modulate the selective cortical signals that encode the result of the computation, the belief state.

The computation of the decision variable entails recurrent network interactions that produce timing jitter. Consequently, neural signatures of the decision variable may not be precisely phase locked to the onset of evidence samples but rather manifest in slower variations of the amplitude of local field potential oscillations picked up by MEG. Further, cognitive variables are often encoded in specific frequency bands, sometimes with opposite sign between bands.

We therefore reasoned that our frequency-resolved encoding analysis should be well suited for detecting correlates of the decision variable. Our results corroborate this reasoning.

Our results also support the notion that attractor dynamics in frontal and parietal cortical circuits implement the formation and maintenance of decision states; a circuit model with attractor dynamics accurately reproduced the detailed characteristics of both behavior and build-up activity in the parietal and frontal regions involved in action planning. One signature of attractor dynamics during decision making is the sigmoidal relationships of cortical activity and model decision variables found in Fig. 3e and elsewhere. Here, we identified an adaptive function of these dynamics: balancing stable accumulation of the evidence and flexible response to change. Yet, these specific circuit dynamics will only be beneficial in a limited range of environments. Changes in environmental statistics may require a tuning (for example, through neuromodulatory input) or reconfiguration (for example, through cortical feedback) of the circuit properties.

Feedback signals to the sensory cortex from downstream cortical regions during perceptual evidence accumulation could convey predictions that are compared with incoming sensory data in an iterative fashion. Feedback may also mediate the impact of expectations inherited from previous trials. When originating from downstream accumulator circuits, feedback may also consolidate the evolving network state. We find that decision feedback emerges during, not before, evidence accumulation and depends on environmental volatility. These findings support the notion of an adaptive feedback mechanism stabilizing evolving belief states when needed. The particularly pronounced signatures of decision feedback in V1 (stronger than in any other visual cortical area) may point to a special role for the early visual cortex in perceptual evidence accumulation.

An influential account holds that surprise-related phasic responses of the brainstem noradrenergic system cause a shift toward more ‘bottom–up’ signaling relative to ‘top–down’ signaling and thus a greater impact of new evidence on the evolving belief. This is broadly consistent with our observation that pupil responses to evidence indicative of environmental state change predicted an upweighting of the impact of that sample on choice. However, if boosting bottom–up signaling, the associated modulations of cortical evidence encoding may have been expected in visual cortical gamma-band responses and should also affect action-related regions (that is, stronger evidence encoding in motor alpha- and beta-activity). Instead, we found pupil-dependent modulations of evidence encoding only in the visual cortex, not in downstream regions, and only in the alpha band, not in the gamma band. It is tempting to speculate that, in the current settings, an enhancement of feedback signals to the visual cortex by phasic arousal stabilizes the updated belief state induced by evidence indicative of change.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-021-00839-z.

Received: 9 February 2020; Accepted: 12 March 2021; Published online: 26 April 2021
References

1. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. Psychol. Rev. 113, 700–765 (2006).
2. Yang, T. & Shadlen, M. N. Probabilistic reasoning by neurons. Nature 447, 1075–1080 (2007).
3. Brody, C. D. & Hanks, T. D. Neural underpinnings of the evidence accumulator. Curr. Opin. Neurobiol. 37, 149–157 (2016).
4. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. Annu. Rev. Neurosci. 30, 535–574 (2007).

5. de Lange, F. P., Rahnev, D. A., Donner, T. H. & Lau, H. Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. J. Neurosci. 33, 1400–1410 (2013).
6. Donner, T. H., Siegel, M., Fries, P. & Engel, A. K. Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Curr. Biol. 19, 1581–1585 (2009).
7. Fischer, A. G., Nigbur, R., Klein, T. A., Danielmeier, C. & Ullsperger, M. Cortical beta power reflects decision dynamics and uncovers multiple facets of post-error adaptation. Nat. Commun. 9, 5038 (2018).
8. Hanks, T. D. & et al. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature 520, 220–223 (2015).
9. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503, 78–84 (2013).
10. Wyart, V., de Gardelle, V., Scholl, J. & Summerfield, C. Rhythmic fluctuations in evidence accumulation during decision making in the human brain. Neuron 76, 874–887 (2012).
11. O’Connell, R. G., Dockree, P. M. & Kelly, S. P. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nat. Neurosci. 15, 1729–1735 (2012).
12. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. Neuron 46, 681–692 (2005).
13. Glaze, C. M., Kable, J. W. & Gold, J. I. Normative evidence accumulation in unpredictable environments. eLife 4, e08825 (2015).
14. Ossmy, O. et al. The timescale of perceptual evidence integration can be adapted to the environment. Curr. Biol. 23, 981–986 (2013).
15. Wimmer, K. et al. Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. Nat. Commun. 6, 6177 (2015).
16. Wang, X. J. Probabilistic decision making by slow reverberation in cortical circuits. Neuron 56, 955–968 (2002).
17. Nienborg, H. & Cumming, B. G. Decision-related activity in sensory neurons reflects more than a neuron’s causal effect. Nature 459, 89–92 (2009).
18. Kiani, R., Hanks, T. D. & Shadlen, M. N. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. J. Neurosci. 28, 3017–3029 (2008).
19. Piet, A. T., El Hady, A. B. & Brody, C. D. Rats adopt the optimal timescale for evidence integration in a dynamic environment. Nat. Commun. 9, 4265 (2018).
20. Cheadle, S. et al. Adaptive gain control during human perceptual choice. Neuron 81, 1429–1441 (2014).
21. Talluri, B. C., Urai, A. E., Tsetsos, K., Usher, M. & Donner, T. H. Confirmation bias through selective overweighting of choice-consistent evidence. Curr. Biol. 28, 3128–3135 (2018).
22. Tsetsos, K., Chater, N. & Usher, M. Salience driven value integration explains decision biases and preference reversal. Proc. Natl Acad. Sci. USA 109, 9659–9664 (2012).
23. Weiss, A., Chambler, V. Lee, J. K., Drugowitsch, J. & Wyart, V. Interacting with volatile environments stabilizes hidden-state inference and its brain signatures. Nat. Commun. 12, 2228 (2021).
24. Murphy, P. R., Boonstra, E. & Nieuwenhuis, S. Global gain modulation generates time-dependent urgency during perceptual choice in humans. Nat. Commun. 7, 13526 (2016).
25. Murray, J. D. et al. A hierarchy of intrinsic timescales across primate cortex. Nature 17, 1661–1664 (2013).
26. Wang, X. J. Macroscopic gradients of synaptic excitation and inhibition in the neocortex. Nat. Rev. Neurosci. 21, 169–178 (2020).
27. Wilming, N., Murphy, P. R., Meynial, F. & Donner, T. H. Large-scale dynamics of perceptual decision information across human cortex. Nat. Commun. 11, 5109 (2020).
28. McGinley, M. J. et al. Waking state: rapid variations modulate neural and behavioral responses. Neuron 87, 1143–1161 (2015).
29. Bastos, A. M. et al. Visual areas exert feedforward and feedback influences through distinct frequency channels. Neuron 85, 390–401 (2015).
30. van Kerkoerle, T. et al. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. Proc. Natl Acad. Sci. USA 111, 14332–14341 (2014).
31. Felleman, D. J. & van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47 (1991).
32. Glaze, C. M., Kable, J. W. & Gold, J. I. Learning the value of information in an uncertain world. Nat. Neurosci. 15, 1040–1046 (2012).
33. de Gee, J. W. et al. Dynamic modulation of decision biases by brainstem arousal systems. eLife 6, e23323 (2017).
34. Joshi, S., Li, Y., Kalwani, R. M. & Gold, J. I. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 89, 221–234 (2016).
35. Reimer, J. et al. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat. Commun. 7, 13289 (2016).
36. Mobbs, D., Trimmer, P. C., Blumstein, D. T. & Dayan, P. Foraging for foundations in decision neuroscience: insights from ethology. Nat. Rev. Neurosci. 19, 419–427 (2018).
37. Marr, D. & Poggio, T. From Understanding Computation to Understanding Neural Circuitry. (Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 1976).
38. Harris, K. D. & Thiele, A. Cortical state and attention. Nat. Rev. Neurosci. 12, 509–523 (2011).
39. Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. Nat. Neurosci. 10, 1214–1221 (2007).
40. McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning rate in a dynamic environment. Neuron 84, 870–881 (2014).
41. Donner, T. H. & Siegel, M. A framework for local cortical oscillation patterns. Trends Cogn. Sci. 15, 191–199 (2011).
42. Amit, D. J. & Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb. Cortex 7, 237–252 (1997).
43. Inagaki, H. K., Fontolan, L., Romani, S. & Svoboda, K. Discrete attractor dynamics underlies persistent activity in the frontal cortex. Nature 566, 212–217 (2019).
44. Friston, K. J. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138 (2010).
Methods

We report data from three experiments using different variants of the behavioral task from Fig. 1: experiment 1 was the main behavioral and MEG experiment of this study (Figs. 1–6 and Extended Data Figs. 1, 2, 4–7, 9 and 10). Experiment 2 was a behavioral control experiment (Extended Data Fig. 3). Experiment 3 was a MEG experiment following up on experiment 1 (Fig. 7 and Extended Data Fig. 8).

Participants. The study was approved by the ethics committee of the Hamburg Medical Association and included a total of 51 human participants including the first author. All participants provided written informed consent, had normal or corrected-to-normal vision and no history of psychiatric or neurological diagnosis. Participants received remuneration in the form of an hourly rate, a performance-dependent bonus and a bonus for completing all planned sessions. For experiment 1, 17 participants (mean ± s.d. age of 28.22 ± 3.89 years, range 23–36; 11 females) completed one training session (120 min), three (16 participants) or four (the first author) sessions of the main experiment (150 min each) and one session to obtain a structural magnetic resonance image (MRI) (30 min).

For experiment 2, four participants (ages of 25, 26, 27 and 28 years; two males) completed one training session (120 min) and three (one participant) or four (three participants) experimental sessions (120 min each).

For experiment 3, thirty participants (age, 26.07 ± 3.49 years; range, 20–35 years; 18 females) completed one training session (120 min), three sessions of the main experiment (480 min each) and one session to obtain a structural MRI (30 min). This experiment also included a randomized, double-blind pharmacological intervention (cross-over design, 40 mg atomoxetine, 40 mg donepezil or placebo). Drug-related participant exclusion criteria were smoking, consumption of >15 units of alcohol per week, illegal drug use, pregnancy, medication intake, glaucoma, phaeochromocytoma, cardiovascular disease or known hypersensitivity to atomoxetine or donepezil. Here, we report results pooling data over drug conditions; drug effects will be reported in a separate paper.

Behavioral tasks. Stimuli were generated using Psychtoolbox-3 for MATLAB11. Visual stimuli were back projected on a transparent screen using a Sanyo PLC-XP15 projector at 60 Hz during MEG and on a VIEWPixx monitor in a psychophysics room. Participants were seated 61 cm from the screen during MEG or with their head in a chinrest 60 cm from the monitor in the psychophysics room (training; all of experiment 2). For experiments 1 and 2, the experimental sessions took place on consecutive days. They comprised between seven and nine blocks (76 trials each) with 2–10-min breaks between blocks. For experiment 3, sessions comprised 16 blocks (86 trials each), including 2–10-min breaks between blocks and a 45–60-min break between blocks 8 and 9. Each block was followed by feedback about the mean choice accuracy for that block and a running mean for that session. Participants completed 22–33 blocks (1,628–2,508 trials) in experiment 1, 33–47 blocks (2,100–3,174 trials) in experiment 2 and 42–51 blocks (3,612–4,386 trials) in experiment 3.

Main task. The main task across experiments was a two-alternative forced choice task in which the generative task state $S = \text{left}$ (right) could change unpredictably. Participants were asked to maintain fixation at a centrally presented mark throughout the trials, monitor a sequence of evidence samples and report their inference about $S$ at the end of the sequence.

Stimuli were presented against a gray background. Three placeholders were present throughout each trial: a light-gray vertical line extending downward from the fixation point; a colored half-ring in the lower visual hemifield (polar angle, from $-90$ to $90^\circ$; eccentricity, $8.8^\circ$), which depicted the LLR associated with each possible sample location; and a fixation mark as a black disc $0.18^\circ$ in diameter superimposed onto a disk $0.36^\circ$ in diameter with varying color informing participants about intervals. The colors comprising this half-ring and the fixation point were selected from the Teufel colors52. Evidence samples consisted of achromatic, flickering checkerboards (temporal frequency, 10 Hz; spatial frequency, $2^\circ$) within a circular aperture (diameter $=0.8^\circ$) and varied in polar angle (constant eccentricity of $8.1^\circ$).

In experiment 1, samples were presented for 300 ms (sample-onset asynchrony, SOA), 400 ms. Samples were centered on polar angles $\theta_n$, drawn from one of two truncated Gaussian distributions $p(\theta | \sigma_n)$ with variance $\sigma_n = 29^\circ$ and means symmetric about the vertical meridian ($\mu_n = -17^\circ$, $\mu_n = 17^\circ$). If a draw $\theta_n$ was $<-90^\circ$ ($>90^\circ$), it was replaced with $-30^\circ$ ($+30^\circ$). $S$ was chosen at random in each trial, and could change with a hazard rate $H = p(S = \text{right} | S_n = \text{left}) = p(S = \text{left} | S_n = \text{right}) = 0.08$ after each sample. Seventy-five percent of sequences contained 12 samples. The remaining 25% (randomly distributed in each block) contained (with equal probability) 2–11 samples and were introduced to encourage participants to attend to all samples. Each trial began with a variable fixation baseline period (uniform between 0.5–2.0 s) during which a stationary checkerboard patch was presented at a polar angle of 0°. This checkerboard then began to flicker, followed after 400 ms by the first evidence sample. The final sample of the sequence was then replaced by the stationary 0° patch, and a ‘Go’ cue instructed participants to report their choice 1.0–1.5 s (uniform) after the sequence end. They pressed a button with their left or right thumb to indicating left or right, respectively. Auditory feedback 0.1 s after the response informed participants whether their choice corresponded to the true $S$ at the sequence end: an ascending 350–950 Hz tone (0.25 s) for correct choice and a descending 950–350 Hz tone for incorrect choice. This was followed by an intertrial interval of 2 s when participants could blink. During the preparatory interval, sample sequence and subsequent delay, the color of the second fixation disk was light red, the ‘Go’ cue was this disk becoming light green, and the intertrial period was indicated by the disk becoming light blue.

Experiments 2 and 3 used the same task, with the following exceptions. In experiment 2, samples were presented either for 100 ms with 200 ms SOA or for 500 ms with 600 ms SOA. Presentation duration and SOA alternated between task blocks. In experiment 3, $H$ was pseudorandomly assigned to be 0.1 or 0.9 for each session, under the constraint that the starting $H$ be chosen equally often across the group for each session, and every participant would start with each $H$ at least once. It then remained fixed for eight task blocks and switched for the remaining eight blocks. The generative distributions in experiment 3 were $\mu_n = -28^\circ$, $\mu_n = +28^\circ$, $\sigma_n = \sigma_w = 28^\circ$. In experiment 3, 65% of sample sequences contained ten samples, the remaining 35% ranged between 2–9 samples (uniformly), the preparatory interval duration was 0.5–1.5 s, the time between sequence completion and the ‘Go’ cue was 0.7–1.2 s, and the intertrial interval was 1.2 s.

In all experiments, participants practiced the task in a psychophysics room a few days before the first session, with gradual increases in complexity and detailed information from the experimenter about task statistics.

Localizer task. In each MEG session, participants also completed one block of a localizer task for measuring motor preparatory activity for the hand movement used to report choice. Participants fixated the central fixation mark while a sequence of lexical cues and subsequent ‘Go’ cues for responding were presented. A trial began with presentation of one of two lexical cues (‘LEFT’ or ‘RIGHT’; 15-point white Trebuchet font; 0.3 s) 1.25° above fixation. The task was to execute a button press with the corresponding thumb after a 1-s delay-fixation interval, marked by a color change at fixation. The block was comprised of 60 trials (30 left and 30 right responses, randomly distributed) and administered after the sixth block of the decision-making task in each session.

Statistics and reproducibility. The sample size was chosen based on recent work using comparable methods (for example13). Before data collection, we verified through simulations that our combinations of sample size and trial counts were sufficient to produce key behavioral effects even for high noise.

No participant was excluded from analysis in experiments 1 and 2. In experiment 3, one participant had subjective $H$ values approaching 0.5, consistent with basting decisions only on the final evidence sample. Regression model fits for this participant did not converge. Consequently, they had to be excluded from all related analyses. We used standard criteria based on previous studies to exclude single trials in which MEG or eye-tracking data were contaminated by artifacts.

All experimental manipulations analyzed here were repeated within individual participants and administered without blindings or switches for the experimenters. All participants were allocated to the same group for analysis. Trial orders were fully randomized within and across participants using a random seed for each session.

Unless otherwise stated, permutation testing was used for statistical analysis (two-tailed; 10,000 permutations). When applied to time series or time–frequency maps, we implemented cluster level corrected for multiple comparisons (cluster-forming threshold, $P < 0.05$). In cases in which the reliability of an estimator depended on the strength of an underlying effect (latencies and timescales of LLR encoding and decoding traces; see below), we used weighted permutation tests. Each participant $p$ was assigned a weight, $w_p$, determined by effect strength (described in a case-specific manner below), constructed such that $\sum_p w_p = 1$. $w_p$ determined the contribution of each participant to the weighted mean over participants and the probability that they would be drawn, with replacement, in iterations of the permutation test.

Normative model. The normative model for the main task prescribes the following computation:

$$L_n = w_p + LLR_n$$

$$w_p = \text{logit} \left( \frac{1 + H}{H} \exp \left( -L_{n-1} \right) \right) - \log \left( \frac{1 + H}{H} \exp \left( -L_{n-1} \right) \right)$$

Here, $L_n$ was the observer belief after encountering the evidence sample $x_n$ expressed in log-posterior odds of the alternative task states; $LLR_n$ was the log-likelihood ratio reflecting the relative evidence for each alternative carried by $x_n$ ($LLR_n = \text{logit} (p(x_n | \text{right})/p(x_n | \text{left}))$); and $w_p$ was the prior expectation of the observer before checking $x_n$. We used this model to derive two computational quantities: $\text{CPP}$ and $-|\text{PP}|$. $\text{CPP}$ was the posterior probability that a change in generative task state has just occurred, given the expected $H$, the evidence carried by $x_n$ and the observer’s belief before encountering that sample $L_{n-1}$ (see Supplementary Note 1 for derivation). Uncertainty was defined as the negative
absolute value of the prior (−|ψ|), reflecting uncertainty about the generative state before observing x (Fig. 1c).

We evaluated the impact of CPP and −|ψ| on evidence accumulation as a function H and the SNR of the generative distributions (difference between distribution means divided by their s.d.). For each point on a 5×5 grid (H = (0.01, 0.03, 0.08, 0.20, 0.40), SNR = (0.4, 0.7, 1.2, 2.0, 5.0)), we simulated 10,000,000 observations, passed these through equations (1) and (2), calculated CPP and −|ψ| and then fitted the following model (model 1):

\[ \psi_{t+1} = \beta_1 + \beta_1 \times x_t + \beta_2 \times \text{LLR}_t + \beta_3 \times \text{LLR}_t \times \text{CPP}_t + \beta_4 \times \text{LLR}_t \times (-|\psi|) \]

where CPP was log transformed to reduce skew, \(\log(\text{CPP})\), and −|ψ|, were z scored, and all β are fitted regression coefficients. We assessed the contribution of each term to the updated prior \(\psi_{t+1}\), by calculating their coefficients of partial determination:

\[ R^2_{\text{partial}} = \frac{\text{SSR}_{\text{reduced}} - \text{SSR}_{\text{full}}}{\text{SSR}_{\text{reduced}}} \]

where SSR_reduced was the sum of squared residuals of equation (3), and SSR_full was the sum of squared residuals of a model excluding the term of interest. We repeated the analysis for \(H = (0.01, 0.09)\) used in experiment 3, here submitting CPP to the logit transform and for two other quantities defined in Extended Data Fig. 10: 'unconditional' and 'conditional' Shannon surprise.

Modeling of behavior. Normative model. We first computed the accuracy of participants' choices with respect to the true final generative state and compared this to the accuracy yielded by three idealized decision processes presented with identical stimulus sequences: normative accumulation (equations (1) and (2)), perfect accumulation and only using the final evidence sample. For each trial, choice \(r_t\) (left = −1, right = +1) was determined by the sign of the log posterior odds after observing all samples: \(r_t = \text{sign}(L_{n,t})\) for normative, \(r_{t-1} = \text{sign}(\sum_{t-1}^n \text{LLR}_n)\) for perfect and \(r_t = \text{sign}(\text{LLR}_t)\) for the last sample, where \(n\) indicated the number of samples presented on trial \(n\). Second, for each strategy and participant, we computed choice accuracy as a function of duration of the final state on each trial (from 1 on trials in which a change point occurred immediately before the final sample to 12 on full-length trials without change points). Third, we assessed the consistency between the choices of participants and each idealized strategy, computing the slope of a psychometric function relating log-posterior odds to human choice. For each strategy and participant, we normalized log-posterior odds across trials and described the probability \(\tau\) of a making a right choice as

\[ r_{t,\text{fitted}} = \tau + \frac{1 - \tau - \lambda}{1 + e^{-\lambda + \alpha x}} \]

where \(\tau\) and \(\lambda\) were lapse parameters, \(\alpha\) was a bias, \(D V_{t,\text{norm}}\) was the z-scored log posterior odds, and \(\alpha\) was the slope, reflecting consistency between the choices produced by a strategy and those of a participant. We estimated \(\gamma, \delta\) and \(\alpha\) by minimizing the negative-log likelihood of the data using the Nelder–Mead simplex search routine and assumed \(\gamma, \lambda\). We then fit variants of the normative model to participants’ behavior, assuming that choices were based on the log posterior odds \(L_{n,t}\) for the observed stimulus sequence on each trial. \(L_{n,t}\) was corrupted by a noise term \(\epsilon\), such that choice probability \(\tau\) was computed as follows.

\[ r_{t,\text{fitted}} = \frac{1}{2} \left( 1 + \frac{1}{2} \text{erf} \left( \frac{L_{n,t} + \epsilon}{\sqrt{2\epsilon}} \right) \right) \]

We also allowed for misestimation of \(H\), a bias in the mapping of stimulus location to LLR and a bias in the weighting of evidence samples that were (in)consistent with the existing belief. We fitted the parameters by minimizing the cross-entropy between participant and model choices:

\[ e = - \sum_{n=1}^{N_{\text{trial}}} (1 - r_{n,t}) \log(1 - r_{n,t}) + r_{n,t} \log(\hat{r}_{n,t}) \]

where \(r_{n,t}\) was the participant choice. The sum of e with any regularization penalty terms was minimized via particle swarm optimization (PSO toolbox”), setting wide bounds on all parameters and running 300 pseudorandomly initialized particles for 1,500 search iterations. The relative goodness of fit of different model variants was assessed by the BIC:

\[ \text{BIC} = 2e + k \log(n) \]

where \(k\) was the number of free parameters, and \(n\) was the number of trials.

The motivation for the inclusion of terms in our model fits that allowed for deviations from the ideal strategy was as follows.

For selection noise, a noise term was only applied to the final log posterior, corrupting the translation of final belief into choice rather than the preceding accumulation process. The latter is also likely corrupted by noise, as evident in our analyses of residual MEG and pupil fluctuations. We used selection noise for consistency with previous implementations and because this yielded decent fits. The exact nature of internal noise is beyond the scope of the present study.

For subjective \(H\), in line with previous work\(^1\), we allowed for subjectivity in \(H\). The model fits indicated that participants in experiment 1 tended to underestimate \(H\) (subjective \(H = 0.039 \pm 0.005 \) (s.e.m.); Extended Data Fig. 5b).

For non-linear stimulus-to-LLR mapping, we allowed for a non-linearity in the mapping (of polar angle onto subjective LLR (LLR)) as a non-parametric function fit to participants’ choices. LLR was expressed as a cubic–spline-interpolated function of polar angle \(x\), estimating values at \(x = (12.5, 25, 37.5, 50, 62.5, 75, 82.5, 900)\) and assuming symmetry around \(x = 0\) and \(\text{LLR}(x = 0) = 0\). We used Tikhonov regularization of the first derivative of the function by adding a penalty term to the objective function \(\sum_{i} (\Delta \text{LLR}_i)^2\), where \(i\) indexed the interval of \(x\) at which \(\text{LLR}\) was estimated, and \(\gamma = 1 - 2\) was determined ad hoc.

In (in)consistency bias, based on previous work\(^2\), we also allowed for larger weights for samples (in)consistent with the existing belief. A multiplicative gain factor \(g\) was selectively applied to such samples, such that \(\text{LLR}_{n,t}^\text{new} = \text{LLR}_{n,t}^\text{old} \times g\) for any sample \(n\) where sign (\(\text{LLR}_{n,t}\)) ≠ sign (\(\psi_n\)). Indeed, participants assigned higher-than-optimal weight to samples that were inconsistent with their existing belief (fitted \(g = 1.40 \pm 0.04\) (s.e.m.); Extended Data Fig. 5d).

The full model fits to the data from experiment 1 consisted of 11 free parameters (noise, subjective \(H\), eight stimulus-to-LLR mapping parameters and inconsistency gain). We compared these to fits of more constrained models lacking combinations of the above parameters (Extended Data Fig. 4a). Here, we reparameterized the stimulus-to-LLR mapping function for each sample \(n\) as follows:

\[ \text{LLR}_{n,t} = \text{LLR}_{n,t}^\text{old} + \delta \times \psi_n \]

which was more constrained than the interpolated function but could produce convex (\(x > 1\), concave (\(x < 1\) or linear (\(x = 1\)) mapping functions using only two free parameters, \(\delta\) and \(\psi\)).

We used the same general approach to fit model variants to the data from experiment 3, with the following exceptions: (1) subjective \(H\) could vary across objective \(H\) conditions, (2) we only allowed for linear scaling to be applied to true LLRs (\(\text{LLR}_n^\text{new} = \beta \times \text{LLR}_n\)) and fixed \(\beta\) across \(H\) conditions, (3) we assessed whether \(g\) was best defined relative to the sign of \(\psi_t\) or \(L_{n,t}\), by comparing models assuming either definition (the normative model prescribes a sign flip in the transformation of \(L_{n,t}\) into \(\psi_t\), for \(H = 0.9\) but not for \(H = 0.1\)), and (4) we fit model variants to assess whether noise or \(g\) might vary across \(H\) conditions.

We also fit a variant that estimated the \(L_{n,t}^\text{new} to \(\psi_t\) mapping directly from participants’ choices. It was estimated as an interpolated non-parametric function assuming symmetry around \(L_{n,t} = 0\) and \(\psi(L_{n,t} = 0) = 0\) and replaced subjective \(\psi_t\). \(\psi_t\) was estimated for values of \(L_{n,t}\) that were spread evenly between 1–10 in steps of 1. We applied Tikhonov regularization to the first derivative with \(\gamma = 1 + 2\).

Alternative models. We fit three other accumulator models to the data that lacked key characteristics of the normative process. The first was a perfect accumulator that linearly integrated all evidence samples without loss. The second was a leaky accumulator that substituted equation (2) with the following \(L\)-to-\(\psi\) mapping:

\[ \psi_{t+1} = \Delta \text{LLR}_t \times (1 - \lambda) \]

where \(\lambda\) was the leak. The third model employed perfect accumulation toward non-absorbing bounds, substituting equation (2) for

\[ \psi_{t+1} = \begin{cases} -A & L_{n,t-1} \leq -A \\ L_{n,t-1} - A < \psi_{t+1} < A \\ A \\ L_{n,t-1} \geq A \end{cases} \]

where \(A\) was the bound height. Fits of these models variously incorporated noise, non-linear LLR mapping and an inconsistency bias. The latter two models had the same number of free parameters as our fits of the normative model, while the former had one parameter less.

Psychophysical kernels. We quantified the time course of the impact of evidence on choice using logistic regression (model 2):
logit (\(p_r(n = \text{right})\)) = \(\beta_0 + \sum_{i=1}^{n} \beta_{ix} \times LLR_{ip} + \sum_{j=0}^{m} \beta_{jx} \times LLR_{jp} + (-|\psi| r_{D})\) (13)

where \(i\) and \(j\) indexed sample position within sequences of \(n\) samples (12 for experiments 1 and 2, ten for experiment 3), and \(LLR\) was the true LLR. For experiment 3, we excluded \(-|\psi|\) terms because they did not contribute to normative belief updating in this setting (Extended Data Fig. 8a). The dependent variable was participants’ choice (left = 0, right = 1) or model choice probability. The set \(\beta_i\) quantified the impact of evidence at each position on choice, and the sets \(\beta_j\) and \(\beta_k\) quantified modulations of evidence weighting by CPP and \(-|\psi|\). CPP was log transformed (experiments 1 and 2) or logit transformed (experiment 3, see above), and both variables were then \(z\) scored. Additionally, all regressors were \(z\) scored across the trial dimension.

Circuit model. Parameterization of the cortical circuit model was performed as described in ref. 16, except for the following. First, the stimulus input function from the original model produced excessive primacy in evidence weighting, inconsistent with our data. We used a threshold-linear function symmetric around 40 Hz and imposed no upper bound on the input:

\[\text{Input}_0 = \max(0, 40 - LLR \times c)\]

\[\text{Input}_o = \max(0, 40 + LLR \times c)\] (14)

where \(\text{Input}_0\) was the input to choose-selective excitatory populations \(D_i\) and \(D_o\), and \(c\) was multiplicatively scaled \(LLR\). We set \(c = 19\), yielding an input of \(-110\) Hz for the strongest evidence strength, higher than that in ref. 16. Second, the choice-selective recurrent connectivity from the original model \((w_r\text{, set to 1.7})\) produced strong and stable attractor states that prohibited changes of mind for all but the strongest evidence strengths. We set \(w_r = 1.68\), which allowed the model to better track change points in our task. Third, we set the integration time step \(dt = 0.2\) ms (0.02 ms in the original model), verifying through simulations that this did not significantly change model behavior or dynamics. The model was simulated for all 12-sample trials presented to participants in experiment 1, from 0.4 s before sequence onset to 0.4 s after sequence offset. Stimulus input to \(D_i\) and \(D_o\) was present from 0 to 0.4 s from sample onset and was zero elsewhere. The model decision variable \(X\) was defined as \(X = r_{D_o} - r_{D_i}\), with \(r_{D_o}\) and \(r_{D_i}\) as the population firing rates (50-ms window). Model choice was determined by sign(X) at the end of each trial. The updated decision variable given a new sample was \(X\) at 0.4 s after sample onset.

We estimated the model psychophysical kernels using CPP and \(-|\psi|\) from the ideal observer. Noise in the model was low relative to evidence SNR (we did not manipulate the model noise because we aimed to test whether the model could reproduce the qualitative features of the data). Thus, kernel magnitudes were large compared to the participant kernels. We applied a single multiplicative scaling factor (0.375, chosen manually) to the model kernels, which approximately matched kernel magnitudes with the participants while preserving relative differences between kernel types. We assessed different regimes of a reduction of the circuit model described by diffusion of a decision variable in a double-well potential10 in Extended Data Fig. 6.

Data acquisition. MEG data were acquired with 275 axial gradiometers (equipment and acquisition software, CTF Systems) in a magnetically shielded room (sampling rate, 1,200 Hz). Participants were asked to minimize movement. Head location was recorded in real-time using fiducial coils at the nasal bridge and each ear canal. A template head position was registered at the beginning of the first session, and the participant was guided back into that position before each task block. Ag–AgCl electrodes measured the ECG, EOG and EEG (from 10–20 locations Fz, Cz, Pz) with a nasion reference. Eye movements and pupil diameters were recorded at 1,000 Hz with an EyeLink 1000 Long Range Mount 10–20 locations Fz, Cz, Pz) with a nasion reference. Eye movements and pupil diameters were recorded at 1,000 Hz with an EyeLink 1000 Long Range Mount.

Cluster-based permutation testing at the group level was participants’ choice (left \(=\) 1, right \(=\) 0) indicated the time point relative to trial onset, \(f\) indicated frequency, \(c\) indicated the spatial filter, \(\beta\) indicated the behavioral response, and session \(s\) was the motor preparatory activity (motor filters). Any trial on which a response was executed before the ‘Go’ cue was excluded from analysis. This was the case on the majority of trials for five participants, and their data were not used. Data from the remaining 12 participants were split into sensors covering the left (ML+) in the CTF system) and right (MR+) cerebral hemispheres. For each of 131 matching left–right pairs, we calculated a lateralization index LI for each time point, frequency and trial as the difference in \(db\_LI\) between the left and right sensor. We then fit the following linear regression model:

\[LI_{f,s,trl} = \beta_0 + \beta_1 \times f_{s0} + \sum_{i=1}^{n-1} \beta_{i} \times LLR_{i} + \sum_{j=0}^{m} \beta_{j} \times LLR_{jp} + (-|\psi| r_{D}) \times session\_otr\] (15)

where \(f_{s0}\) indicated the time point relative to trial onset, \(f\) indicated frequency, \(c\) indicated the spatial filter, \(\beta\) indicated the behavioral response, and session \(s\) was the motor preparatory activity (motor filters). Any trial on which a response was executed before the ‘Go’ cue was excluded from analysis. This was the case on the majority of trials for five participants, and their data were not used. Data from the remaining 12 participants were split into sensors covering the left (ML+) in the CTF system) and right (MR+) cerebral hemispheres. For each of 131 matching left–right pairs, we calculated a lateralization index LI for each time point, frequency and trial as the difference in \(db\_LI\) between the left and right sensor. We then fit the following linear regression model:

\[w_f = \frac{\sum_{j=1}^{n-1} M_{f,j} \times M_{f,j}}{\sum_{j=1}^{n-1} M_{f,j} \times M_{f,j}}\] (16)

Likewise, integrating \(M\) over frequency yielded a spatial filter. Not integrating at all yielded a spatio–spectral filter.

Modeling motor preparatory activity. The motor filters were applied to LI data from the main task by computing the dot product between data and filter along the desired dimension(s). This yielded a motor preparatory signal (motor below). Application of the spatial filter yielded TFRs of motor preparatory activity, which were segmented from 0 to 1 s relative to the onset of each of the 12 samples per full-length trial. The resulting time \(\times\) frequency \(\times\) sample \(\times\) trial matrix was used to assess the sensitivity of motor preparatory activity to computational variables. For simplicity, we derived computational variables from model fits in which the stimulus-to-LLR mapping was constrained to be linear (such that \(LLR_\beta = \beta \times LLR\)). Because \(\beta\) set the slope of the mapping function and was a free parameter. Although this model-based variant yielded marginally worse goodness of fit than the full model, the computational variables generated by each were highly correlated (average \(R^2\) 0.95 for \(w, L, CPP\) and \(-|\psi|\)). We then fit the following linear model (model 5):

\[\text{motor}_{f,s,trl} = \beta_0 + \beta_1 \times LLR_{ip} \times |\psi| r_{D} + \beta_2 \times LLR_{ip} \times CPP_{ip} + \beta_3 \times LLR_{ip} \times CPP_{ip} + \beta_4 \times LLR_{ip} \times session\_otr\] (17)
 Articles

Nature Neuroscience

We also constructed models containing complementary parts of the 'full' model from equation (17) (models 4 and 5).

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(18)

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(19)

For each model, we computed a 'super-BIC' score quantifying the model's group-level goodness of fit as a function of \( t \) and \( s \). This metric was calculated with the negative-log-likelihood, number of free parameters and number of observations determined by the sum of these values across participants. We averaged across the sample dimension and subtracted BIC scores for the 'belief model' (equation (18)) from those for the 'evidence model' (equation (19)) to generate a measure of the model's goodness of fit.

We characterized the relationship between motor_{\text{full}} and model-derived belief state by applying the spatio–spectral filter to the main task data, yielding a scalar measure of motor preparation per trial onset. We averaged from 0.4–0.6 s and z-scored across trials separately for each sample position \( i \) and session. We normalized the posterior log odds \( L_i \), to an analogous fashion to remove any differences in the range of \( L \) across participant-specific model fits. For each sample position \( i \) and participant, we sorted the signal by normalized \( L \), into 11 equal-sized bins and calculated the mean of both metrics per bin. Averaging across sample positions and participants yielded a group average belief encoding function, which we rescaled to the range \([-1, 1]\). We repeated the procedure, replacing the motor preparation signal with \( \psi_i \), from the same model 6 (excluding the L-to-W mapping for the normative model participant normalization, binning and averaging). We then repeated the procedure using the updated decision variable of the circuit model.

Single-trial correlations between motor preparatory activity and decision variables were quantified for each full-length trial by extracting scalar values of the relevant variables for each sample position \( s \) (motor, averaged from 0.4–0.6 s after sample onset) and by computing the Pearson correlation over sample positions between motor and model variables.

Source analyses. We used linearly constrained minimum variance (LCMV) beamforming to estimate activity time courses at the level of cortical sources\(^6\). We constructed individual three-layer head models from structural MRI scans using FieldTrip\(^6\). Four participants lacked individual MRIs, in which case we used the FreeSurfer average participant. Head models were aligned to the MEG data by a session-specific transformation matrix generated with MNE\(^6\). We reconstructed cortical surfaces from MRIs using FreeSurfer\(^6\) and aligned anatomical atlases (see below) to each surface. We then used MNE to compute LCMV filters confined to the cortical surface (4,096 vertices per hemisphere, recursively subdivided octahedron) from a covariance matrix computed on the time points from the cleaned and segmented data. The covariance matrix was computed across trials, from trial onset to 6.2 s (experiment 1) or 5.4 s later (experiment 3). We chose the source orientation with maximum output source power at each cortical location. The LCMV filters were used to project single trial time series as well as their complex-valued spectrograms into source space. We computed TFRs of power at cleaned and segmented data. The covariance matrix was computed across trials, timescale of the normalized trajectories, \( \tau \), and then converting the complex Fourier coefficients into power. We then averaged from trial onset to 6.2 s (experiment 1) or 5.4 s later (experiment 3). We chose the epoch duration for signal leakage\(^6\) (due to limited filter resolution or volume conduction).

We characterized the relationship between motor_{\text{full}} and model-derived belief state by applying the spatio–spectral filter to the main task data, yielding a scalar measure of motor preparation per trial onset. We averaged from 0.4–0.6 s and z-scored across trials separately for each sample position \( i \) and session. We normalized the posterior log odds \( L_i \), to an analogous fashion to remove any differences in the range of \( L \) across participant-specific model fits. For each sample position \( i \) and participant, we sorted the signal by normalized \( L \), into 11 equal-sized bins and calculated the mean of both metrics per bin. Averaging across sample positions and participants yielded a group average belief encoding function, which we rescaled to the range \([-1, 1]\). We repeated the procedure, replacing the motor preparation signal with \( \psi_i \), from the same model 6 (excluding the L-to-W mapping for the normative model participant normalization, binning and averaging). We then repeated the procedure using the updated decision variable of the circuit model.

Single-trial correlations between motor preparatory activity and decision variables were quantified for each full-length trial by extracting scalar values of the relevant variables for each sample position \( s \) (motor, averaged from 0.4–0.6 s after sample onset) and by computing the Pearson correlation over sample positions between motor and model variables.

Source analyses. We used linearly constrained minimum variance (LCMV) beamforming to estimate activity time courses at the level of cortical sources\(^6\). We constructed individual three-layer head models from structural MRI scans using FieldTrip\(^6\). Four participants lacked individual MRIs, in which case we used the FreeSurfer average participant. Head models were aligned to the MEG data by a session-specific transformation matrix generated with MNE\(^6\). We reconstructed cortical surfaces from MRIs using FreeSurfer\(^6\) and aligned anatomical atlases (see below) to each surface. We then used MNE to compute LCMV filters confined to the cortical surface (4,096 vertices per hemisphere, recursively subdivided octahedron) from a covariance matrix computed on the time points from the cleaned and segmented data. The covariance matrix was computed across trials, from trial onset to 6.2 s (experiment 1) or 5.4 s later (experiment 3). We chose the source orientation with maximum output source power at each cortical location. The LCMV filters were used to project single trial time series as well as their complex-valued spectrograms into source space. We computed TFRs of power at cleaned and segmented data. The covariance matrix was computed across trials, timescale of the normalized trajectories, \( \tau \), and then converting the complex Fourier coefficients into power. We then averaged from trial onset to 6.2 s (experiment 1) or 5.4 s later (experiment 3). We chose the epoch duration for signal leakage\(^6\) (due to limited filter resolution or volume conduction).

We characterized the relationship between motor_{\text{full}} and model-derived belief state by applying the spatio–spectral filter to the main task data, yielding a scalar measure of motor preparation per trial onset. We averaged from 0.4–0.6 s and z-scored across trials separately for each sample position \( i \) and session. We normalized the posterior log odds \( L_i \), to an analogous fashion to remove any differences in the range of \( L \) across participant-specific model fits. For each sample position \( i \) and participant, we sorted the signal by normalized \( L \), into 11 equal-sized bins and calculated the mean of both metrics per bin. Averaging across sample positions and participants yielded a group average belief encoding function, which we rescaled to the range \([-1, 1]\). We repeated the procedure, replacing the motor preparation signal with \( \psi_i \), from the same model 6 (excluding the L-to-W mapping for the normative model participant normalization, binning and averaging). We then repeated the procedure using the updated decision variable of the circuit model.

Single-trial correlations between motor preparatory activity and decision variables were quantified for each full-length trial by extracting scalar values of the relevant variables for each sample position \( s \) (motor, averaged from 0.4–0.6 s after sample onset) and by computing the Pearson correlation over sample positions between motor and model variables.

Source analyses. We used linearly constrained minimum variance (LCMV) beamforming to estimate activity time courses at the level of cortical sources\(^6\). We constructed individual three-layer head models from structural MRI scans using FieldTrip\(^6\). Four participants lacked individual MRIs, in which case we used the FreeSurfer average participant. Head models were aligned to the MEG data by a session-specific transformation matrix generated with MNE\(^6\). We reconstructed cortical surfaces from MRIs using FreeSurfer\(^6\) and aligned anatomical atlases (see below) to each surface. We then used MNE to compute LCMV filters confined to the cortical surface (4,096 vertices per hemisphere, recursively subdivided octahedron) from a covariance matrix computed on the time points from the cleaned and segmented data. The covariance matrix was computed across trials, from trial onset to 6.2 s (experiment 1) or 5.4 s later (experiment 3). We chose the source orientation with maximum output source power at each cortical location. The LCMV filters were used to project single trial time series as well as their complex-valued spectrograms into source space. We computed TFRs of power at each vertex by first aligning the polarity of time series at neighboring vertices (the beamformer output potentially included arbitrary sign flips for different vertices) and then converting the complex Fourier coefficients into power. We then averaged estimated power values across vertices within each ROI and baseline corrected using the DB transform (0.2–0.4 s before trial onset). We computed source-level lateralization indices \( \text{LI}_{\text{baseline}} \) by subtracting the power estimate for each left hemisphere ROI from that for the corresponding right hemisphere ROI.

We analyzed the following fMRI-defined ROI: (1) visual cortical field maps\(^6\), (2) regions exhibiting hand movement–specific lateralization (aIPS, IPS/PCeS and the M1 hand area\(^6\)) and (3) the dorsal–ventral premotor cortex\(^6\). Visual cortical field maps with shared foveal representation were grouped together\(^6\) (Supplementary Table 1), thus increasing the spatial distance between ROI centers and minimizing the impact of signal leakage\(^6\) (due to limited filter resolution or volume conduction).

We repeated model 3 fitting for each ROI, replacing motor_{\text{full}} with source-level lateralization \( \text{CPP} \). We extracted fitted \( \beta \)-coefficients and averaged across sample positions and 8–14 Hz for the following set: V1, V2–V4, V3A/B, IPS0/1, IPS2/3, IPS/PCeS, PMd/v and M1; reversed the sign of the cortical surfaces from MRIs using FreeSurfer\(^6\) and aligned anatomical atlases

\[ \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(19)

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(18)

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(19)

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(18)

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(19)

\[ \text{motor}_{\text{full}} = \beta_0 + \beta_1 \times \text{LLR}_{\text{full}} + \sum_{j=1}^{n-2} \beta_j \times \text{session}_{\text{full}} \]

(18)
where \( t \) indicates time point, \( x_{\text{hor}} \) and \( y_{\text{ver}} \) are horizontal and vertical gaze positions, and 'base' is the 'baseline' pupil diameter (−0.05 to 0.05 s around sample onset). LLR captured a possible relationship with 'unconditional' Shannon surprise (Extended Data Fig. 10). Previous sample CPP, −1ψ and [LLR] were included because the pupil response is slow, meaning correlations with variables from the previous sample may have caused spurious effects. We assessed the relevance of variability in the pupil response for choice by fitting versions of equation (21) in which the final term (\( \beta_i \)) was LLR-\( \text{resid}_{\text{trl}} \), where 'resid' referred to the residuals from equation (22).

The relationship between 'pupil' and neural evidence encoding was assessed by extracting residuals from equation (22) at \( t = 0.57 \) s and fitting the following equation to the ROI-specific LI signal (model 8).

\[
\text{LI}_{i,t,\text{trl}} = \beta_0 + \beta_1 \times \psi_{i,t,\text{trl}} + \beta_2 \times \text{LLR}_{i,t,\text{trl}} + \beta_3 \times \text{LLR}_{i,t,\text{trl}} \times \text{CPP}_{\text{trl}} + \beta_4 \\
\times \text{LLR}_{i,t,\text{trl}} \times (-|\psi_{i,t,\text{trl}}|) + \beta_5 \times \text{LLR}_{i,t,\text{trl}} \times \text{resid}_{\text{trl}} + \sum_{i=1}^{w} \beta_{6_i} \times \text{session}_{i,\text{trl}}
\]

This model was an extension of equation (17), where \( \beta_i \) captured modulation of LLR encoding in LI by the associated pupil response.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

Raw behavioral and eye-tracking data for experiments 1 and 2 are available at [https://doi.org/10.6084/m9.figshare.14039353](https://doi.org/10.6084/m9.figshare.14039353). Raw MEG data for experiment 1 are available at [https://doi.org/10.6084/m9.figshare.14039384](https://doi.org/10.6084/m9.figshare.14039384).

Source-reconstructed MEG data for experiment 1 are available at [https://doi.org/10.6084/m9.figshare.14039465](https://doi.org/10.6084/m9.figshare.14039465). Behavioral, eye-tracking and MEG data for experiment 3 will be made available following publication of a follow-up manuscript currently in preparation that will cover other research questions that this dataset was collected to address.

**Code availability**

All analysis code is available at [https://github.com/DonnerLab/2021_Murphy_Adaptive-Circuit-Dynamics-Across-Human-Cortex](https://github.com/DonnerLab/2021_Murphy_Adaptive-Circuit-Dynamics-Across-Human-Cortex).

**References**

51. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* 10, 433–446 (1997).
52. Teufel, H. J. & Wehrhahn, C. Evidence for the contribution of S cones to the detection of flicker brightness and red-green. *J. Opt. Soc. Am. A* 17, 994–1006 (2000).
53. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190 (2007).
54. Birge, B. in *Proceedings of the IEEE Swarm Intelligence Symposium* [https://ieeexplore.ieee.org/document/1202265](https://ieeexplore.ieee.org/document/1202265) (IEEE Swarm Intelligence Symposium, 2003).
55. Drugowitsch, J., Wyart, V., Devauchelle, A. D. & Koechlin, E. Computational precision of mental inference as critical source of human choice suboptimality. *Neurosci.* 92, 1398–1411 (2016).
56. Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J. Neurosci.* 30, 12366–12378 (2010).
57. Prat-Ortega, G., Wimmer, K., Roxin, A. & de la Rocha, J. Flexible categorization in perceptual decision making. *Nat. Commun.* 12, 1283 (2021).
58. Oostenveld, R., Pries, P., Maris, E. & Schoffelen, J. M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 156869 (2011).
59. Gramfort, A. et al. MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* 7, 267 (2013).
60. Van Veen, B. D., van Drongelen, W., Yuchtman, M. & Suzuki, A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880 (1997).
61. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194 (1999).
62. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207 (1999).
63. Wang, L., Mruczek, R. E., Arcaro, M. J. & Kastner, S. Probabilistic maps of visual topography in human cortex. *Cereb. Cortex* 25, 3911–3931 (2015).
64. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178 (2016).
65. Wandell, B. A., Dumoulin, S. O. & Brewer, A. A. Visual field maps in human cortex. *Neuron* 56, 366–383 (2007).
66. Donoghue, T. et al. Parameterizing neural power spectra into periodic and aperiodic components. *Nat. Neurosci.* 23, 1655–1665 (2020).
67. Pedregosa, F. et al. Scikit-learn: machine learning for Python. *J. Mach. Learn. Res.* 12, 2825–2830 (2011).

**Acknowledgements**

We thank K. Wimmer for discussion and advice on the cortical circuit model and F. Meyniel for comments on the manuscript. This work was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation, DO 1240/3-1, DO 1240/4-1 and SFB 936 [Projekt-Nr. A7]).

**Author contributions**

F.R.M., conceptualization, methodology, investigation, software, formal analysis, visualization, writing (original draft, review and editing); N.W., conceptualization, methodology, software, writing (review and editing); D.C.H.-B., investigation, formal analysis; G.P.O., software, formal analysis, writing (review and editing); T.H.D., conceptualization, methodology, resources, writing (original draft, review and editing), supervision.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Extended data is available for this paper at [https://doi.org/10.1038/s41593-021-00839-z](https://doi.org/10.1038/s41593-021-00839-z).

**Supplementary information**

The online version contains supplementary material available at [https://doi.org/10.1038/s41593-021-00839-z](https://doi.org/10.1038/s41593-021-00839-z).

**Correspondence and requests for materials**

should be addressed to F.R.M. or T.H.D.

**Peer review information**

*Nature Neuroscience* thanks Laurence Hunt, Valentin Wyart, and the other anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information**

is available at [www.nature.com/reprints](https://www.nature.com/reprints).
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | Sensitivity of normative evidence accumulation to change-point probability and uncertainty across a range of generative task statistics. a, Non-linearity in normative model for different hazard rates (H): Posterior belief after accumulating the most recent sample ($L_n$) is converted into prior belief for next sample ($\psi_{n+1}$) through a non-linear transform, which saturates (slope $\approx 0$) for strong $L_n$ and entails more moderate information loss (0<<slope<1) for weak $L_n$. In a static environment (H=0), the model combines new evidence (log-likelihood ratio for sample $n$, LLR$_n$) with prior ($\psi_n$) into updated belief without information loss (that is $L_n=\psi_n+\text{LLR}_n$, $\psi_{n+1}=L_n$). In an unpredictable environment (H=0.5), the new belief is determined solely by the most recent evidence sample (that is $L_n=\text{LLR}_n$, $\psi_{n+1}=0$) so that no evidence accumulation takes place. b, Example trajectories of the model decision variable ($L$) as function of the same evidence stream, but for different levels of H. c, Upper panels of each grid segment: Change point-triggered dynamics of change-point probability (CPP) and uncertainty ($-|\psi|$) derived from normative model as function of H (grid rows) and evidence signal-to-noise ratio (SNR, difference in generative distribution means over their standard deviation; grid columns). Lower panels of each grid point: Contribution of computational variables (including CPP and $-|\psi|$) to normative belief updating, expressed as coefficients of partial determination for terms of a linear model predicting the updated prior belief for a forthcoming sample. Yellow background (center of grid): generative statistics of the task used in Experiment 1; reproduced from main Fig. 1e for comparison. Data for each parameter combination were derived from a simulated sequence of 10 million observations. d, Psychophysical kernels of the normative model for the task used in experiment 1 (12 samples per trial, generative SNR $\approx 1.2$) but for each level of H from panel c. Kernels reflect the time-resolved weight of evidence on final choice (left) and the modulation of this evidence weighting by sample-wise change-point probability (CPP, middle) and uncertainty ($-|\psi|$, right). Kernels were produced by adding a moderate amount of decision noise ($\nu=0.7$, see Methods) to the ideal observer (that is the normative model with perfect knowledge of generative statistics); without noise, coefficients at high H (where choices are based almost entirely on the final evidence sample) are not identifiable.
Extended Data Fig. 2 | Consistency of human choices with those of idealized choice strategies, and dependence of choice accuracy on final state duration. **a**, Choices for \( n = 17 \) independent participants as function of log-posterior odds (z-scored) derived from each of three alternative strategies: basing choices only on final evidence sample (gray); perfect evidence accumulation (magenta); and an ideal observer with perfect knowledge of the generative task statistics and employing the normative accumulation process for the task (blue). Points and error bars show observed data ± s.e.m.; lines and shaded areas show mean ± s.e.m. of fits of sigmoid functions to the data. Slopes of fitted sigmoids were steepest for ideal observer (ideal vs. perfect accumulation: \( t_{16} = 5.9, p < 0.0001 \); ideal vs. last-sample: \( t_{16} = 9.6, p < 10^{-7} \); two-tailed paired t-tests), indicating that human choices were most consistent with those of the ideal observer. **b**, Choice accuracy for the same \( n = 17 \) participants as a function of duration of the final environmental state for the human participants (black), idealized strategies from panel a, fits of the normative model (cyan), and the circuit model (orange). Participants’ choice accuracy increased with the number of samples presented after the final state change on each trial, consistent with temporal accumulation. Error bars and shaded regions indicate s.e.m.
Extended Data Fig. 3 | Effect of sample onset asynchrony (SOA) on diagnostic signatures of adaptive evidence accumulation. Psychophysical kernels reflecting the time-resolved weight of evidence on final choice (left), and the modulation of this evidence weighting by sample-wise change-point probability (CPP, middle) and uncertainty (−ψ, right), separately for the two conditions of experiment 2 (n = 4 independent participants) in which participants performed the decision-making task at fast (0.2 s) and slow (0.6 s) sample onset asynchronies. Thin unsaturated lines are individual participants, thick saturated lines are means across participants. Cluster-based permutation tests (10,000 permutations) of SOA effects on each kernel type revealed no significant effects.
Extended Data Fig. 4 | Fit to human behavior of alternative evidence accumulation schemes. Alternative accumulation schemes are leaky accumulation (gray), where the evolving decision variable is linearly discounted by a fixed leak term after every updating step; and perfect accumulation toward non-absorbing bounds (green), which imposes upper and lower limits on the evolving decision variable without terminating the decision process, thus enabling changes of mind even after the bound has been reached. Both schemes combined can approximate the normative accumulation process across generative settings (ref. 13). All panels based on data from \( n = 17 \) independent participants. a, Choice accuracies for human participants (grey bars) and both model fits. b, Choice accuracy as a function of duration of the final environmental state, for the participants and the models. c, Regression coefficients reflecting the subjective weight of evidence associated with binned stimulus locations, for the participants and the models. d, Psychophysical kernels for the data and the model reflecting the time-resolved weight of evidence on final choice (left), and the modulation of this evidence weighting by sample-wise change-point probability (CPP, middle) and uncertainty (\(-|\psi|\), right). In panels b–d, error bars and shaded regions indicate s.e.m. Fits of the model versions used to generate behavior in all panels included free parameters for both a non-linear stimulus-to-LLR mapping function and a gain factor on inconsistent samples (see Extended Data Fig. 5). Even with these additional degrees of freedom, the leaky accumulator model failed to reproduce the CPP and \(-|\psi|\) modulations characteristic of human behavior. By contrast, like the normative model, perfect accumulation to non-absorbing bounds captured all qualitative features identified in the behavioral data. This is in line with the insight that this form of accumulation closely approximates the normative model in settings like ours, where strong belief states are formed often (that is low noise and/or low \( H \); ref. 13). This form of accumulation also captures a key feature of the dynamics of the circuit model that we interrogate in the main text; that is a saturating decision variable in response to consecutive consistent samples of evidence.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | Model comparison and qualitative signatures indicate approximation of normative belief updating by measured human behavior. All panels based on data from n = 17 independent participants. a, Bayes Information Criteria (BIC) for all models fit to the human choice data, relative to model with lowest group-mean BIC. Lines, group mean; dots, individual participants. Model constraints are specified by the tick labels. Colors denote unbounded perfect accumulator (magenta), leaky accumulator (gray), perfect accumulator with non-absorbing bounds (green), and the normative accumulation process with a subjective hazard rate (cyan). Labels refer to the following: ‘Linear’ (‘Nonlinear’) =linear (non-linear) scaling of stimulus-to-LLR mapping function; ‘Gain on inconsist.’ = multiplicative gain term applied to inconsistent samples; ‘Leak’ = accumulation leak; ‘Bound’ = height of non-absorbing bounds; ‘H’ = subjective hazard rate. ‘Perfect’ is a special case of the leaky accumulator where leak = 0. All models included a noise term applied to the final log-posterior odds per trial. See Methods for additional model details. Labels in black text highlight models plotted in main Fig. 2 and Extended Data Fig. 4. Model with lowest group-mean BIC (‘full normative fit’ in remaining panels) employs normative accumulation with subjectivity in hazard rate, a gain term on inconsistent samples, and non-linear stimulus-to-LLR mapping. b, Subjective hazard rates from full normative fits (cyan), with true hazard rate in blue. Participants underestimated the volatility in the environment (t16 = −7.7, p < 10⁻⁶, two-tailed one-sample t-test of subjective H against true H). c, Non-linearity in evidence accumulation estimated directly from choice data (black) and from full normative model fits (cyan). Non-linearity of the ideal observer shown in blue. Shaded areas, s.e.m. d, Multiplicative gain factors applied to evidence samples that were inconsistent with the existing belief state, estimated from full normative fits. Gain factor > 1 reflects relative up-weighting of inconsistent samples beyond that prescribed by normative model; gain factor < 1 reflects relative down-weighting of inconsistent samples. The ideal observer employing the normative accumulation process uses gain factor = 1 (dashed blue line). Participants assigned higher weight to inconsistent samples than the ideal observer (t16 = 8.2, p < 10⁻⁶, two-tailed one-sample t-test of fitted weights against 1). e, Regression coefficients reflecting modulation of evidence weighting by change-point probability (CPP). Shown are coefficients for the human participants (black line), full normative fits (cyan), normative model fits without gain factor applied to inconsistent evidence (grey), and the ideal observer with matched noise (blue). Error bars and shaded regions, s.e.m. f, Mapping of stimulus location (polar angle; x-axis) onto evidence strength (LLR; y-axis) across the full range of stimulus locations. Blue line reflects the true mapping given the task generative statistics, used by the ideal observer. Cyan line and shaded areas show mean ± s.e.m of subjective mappings used by the human participants, estimated as an interpolated non-parametric function (Methods). g, Regression coefficients reflecting subjective weight of evidence associated with binned stimulus locations, for the human participants (black line), full normative model fits including the non-linear LLR mapping shown in panel f (cyan), normative model fits allowing only a linear LLR mapping (grey), and the ideal observer with matched noise (blue). Error bars and shaded regions, s.e.m.
Extended Data Fig. 6 | Assessment of boundary conditions at which circuit model approximates normative evidence accumulation. A reduction of the spiking-neuron circuit model (ref. 57) was used to explore the impact of different dynamical regimes on behavioral signatures of evidence weighting. Left two columns: shape of model’s energy landscape (‘potential’ $\phi$) described by equations:

$$\frac{dX}{dt} = -\frac{d\phi}{dX} + \sigma \xi_t,$$

$$\phi(X) = -k\mu t X - aX^2 + bX^4$$

in response to ambiguous (LLR = 0, left) or strong evidence (LLR = mean + 1 s.d. of LLRs across experiment = 1.35, second from left). $\mu_t$ was the differential stimulus input to the choice-selective populations at time point $t$ relative to trial onset (in our case the per-sample LLR, which changed every 0.4 s) that was linearly scaled by parameter $k$ (fixed at 2.2); $\xi_t$ was a zero-mean, unit-variance Gaussian noise term that was linearly scaled by parameter $\sigma$ (fixed at 0.8); and $a$ and $b$ shape the potential. Middle-to-right columns: psychophysical kernels and modulations of evidence weighting by CPP and $-|\psi|$.

a, Double-well potential with small barrier between wells ($a = 2$, $b = 1$), corresponding to weak bi-stable attractor dynamics. This model variant featured a saturating decision variable in the face of multiple consistent evidence samples (akin to non-absorbing bounds), wells that maintained commitment states but were sufficiently shallow for changes-of-mind to occur in response to strongly inconsistent evidence, and sensitivity to new input when at the ‘saddle point’ between wells (that is during periods of uncertainty). As a consequence, it produced all key qualitative features of normative evidence accumulation (recency, strong modulation by CPP; weak modulation by uncertainty).

b, Single-well potential with no barrier ($a = 0$, $b = 1$). This model variant also featured a saturating decision variable, but no stable states of commitment. Thus, it exhibited stronger recency in evidence weighting than the one in panel a.

c, Flat potential with no wells ($a = 0$, $b = 0$), corresponding to perfect evidence accumulation without bounds. This model lacked both a saturating decision variable and stable states. As a consequence, it weighed all evidence samples equally (flat psychophysical kernel) and did not produce clear modulations by CPP or uncertainty.

d, Double-well potential with high barrier between wells ($a = 5$, $b = 1$), corresponding to strong bi-stable attractor dynamics. In this model variant, the wells were too deep for changes-of-mind to occur for all but the most extreme inconsistent evidence. Thus, it produced extreme primacy in evidence weighting. The evidence weighting signatures of the regimes in panels c, d were qualitatively inconsistent with the signatures of participants’ behavior (compare Fig. 2b). All three signatures were qualitatively re-produced by the regimes in both a and b, with the closest approximation to human behavior provided by the weak bi-stable attractor regime in panel a (highlighted in yellow).
**Extended Data Fig. 7 | Intrinsic fluctuation dynamics across visual cortical hierarchy.** a. Power spectra and associated model fits of intrinsic fluctuations across cortex (n=17 independent participants). For each hemisphere, ROI, and trial, we computed the power spectrum (1-120 Hz) from a 1-second ‘baseline’ period preceding trial onset and averaged these spectra across trials and hemispheres. Shaded areas show mean ± s.e.m of observed power spectra. We modeled the power spectra as a linear superposition of an aperiodic component (power law scaling) and a variable number of periodic components (band-limited peaks) using the FOOOF toolbox (ref. 66; default constraints, without so-called ‘knees’ which were absent in the measured spectra, presumably due to the short intervals; 49–51 Hz and 99–101 Hz excluded from fits due to contamination by line noise). The fitted aperiodic components only are overlaid as dashed lines. b. Power law scaling exponents of the aperiodic components estimated from fits in a, providing a measure of the relative contributions of fast vs. slow fluctuations to the measured power spectrum. Lines, group mean; dots, individual participants. ‘Forward’ indicates direction of hierarchy inferred from fitted exponents across the dorsal set of visual field map ROIs (V1, V2-V4, V3A/B, IPS0/1, IPS2/3); P-value derived from two-tailed permutation test on slope of line fitted to power law exponents ordered by position along visual cortical hierarchy.
Extended Data Fig. 8 | Normative accumulation, model fits and motor choice signal for experiment 3. a, Contribution of computational variables (including CPP and $-|\psi|$) to normative belief updating in generative contexts from experiment 3, expressed as coefficients of partial determination (upper) and regression coefficients (lower) for terms of a linear model predicting the updated prior belief for a forthcoming sample ($\psi_{n+1}$). Derived from application of the normative model to a single sequence of 10 million simulated samples, generated using the same statistics as the task in experiment 3 ($H = \{0.1, 0.9\}; SNR = 2$). Note that the strength of the contributions of each term is the same across the two levels of $H$, which is expected when the chosen levels of $H$ are symmetric around 0.5 as here. However, the sign of the contributions of prior ($\psi_n$) and new evidence (LLR) are reversed under high $H$, accounting for the flipping of belief sign imposed by the normative non-linearity (Fig. 7b, main text). Note also the minimal contribution of $-|\psi|$ in this generative setting, which prompted us to not consider $-|\psi|$ modulatory effects in analyses of the data from human participants. b, Bayes Information Criteria (BIC) for all models fit to the human choice data from experiment 3 (n = 30 independent participants), relative to model with the lowest BIC score averaged over participants. Lines, group mean; dots, individual participants. Model constraints are specified by the tick labels. All models included at least one subjective hazard rate, linear scaling parameter applied to the stimulus-to-LLR function, and a decision noise term applied to the final log-posterior odds per trial. Models could also include a gain parameter applied to evidence that is inconsistent with either the posterior after the last sample ($L_{n-1}$) or the prior for the next sample ($\psi_n$). We fit models with a variety of different parameter constraints over $H$ conditions, as specified via the tick labels. See Methods for additional model details. Label in black text highlights model used for remaining analyses. This model included a gain on inconsistent evidence relative to $\psi_n$, and allowed $H$ and noise to vary across $H$ conditions. This model produced BIC scores that were marginally, but not significantly, higher than a less parsimonious model in which the gain parameter was also free to vary across $H$ conditions. c, Weight of residual fluctuations in the MEG lateralization signal from M1 on final choice, over and above the weight on choice exerted by key model variables (n = 29 independent participants; 1 participant for whom regression models did not converge excluded). Effects shown are averaged over final 3 evidence samples (positions 8–10) in the trial, and averaged over $H$ conditions. Colors reflect group-level t-scores. Black contours, $p < 0.05$, two-tailed cluster-based permutation test; gray contours, largest sub-threshold clusters ($p < 0.05$, two-tailed t-test). The highlighted frequency band encompassing the majority of the choice-predictive effect in M1 (12–17 Hz) was used to identify choice signals in early visual cortex (main text Fig. 7f).
Extended Data Fig. 9 | Effect of sample-wise pupil response (derivative) on source-localized, hemisphere-averaged MEG signal. Regions of interest (ROIs) are plotted on cortical surface in top left and top right insets, with colors corresponding to those of ROI-specific time-frequency map titles. Color scale reflects group-level t-scores derived from subject-specific regression coefficients (\( n = 17 \) independent participants). Contours, \( p < 0.05 \) (two-tailed cluster-based permutation test). Large pupil responses were associated with a strong decrease in low-frequency (<8 Hz) power across all ROIs, a decrease in alpha- and beta-band (8-35 Hz) power predominantly in intraparietal and (pre)motor regions, and a weak increase in high-frequency (65-100 Hz) power in a restricted set of ROIs (uncorrected two-tailed permutation tests on regression coefficients averaged over time-points: VO1/2, \( p = 0.0004 \); PHC, \( p = 0.025 \); IPS0/1, \( p = 0.049 \); IPS2/3, \( p = 0.010 \); aIPS, \( p = 0.003 \)).
Extended Data Fig. 10 | Change-point probability modulates normative belief updating and phasic arousal responses more strongly than alternative metrics of surprise. **a,** Alternative surprise metrics as a function of posterior belief after previous sample ($L_{n-1}$) and evidence provided by the current sample ($LLR_n$). (**a1**) CPP (see main text). (**a2**) ‘Unconditional’ surprise, $U$, defined as Shannon surprise (negative log probability) associated with new sample location (‘loc$_n$’), given only knowledge of the generative distributions associated with each environmental state $S = \{l,r\}$. In our task, $U$ varied monotonically with absolute evidence strength ($|LLR|$) and was uncorrelated with log(CPP) (Spearman’s $\rho = 0.00$). (**a4**) ‘Conditional’ surprise $Z$: Shannon surprise associated with new sample, conditioned on both knowledge of the generative distributions and one’s current belief about the environmental state. $Z$ was moderately correlated with log(CPP) in our design (Spearman’s $\rho = 0.36$). Unlike CPP, which we use solely to decompose normative evidence accumulation and relate it to neurophysiological signals, in some models this form of surprise serves as the objective function to be minimized by the inference algorithm (for example ref. 50). **b,** Analogue of Dayan & Yu (2006) for 2AFC with time-varying prior, DY. **c,** Of the two $Z$ components, only CPP has strong, positive effect on normative belief updating. **d,** As panel c, but for expanded linear model that included modulations of evidence weighting by CPP and $U$, which combine linearly into $Z$. Only CPP yielded a robust, positive effect on evidence weighting. **e,** Encoding of surprise metrics in pupil response (temporal derivative) to evidence samples ($n = 17$ independent participants). Encoding time course for $Z$ was computed by a regression model that included $X$- and $Y$-gaze positions as nuisance variables. Encoding time courses for CPP and $U$ were derived from application of the normative model to a single sequence of 10 million simulated samples, generated using the same statistics as the task in Fig. 1a (main text). For each surprise metric, we fit a linear model to updated prior belief (same form as model 1 in Methods; inset: $X$, surprise metric). Left: coefficients of partial determination; right: contributions expressed as regression coefficients. CPP exerted the strongest and only positive modulatory effect on evidence weighting, consistent with the observed effects on behavior, cortical dynamics, and pupil (see main text).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a
- Confirmed
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

MEG, eye-tracking and pupillometry data were collected using the acquisition software developed by the system manufacturer (MEG: CTF Systems Inc., version 5.4.2; Eye-tracking/pupil: SR Research, version 4.594). Behavioral data were collected using Matlab version 2016a, using stimulus presentation functions from Psychtoolbox 3.

Data analysis

Matlab version 2015b was used for the majority of analyses, including custom code (behavioral analysis/modelling; MEG analysis; pupillometry analysis), the FieldTrip toolbox version 20160221 (MEG preprocessing, including the infomax algorithm for independent component analysis), the Particle Swarm Optimization toolbox version 1.0.0.0 (model fitting) and the FOOOF toolbox version 1.0.0 (analysis of MEG power spectra). Eye-blinks and saccades in eye-tracking data were detected online by the measurement device (Eyelink 1000) using software from the system manufacturer (SR Research version 4.594). Python 3.6 in combination with FreeSurfer version dev5-20161028 and the MNE toolbox version 0.16.2 were used for MRI-informed source localization of MEG data. Decoding analyses were carried out using Scikit Learn version 0.23.1. All custom code is publicly available at https://github.com/DonnerLab/2021_Murphy_Adaptive-Circuit-Dynamics-Across-Human-Cortex.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw behavioral and eye-tracking data for Experiments 1 and 2 are available at https://doi.org/10.6084/m9.figshare.14035935. Raw MEG data for Experiment 1 are available at https://doi.org/10.6084/m9.figshare.14039384. Source-reconstructed MEG data for Experiment 1 are available at https://doi.org/10.6084/m9.figshare.14170432. Source data for figures are available at https://doi.org/10.6084/m9.figshare.14039465. Behavioral, eye-tracking and MEG data for Experiment 3 will be made available following publication of a follow-up manuscript currently in preparation that will cover other research questions that this dataset was collected to address.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

**Sample size**  
No statistical methods were used to pre-determine sample size but our sample size is equal to or larger than those reported in previous publications for comparable designs (e.g. Siegel et al., 2008, Neuron; Donner et al., 2009, Curr. Biol.; Wilming et al., 2020, Nat. Commun.). Of particular note, each of the 17 subjects from our first experiment attended at least 4 sessions (1 training, 3 MEG) and we collected a median of 1976 analyzable trials per participant (range=1628–2508), which is unusually large relative to existing studies and facilitates meaningful data analysis at the level of single subjects (e.g. Figure 2a of main manuscript). The same was true for our second (n=4) and third (n=30) experiments. Prior to data collection, we also verified through simulations of the normative model for our task that these combinations of sample size and trial counts were more than sufficient to produce the critical behavioral effects (modulation of evidence weighting by change-point probability and uncertainty) even for high levels of internal noise.

**Data exclusions**  
No subjects were excluded from any analysis in Experiments 1 and 2. In Experiment 3, 1 out of 30 subjects was excluded because they used a heuristic strategy for solving our task that rendered certain analyses impossible (fits of regression models for this participant did not converge). This issue was not anticipated and so this exclusion criterion was not specified in advance. At the single subject level, we used pre-established criteria based on previous studies from our lab to exclude single trials in which MEG or eye-tracking data were contaminated by well-known artifacts (e.g. by blinks or saccades during the stimulus sequence).

**Replication**  
In addition to our main dataset (n=17), we collected and analyzed two independent datasets (n=4 and n=30) using the same basic task but with manipulations of either stimulus presentation rates (Experiment 2) or environmental volatility (Experiment 3). All attempts at replication of Experiment 1 effects in these follow-up datasets were successful. Specifically, we observed that key behavioral and model-based effects (recency in evidence weighting, as well as modulations of evidence weighting by change-point probability and, where appropriate, uncertainty) were present in all datasets. In Experiment 3, we also successfully replicated key MEG effects observed in Experiment 1 (encoding of computational variables, and presence of choice-predictive activity fluctuations, in visual cortex).

**Randomization**  
For analyses of the data from each experiment that we report here, subjects were all allocated into a single group for analysis. As such, randomization of subjects to experimental groups was not necessary.

**Blinding**  
Within each experiment and for the analyses that we report here, all subjects were allocated to the same single group. For this reason, blinding of experimenters was not necessary.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.
Human research participants

Policy information about studies involving human research participants

Population characteristics

| Experiment 1: Seventeen participants; mean ± s.d. age of 28.22 ± 3.89 years, range 23–36; 11 females. |
| Experiment 2: Four participants; ages of 25, 26, 27 and 28 years; 2 males. |
| Experiment 3: Thirty participants; mean ± s.d. age of 26.07 ± 3.49 years, range 20–35; 18 females. |

All had normal or corrected-to-normal vision and no history of psychiatric or neurological diagnosis.

Recruitment

| Experiment 1: 15 of 17 participants were recruited from a large database maintained by the Department of Neurophysiology and Pathophysiology at the University Medical Center Hamburg-Eppendorf. The remaining two participants were members of our laboratory - one was naive to the study aims, the other was the first author of the manuscript. |
| Experiment 2: 3 of 4 participants were recruited from the departmental database, one was a member of our laboratory who was naive to the study aims. |
| Experiment 3: 27 of 30 participants were recruited from the departmental database, three were members of our laboratory. Of the three laboratory members, two were already familiar with the nature of the experimental task, but none had performed it at the generative settings of Experiment 3 before participating. |

It is conceivable that familiarity with the aims of the research project may have predisposed those members of our laboratory that participated, including the first author, to intentionally try to perform the task more like the normative model - a form of self-selection bias. However, we verified that the behavior produced by these participants was not qualitatively different from those of the remaining sample; indeed, almost all participants clearly displayed the key behavioral signatures suggestive of an approximation of the normative strategy (recency in evidence weighting, as well as positive modulations of evidence weighting by change-point probability and, where appropriate, uncertainty). Thus, we do not think that inclusion of members of our laboratory as participants created any significant bias in the reported results. A separate source of bias that we acknowledge here is that the participant database we used for the majority of the recruitment contains mostly students who studied medicine at the University Medical Center Hamburg-Eppendorf, thereby not forming a representative sample from the general population. However, our study investigates low-level decision processes which we believe to be largely unaffected by this possible selection bias. All participants, with the exception of the first author and laboratory members, received remuneration for their participation in the form of an hourly rate, a study completion bonus and an additional performance-dependent bonus.

Ethics oversight

| The experiments were approved by the ethics committee of the Hamburg Medical Association and conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.