Dissociable effects of surprise and model update in parietal and anterior cingulate cortex

Jill X. O'Reilly[a,1,2], Urs Schüffelgen[b,1], Steven F. Cuell[b], Timothy E. J. Behrens[a,c], Rogier B. Mars[a,b], and Matthew F. S. Rushworth[a,b]

[a]Centre for Functional MRI of the Brain (FMRIB), John Radcliffe Hospital, Oxford University, Oxford OX3 9DU, United Kingdom; [b]Department of Experimental Psychology, Oxford University, Oxford OX1 3UD, United Kingdom; and [c]Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom

Brains use predictive models to facilitate the processing of expected stimuli or planned actions. Under a predictive model, surprising (low probability) stimuli or actions necessitate the immediate reallocation of processing resources, but they can also signal the need to update the underlying predictive model to reflect changes in the environment. Surprise and updating are often correlated in experimental paradigms but are, in fact, distinct constructs that can be formally defined as the Shannon information (I_S) and Kullback–Leibler divergence (D_{KL}) associated with an observation. In a saccadic planning task, we observed that distinct behaviors and brain regions are associated with surprise/I_S and updating/D_{KL}. Although surprise/I_S was associated with behavioral reprogramming as indexed by slower reaction times, as well as with activity in the posterior parietal cortex [human lateral intraparietal area (LIP)], the anterior cingulate cortex (ACC) was specifically activated during updating of the predictive model (D_{KL}). A second saccade-sensitive region in the inferior posterior parietal cortex (human 7a), which has connections to both LIP and ACC, was activated by surprise and modulated by updating. Pupillometry revealed a further dissociation between surprise and updating with an early positive effect of surprise and late negative effect of updating on pupil area. These results give a computational account of the roles of the ACC and two parietal saccade regions, LIP and 7a, by which their involvement in diverse tasks can be understood mechanistically. The dissociation of functional roles between regions within the reorienting/reprogramming network may also inform models of neurological phenomena, such as extinction and Balint syndrome, and neglect.

eye movement | prediction | attention | learning | Bayes

In a nonrandom environment, brains can and should make use of past experience to facilitate the processing of incoming sensory information and the selection of actions, through prediction (1, 2). An important aspect of brain function is therefore the construction and tuning of internal models to represent statistics of the environment that are relevant for future behavior.

The use of predictive internal models implies that not only are some events well predicted (high probability under the model) but, conversely, some events (which have a low probability under the model) are surprising (3). Surprising events may be associated with behavioral costs; for example, although valid attentional cues speed reaction times (RTs), invalid cues lengthen them (3, 4). However, surprising events can have a further significance for the observer in that they sometimes provide evidence for a change in the environment, which would imply a need to update the brain's internal models to predict future events accurately.

Here, we explore the possibility that the brain carries out at least two distinct operations when a surprising event occurs: (*i*) within trial reorienting processes evoked by surprise, including reallocation of resources to a previously deprioritized region of space and/or replanning a motor response to an unexpected stimulus, and (*ii*) between-trial processes, particularly the possible

need to update the internal model to predict future observations accurately in a changeable environment.

The within- and between-trial processes could be broadly characterized in terms of surprise (or rather the reprogramming/reorienting response caused by the surprising stimulus) and updating (of the entire model), respectively. Consider, for example, the classic Posner orienting task (3), in which the locations of visual events are predicted either explicitly by symbolic cues or implicitly by the fact that targets appear more frequently in certain locations. Invalidly cued (surprising) targets evoke behavioral reorienting (redirecting of attention or gaze to the surprising target location), but they may also cause the participant to update his/her beliefs about the probable locations of future targets.

Information theory gives distinct definitions of surprise and updating that formalize the distinction between the surprise (and consequent behavioral reorienting) evoked by a particular stimulus and updating of the overall model of the environment.

In information theory, the surprise associated with a particular stimulus value, α, is characterized by its Shannon information [$I_S(\alpha)$]:

$$I_S(\alpha) = -\log p(\alpha|\text{prior}), \qquad [1]$$

where $p(\alpha|\text{prior})$ is the prior probability that the observation α would be made, given the brain's internal model just before the data point was observed. Therefore, the I_S captures how unexpected or unlikely a particular observation is, given the internal model.

Significance

This study investigates the brain mechanisms by which people disregard their previous beliefs about their environment and start forming new beliefs. Surprising events are often a signal that one's previous beliefs are no longer valid. Using brain imaging, we identified separate brain systems involved in dealing with the immediate consequences of surprise (i.e., reprogramming actions) and in updating one's beliefs about the environment to predict future events accurately. We present a mathematical and neuroanatomical model of how brains adjust to change in their environment that may inform our understanding of neurological disorders in which this adjustment process fails.

In contrast, updating of the internal model is captured by the Kullback–Leibler divergence (D_{KL}) between the posterior and the prior:

$$D_{KL}(\text{post}\|\text{prior}) = \sum_{\alpha} p(\alpha|\text{prior})[\log p(\alpha|\text{prior}) - \log p(\alpha|\text{post})],$$

[2]

where $p(\alpha|\text{prior})$ is the probability that the observation α would be made, given the model just before α was observed, and $p(\alpha|\text{post})$ is the same quantity, given the updated model just after α was observed.

As evident in Eq. 2, the D_{KL} is the probability-weighted average change in the I_S across all possible stimuli as a consequence of updating the model. Hence, although the I_S describes the degree of surprise evoked by observing a particular data point α, the D_{KL} describes how the model, as a whole, is updated as a consequence of observing α.

Although surprise and updating have distinct computational definitions, they are usually strongly correlated in experimental paradigms. As already noted, invalid targets in a Posner paradigm could evoke just surprise or both surprise and updating, depending on the task design. More theoretically, in temporal difference learning (5), the updating of action or stimulus values is driven by prediction error (surprise). Furthermore, the terminology used is confusing because the D_{KL} (updating of the model) has been described as "Bayesian" surprise (6). Note, however, that the constructs are behaviorally dissociable: Surprise need not necessarily lead to updating, and updating can be triggered without surprise. The relationship between surprise and updating depends, among other things, on the learning rate (7), the degree of expected stochasticity in the environment (8, 9), and the expected frequency or rate of change in the underlying environment (10). Furthermore, there exist scenarios in which we might hypothesize that updating should be triggered in the absence of surprising observations; for example, if the observer moves into a new context in which he knows a priori that his old internal models are unlikely to be valid (11).

In the present study, we wished to investigate the neural mechanisms by which expectations (internal models) are updated, as distinct from the mechanisms by which the immediate behavioral response is reprogrammed to a surprising stimulus. We used saccadic planning as a context in which to investigate this problem.

The choice of a saccadic planning task was motivated as follows. First, spatial attention and saccadic planning are clear examples of predictive models, in which the violation of predictions (as in invalid trials on a Posner task) evokes a process of behavioral reorienting that is measurable in terms of RT costs (3).

Second, the neural nature of predictive models over saccadic target locations (the representations to be updated) is relatively well understood. Networks of spatially tuned cells in macaque parietal cortex represent maps over space of variables relevant to saccadic planning, such as the probability of targets appearing or the reward value associated with making a saccade to different points in space (12–15).

Finally, the contrast between saccadic reprogramming (as evoked by surprise) and updating of internal models has intriguing parallels in the neurological literature. Although it is generally recognized that the posterior parietal cortex plays a key role in control of eye movements and spatial attention, it has long been noted that there is a fundamental difference between the neurological syndromes that follow damage to the intraparietal sulcus (IPS), including the lateral intraparietal area (LIP), and to the inferior parietal lobule (IPL) (16, 17). On the one hand, impairments in the ability to reorient or redirect the eyes (18) in Balint syndrome or the covert focus of attention (4) in the extinction syndrome, for example, are associated with damage to the IPS that includes the LIP. By contrast, the quite distinct

syndrome of neglect is associated with damage to the IPL (16, 19). In the neglect syndrome, patients' problems cannot be described as simply the inability to reorient to the neglected visual field when two stimuli are present as is the case in extinction. Instead, patients with the neglect syndrome are unable to acquire, despite experience and instruction, an expectation of any event of significance in the neglected field.

We developed a paradigm in which participants made speeded saccades to visual targets, the locations of which could be predicted by estimating their underlying spatial distribution. To dissociate the neural processes associated with surprise and updating in the experimental paradigm, we used a simple manipulation of the relevance of surprising targets in terms of the extent to which they predicted the locations of future targets, so that both surprise (probability of observations under the internal model) and updating (change in the internal model) varied independently across trials. Using functional MRI (fMRI) and behavioral measures, including pupillometry, we observed both neural and behavioral dissociations between surprise and the updating of an internal model.

Results

Seventeen participants completed two sessions of the same task: a behavioral session in which high-quality eye tracking data were acquired and a session with fMRI and concurrent eye tracking.

We sought to distinguish the neural response to surprise per se from neural activity occurring when an internal model was to be updated. To do this, we designed a saccade task in which the surprise associated with target locations (I_S) and the updating of the internal model (D_{KL}) were dissociated by a relevance manipulation.

The task was a simple saccadic eye movement response task in which participants could use prior knowledge about the spatial distribution of saccadic targets to facilitate speeded eye movement responses to them. On each trial, participants began by fixating a central cross. First, a warning signal appeared (the cross brightened for 400 ms); immediately afterward, a colored dot, the saccadic target, appeared for 350 ms. Participants were instructed to move their eyes as quickly as possible to look at the dot and then return fixation to the central cross. We measured saccadic RT using an IR eye tracker [EyeLink 2000 (SR Research) sampling at 500 Hz].

Participants could anticipate the location of the target dot because the dots appeared in similar locations over a run of several trials (mean run length of 15 trials, range of 10–20 trials). The distance of the dot from the central fixation point was fixed (so all targets appeared on a circular perimeter); hence, the position of the dot was governed by a single parameter, angle from vertical. This angle, α, followed a circular Gaussian distribution with a mean and variance that remained fixed during each run but moved abruptly to new values between runs. Runs were not temporally separated (i.e., the first trial of run n followed directly from the last trial of run $n - 1$), but the start of each new run was explicitly signaled by a change in dot color; thus, all the targets in run n may have been red and all the targets in run $n + 1$ may have been blue, etc.

Although most trials (75%) were drawn from Gaussian distributions as described above, we included another set of trials, called "one-offs" (25% of trials), that had locations randomly selected from a uniform distribution over the whole circle and were randomly interspersed with the other trial types. Hence, the generative probability density function over dot locations was the sum of a normalized Gaussian scaled by 75% and a normalized uniform over the whole circle scaled by 25%:

$$p(\alpha) = 0.75 \ p(\alpha|\alpha \sim \mathcal{N}(\mu, \sigma^2)) + 0.25 \ p(\alpha|\alpha \sim U(0°, 360°)).$$

[3]

Importantly, participants received explicit signals informing them when the current trial was the start of a new run or a one-off trial. Most target dots were chromatically colored, and the beginning of a new run was signaled explicitly by a change in the color of the target dots (so all the dots in one run would be, say, red and all the dots in the next run would be, say, blue). In contrast, the dots on one-off trials were always colored gray. Therefore, although both one-off and update trials shared the feature of the target dot appearing in an unexpected location, the two types could be easily distinguished; the target dot was gray on update trials but chromatically colored on trials drawn from the Gaussian distribution and took a new chromatic color on update trials. The task is illustrated in Fig. 1.

Behavioral Confirmation of Task Strategy. The logic of including one-off trials was to dissociate the effects of surprise and behavioral reprogramming from the process of updating an internal model. On both update and one-off trials, we expected participants to be surprised by the target location and to engage motor reprogramming to cancel the planned saccade and plan a saccade to the new target location. However, updating should only occur on the update trials, because participants were explicitly informed that one-off trials had no predictive value relating to future trials.

Participants learn on update trials but not on one-off trials. RT data confirmed that participants did learn on update trials (Fig. 2A). Although participants' gaze took longer to reach the target on the first trial of a new run (the update trial) compared with "expected" (non-one-off, non-update) trials ($t_{16} = 4.7, P = 0.0001$, paired samples t test), there was no further behavioral cost on subsequent trials; on the second and subsequent trials of a new run, there was no significant difference in time for the gaze to reach the target compared with the average of expected trials ($t_{16} = 0.48, P = 0.32$). However, RTs were significantly ($t_{16} = 3.0$, $P = 0.0045$) faster than average on the last trial of the block (trial −1 in Fig. 2A), perhaps because of a gradual increase in response speed across the block.

As well as confirming that participants learned on update trials, it was important to confirm that they did not learn on one-off trials. One possible problem with our design would be if participants, even involuntarily, used the location of the target on one-off trials to update their internal model of the underlying distribution for non–one-off trials (because we would then expect update-related brain activity on one-off trials, potentially leading to a false-negative result). If this were the case, we would expect to see slower RTs on trials that immediately followed one-off trials. This is because if the internal model had been updated on a one-off trial, participants should now expect subsequent trials to appear near the one-off target. However, we did not observe slowing on responses following a one-off trial; although participants' gaze took longer to reach the target on one-off trials than on expected trials ($t_{16} = 2.0, P = 0.03$, paired samples t test), RTs on trials following a one-off were very similar to the average expected trial (mean time to reach target for a trial immediately following a one-off was 408 ms, and mean time for all expected trials was 402 ms; $t_{16} = 0.48, P = 0.32$ for a paired samples t test between these values across participants). This suggests that participants successfully refrained from updating their internal model for chromatic dots based on the one-off trials. Equivalently, one could say that there was no evidence that the learning rate on one-off trials was above zero.

RTs following update- and one-off trials are illustrated in Fig. 2A, which shows RT (time to fixate the target) by trial number for one-off and update trials, where trials are sorted so that trial 0 is the one-off or update trial and trial +1 is the subsequent trial, followed by trials +2, +3, etc. Trial −1 is the trial before the update or one-off trial. To reiterate, although the plots for one-off and update trials appear similar, they logically indicate opposite effects. To predict the position of the dot on trial +1 correctly, participants would need to update their internal model on update trials but not on one-off trials.

Analysis of RTs on the update and one-off trials themselves indicated no overall difference in time to reach the target be-
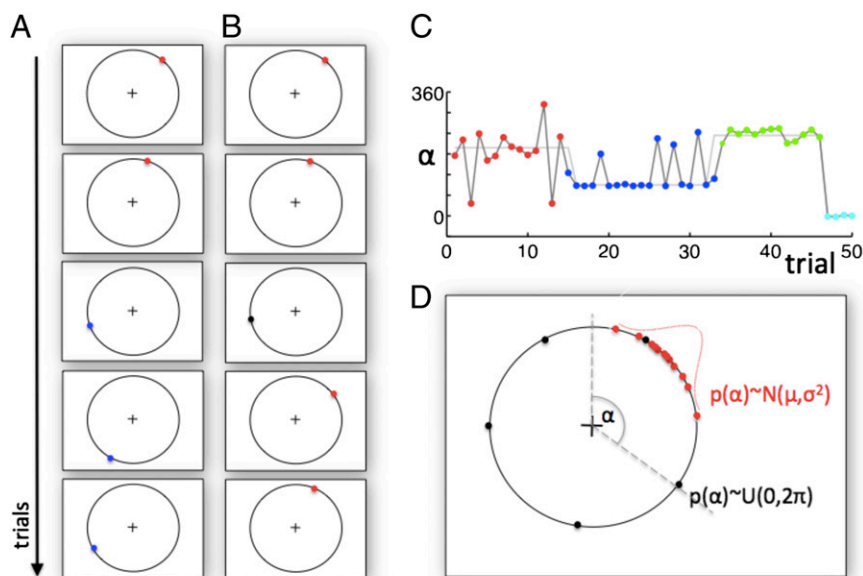

Fig. 1. On each trial of the task, participants began by fixating a central cross. A target (colored dot) appeared on a circular perimeter. Its location was predictable because target locations were similar over runs of 10–20 trials (mean of 15 trials). Two types of unexpected target locations could be observed, as illustrated in A and B. (A) Transition between runs (update trial). Initially, red dots are observed in the upper right; subsequently, blue dots are observed in the lower left. (B) One-off trial. A target appears in an unexpected location (targets are expected in the upper right, but the one-off target appears in the lower left). However, this is a one-off trial, and future targets revert to the original distribution in the upper right. Targets on one-off trials are colored gray. (C) Plot of target locations (angle α from vertical) over 150 trials. Different colored targets are from different runs. One-off targets are shown in gray. (D) Distribution of target locations within a run is a combination of a circular Gaussian, shown in red, and a uniform distribution, shown in black, from which one-off trials are drawn.
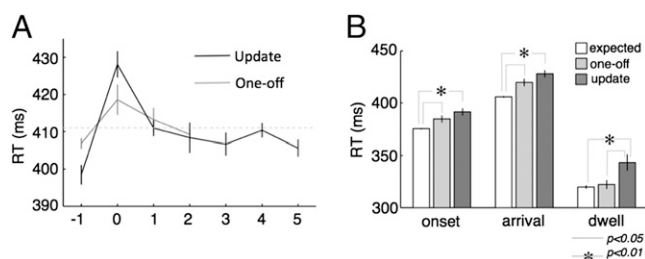
Fig. 2. RTs indicate that participants performed the task as instructed. (*A*) RTs (time for gaze to arrive at the target) on update and one-off trials, as well as on surrounding trials. The dashed line is the mean RT for all trials. (*B*) Breakdown of saccadic RTs into onset time (time to leave central fixation), arrival at target (onset of fixation at the target), and dwell time (duration of fixation at the target).

tween update and one-off trials (t_{16} = 1.22, P = 0.12, paired samples t test).

Saccadic RTs reflect surprising target locations, but dwell time reflects updating. We conducted a more detailed analysis of saccadic responses in terms of RTs for saccadic onset; arrival at the target (as reported above); and dwell time, the time spent looking at the target. These comparisons are illustrated in Fig. 2*B*.

In line with the hypothesis that within-trial response production is slowed by unexpected target locations, saccade onset and arrival times were slower on both of the trial types containing surprising target locations (update and one-off trials), with both update and one-off trials having slower RTs than expected trials (paired sample t test for saccade onset time: t_{16} = 4.3, P = 0.0003 and t_{16} = 2.24, P = 0.02 for update and one-off trials vs. expected trials, respectively; no difference between update and expected trials: t_{16} = 0.86, P = 0.20).

In contrast, dwell time was specifically affected by updating, with update trials differing significantly from the other trial types: Dwell time was longer on update trials than on one-off trials (t_{16} = 2.02, P = 0.03). Dwell time was also longer on update trials compared with expected trials (t_{16} = 2.6, P = 0.01), and there was no significant difference in dwell time between one-off and expected trials (t_{16} = 0.26, P = 0.6).

RTs were driven by spatial, not feature-based, surprise. It could be argued that RT costs were partly driven by surprise associated with nonspatial, or nonoculomotor, aspects of the one-off and update stimuli, notably their color. However, behavioral data suggest this was not the case. Because one-off trial locations were uniformly distributed around the target circle, some of the one-off targets fell within the expected spatial range. We compared RTs on one-off trials that fell within the expected spatial range (hence evoking feature-based but not spatial surprise) with RTs on one-off trials that fell outside the expected spatial range (hence evoking spatial surprise as well as feature-based surprise). Only on one-off trials that were spatially surprising were RTs slower than on expected trials, suggesting that feature-based surprise did not affect behavior (Fig. S1).

Model-Based Measures of Surprise and Updating. The RT data suggest that participants performed the task as instructed [i.e., they used prior knowledge of the distribution of target locations to facilitate speeded eye movements (because time to arrive at the target was longer when the target location was unexpected)] and, furthermore, that participants updated their internal model, as instructed, on update trials but not on one-off trials, as indicated by the data in Fig. 2*A*. They also indicated that whereas the production of the saccadic response within trials was slowed for all surprising target locations (linking surprise and behavioral reorienting), the processing of the target location (indexed by dwell time) was specifically affected by updating.

To give a formal account of the two processes (surprise/reorienting and updating), we wished to analyze the behavioral and neural correlates of two information theoretical measures, I_S (surprise) and D_{KL} (updating), as described in the Introduction. These measures must be calculated with reference to a model of the state of the environment; for example, the I_S is the log probability of an observed data point, given some model. Specifically, we wished to calculate the I_S and D_{KL} on each trial with reference to the participants' internal model of the target distribution.

Participants' internal models of the targets' distribution could be expected to differ from the true (generative) distribution of target locations because they observed a limited number of data points and data were observed sequentially; hence, for example, in the first few trials after a change of distribution, the internal model would be based on relatively few data points.

To obtain an estimate of a participant's beliefs (i.e., the state of the participant's internal model of the environment) on a trial-to-trial basis, we constructed a normative Bayesian learner. The model is described in detail in *Methods*; however, briefly, its key features were that it updated its estimate of the underlying distribution using data only from non–one-off trials and the estimated distribution on each trial was updated with the new data point using Bayes' rule. Two features of the model merit emphasis here.

First, no updating occurred on one-off trials (equivalently, one might say the learning rate was set to zero on one-off trials). This was in accordance with the instructions given to participants (that the locations of one-off targets did not predict the location of future targets). Behavioral evidence suggests participants complied with these instructions, because RTs on the trials following one-off trials were not slowed, as they should have been if participants had erroneously updated their model on the one-off trial. However, it is worth noting that if participants did update on one-off trials, this could only lead to a false-negative result in terms of detecting update-specific brain activation because it would reduce the contrast between update and nonupdate trials.

Second, on the first trial of a new run, the prior in force before the update was first "blanked" (replaced with a uniform distribution); hence, the new posterior model for the update trial (on which the prior for trial +1 was based) was obtained by updating from a uniform distribution using Bayes' rule. This "blanking" step was introduced to reflect a feature of the task design, of which participants were explicitly informed: that the location of one "block" of target dots and the next block (after an update trial) were totally independent, such that expectations about target locations based on previous targets should be disregarded.

The model was Bayes' optimal, implying that it returned the best possible estimate of the underlying distribution based on the data points it was given (the same data points that human participants observed) (20). Furthermore, the model was optimal in that it was supplied with correct assumptions about the structure of the world. For example, the model correctly assumed that when an update occurred, the new distribution of target locations did not depend on the previous distribution; this reflected the true form of the generative process by which target locations were chosen in the experiment. Therefore, the internal model generated by our Bayesian learning algorithm represents an upper bound on the accuracy with which any participant could have estimated the underlying distribution.

Behavioral correlates of information theoretical measures. We ran the Bayesian learning model with the same sets of target locations experienced by our human subjects to obtain an estimate of what their internal models of the environment would be on each trial. This allowed us to determine the surprise (I_S) associated with each stimulus and the degree of updating (D_{KL}) on each trial. In addition, we modeled the strength of participants' prior expectations on each trial as the Shannon entropy of the prior distribution (H_S) because behavioral work indicates that saccadic RTs

may depend on this factor (21). The H_S depended on both the variance of the generative targets' distribution, which was experimentally manipulated, and on learning. The formal definitions of these quantities in terms of the model are given in *Methods*.

RT data indicate that both surprise and updating influence RT. As before, we broke down RTs into onset, arrival, and dwell times, and we modeled these values using a general linear model (GLM) analysis in which RT was modeled as a linear combination of effects of surprise/I_S, updating/D_{KL}, and H_S. The results are shown in Fig. 3; there was an effect of both surprise and updating on all three measures: onset, arrival, and dwell times (statistics for effect of surprise are $t_{16} > 5.1$, $P < 0.00005$; $t_{16} > 5.1$, $P < 0.00005$; and $t_{16} = 1.9$, $P = 0.037$ and statistics for effect of updating are $t_{16} = 2.8$, $P = 0.0059$; $t_{16} = 4.3$, $P = 0.0003$; and $t_{16} = 2.9$, $P = 0.0049$ for saccadic onset, arrival, and dwell times, respectively, in one-sample t tests against zero). The entropy of the prior had no significant effect on any measure ($t_{16} = 1.3$, $P = 0.90$; $t_{16} = 1.3$, $P = 0.90$; and $t_{16} = 0.76$, $P = 0.23$ for one-sample t tests against zero as above).

Is the updating effect simply due to a quantitative enhancement of surprise? Given that RTs showed a parametric effect of surprise with an additional effect of updating, it might be argued that the occurrence of an update stimulus simply caused a quantitative enhancement of the surprise effect. However, several results argue that, in fact, the additional RT cost on update trials reflects a qualitatively different neural process to the effect of surprise. First, participants' behavior on the trial following an update or one-off trial (Fig. 2*A*) indicates that participants adjusted their internal model of targets' distribution on update trials but not on one-off trials. Second, dwell time was enhanced on update trials compared with one-off trials, a possible behavioral correlate of updating. Finally, there were dissociable effects of surprise and updating on both pupil diameter and brain activity. We will now consider these dissociations.

Pupil dilation shows opposite responses to surprise and updating. Recent work has linked pupil dilation responses to psychological constructs similar to those investigated in this study. For example, Nassar et al. (22) observed that pupil dilation increased on trials when participants detected change points in the environment (which should trigger both surprise and updating), whereas Preuschoff et al. (23) have linked pupil dilation to uncertainty about the state of the environment and to changes in that uncertainty (and hence to learning).

To investigate whether pupillometric effects were correlated with surprise, updating, or both, we analyzed pupil dilation data in our task. We again used a GLM approach in which the effects of surprise, updating, prior entropy, and the main effect of task were modeled as a function of time in each trial (*Methods*); these were the same regressors used in the fMRI analysis (below) and in the analysis of RT data presented in Fig. 3. Pupillometric
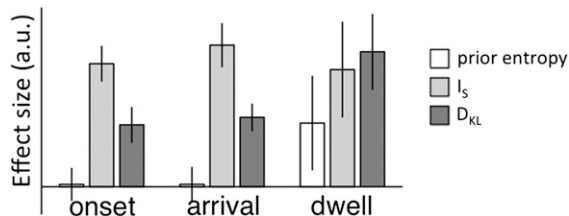

Fig. 3. Effect of prior entropy, surprise, and updating on saccadic responses. Results of a GLM analysis on saccadic onset, arrival, and dwell times in which the regressors were entropy of the prior (the strength by which the stimulus location was predicted), I_S, and D_{KL}. The bar height is the mean effect size (beta value) for each parameter across participants, and error bars are the group SEM. a.u., arbitrary units.

effects are shown in Fig. 4, whereas the raw pupil dilation data (raw pupil area on each trial type) are shown in Fig. S2.

There was a positive effect of surprise on pupil diameter. On trials with higher values of I_S, there was a greater dilation of the pupil after viewing the target. This effect is in line with previous work that reported increases in pupil dilation when participants make observations that have a low probability, given their current model of the environment (22).

However, we also observed that updating was associated with a negative effect on pupil diameter [i.e., there was a relative decrease in pupil size on trials with a high D_{KL} (update trials)]. This effect peaked at a slightly later time than the surprise effect (Fig. 4). Inspection of the raw data (Fig. S2) indicates that pupil diameter was actually decreased in this period on update trials compared with expected trials (and one-off trials).

Dissociable effects of surprise and updating on pupil diameter have not previously been reported, probably because these parameters are generally strongly correlated in most experimental designs. However, the negative effect of D_{KL}/updating on pupil diameter was somewhat unexpected because of theoretical work linking pupil dilation and noradrenaline/norepinephrine to learning and uncertainty. Pupil dilation in the absence of luminance changes may correlate with the firing of cells in the locus coeruleus, and hence with the release of noradrenaline, although the strength of the correlation is debated (23–25). Theoretically, it has been argued that noradrenaline signals uncertainty about the state of the environment (i.e., estimation uncertainty) (11, 26) or that it acts as a kind of "reset" signal for internal models (25, 27).

To replicate the unexpected finding that updating negatively modulates pupil diameter, we conducted an additional behavioral experiment with 18 participants to check potential confounding factors that could have driven the pupillometric effect. In the replication experiment, participants could predict target positions as in the main task, but the task was to make orientation judgments about isoluminant Gabor patches while fixating centrally, eliminating potential confounds of target color and eye movement. Details of this experiment are described in *SI Methods*; in summary, we observed a very similar pattern of opposite effects of surprise and updating as in the main experiment (Fig. S3).

It is unclear how the update-specific decrease in pupil area fits with computational theories of noradrenaline (24, 25), which have linked learning, uncertainty, noradrenaline release, and pupil dilation and would tend to predict an increase in pupil dilation on update trials. A key difference between the present paradigm and probabilistic learning tasks used in previous studies (22) is that in the present task, unlike in probabilistic learning tasks, updating is not driven by increased uncertainty; in fact, at the same moment that the observer detects a change in the environment (an update target), he also gains information about the new target distribution, reducing uncertainty. This interpretation would suggest that pupil increases in learning tasks are driven by uncertainty, or the influence of uncertainty on learning, rather than by learning or change per se.

The observation of a decrease in pupil diameter on update trials, at a different time point from previously reported uncertainty-related pupil dilations, is unique (previous experiments have not reported results in this time period; indeed, previous studies have not dissociated surprise and updating). The timing of the effect raises the intriguing possibility that the late dip in pupil diameter on update trials reflects neural processing of the stimulus itself (modulated by its relevance) rather than a modulation of the prior preparation for learning (24, 25).

In relation to these results, it should be noted that multiple psychological factors, including arousal and attention, can affect pupil dilation and that the effect should not be overinterpreted in the absence of pharmacological work. Nonetheless, the effect is replicable and provides robust evidence for differential processing update and nonupdate stimuli.
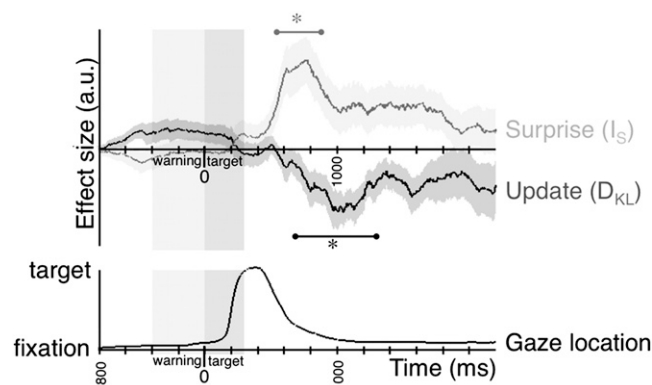
Fig. 4. Pupil dilation is affected by both surprise and updating. The data are plotted as a function of time in the trial (shown on the x axis; labels in the lower panel apply also to the upper panel). The shaded regions represent the period in which the warning stimulus (a brightening of the central fixation cross) and the target were present on the screen. (*Lower*) Average eye displacement as a function of time, which gives an indication of when participants were looking at the target. (*Upper*) Results of a GLM in which pupil dilation at each time point was modeled with the same regressors used in the analysis of RTs (Fig. 3) and fMRI data (Fig. 5): entropy of the prior, surprise (I_S), and updating (D_{KL}) on the trial. The main effect (mean pupil dilation at that time point across all trials) was also modeled. The results show that although surprise (I_S) was correlated with increased pupil diameter, updating (D_{KL}) was associated with a relative decrease in pupil diameter. The time periods in which these effects were significant [Z-score is >2.3 (i.e., $P <$ 0.01 uncorrected, two-tailed)] are indicated by the starred bars above and below the plot. The effect of D_{KL} is significant in an interval 730–1,240 ms after target onset; for I_S, the interval is 540–908 ms.

Distinct Brain Networks for Surprise and Updating. Our behavioral results suggested that although the within-trial behavior (saccadic RT) was affected by both surprise and updating, stimuli that elicited updating (high D_{KL}) were processed differently: (*i*) because behavior on subsequent trials supported the hypothesis that participants updated on update but not one-off trials; (*ii*) because of increased dwell times on update trials; and (*iii*) because the D_{KL} produced an opposite effect on pupil dilation to the I_S.

To determine the neural correlates of surprise and updating, we collected fMRI data from the same group of 17 participants for whom behavior was presented while performing the task. Note that eye tracking data were acquired during the fMRI session and that these data show a very similar pattern of RT effects to the data reported above (replications of Figs. 2 and 3, using the eye tracking data from the fMRI session, are provided in Fig. S4); however, pupillometry was not possible with the in-scanner setup. Details of fMRI data acquisition, preprocessing, and analysis are described in *Methods*.

We began by analyzing the pattern of activity associated with surprise and updating across the whole brain, by means of a GLM analysis implemented using the Centre for Functional MRI of the Brain (FMRIB) software library (28). The task was modeled in terms of four model-derived regressors: main effect of task (i.e., all trials, modeled as a delta function at the time of the target appearance) and three information theoretical parameters, the prior entropy (the predictability of the environment), the I_S of the observed target location (surprise), and the D_{KL} of the posterior from the prior (updating). Quantitative definitions of these regressors are given in *Methods*.

We observed a spatial dissociation between activity correlated with surprise (I_S) and activity correlated with updating (D_{KL}); no significant activity associated with prior entropy was observed (in line with the lack of behavioral effects relating to this parameter).

Surprise (I_S) was strongly correlated with activity in the posterior parietal cortex, with the peak activation in the superior parietal lobule (SPL) and extending along the IPS. In contrast, updating (D_{KL}) was strongly correlated with activity in the anterior cingulate cortex (ACC), particularly the rostral cingulate motor area (rCMA) and extending into the ventral part of the adjacent presupplementary motor area (pre-SMA). These spatially distinct effects, which were the only effects to survive multiple comparisons correction from the whole-brain analysis, are shown in Figs. 5*A* and 6*A*. For illustrative purposes, the uncorrected statistical maps from which the corrected statistics were drawn are shown in Fig. S5. Note that all results presented in the main text are corrected for multiple comparisons.

Although we had specific parametric predictions about brain activity based on the I_S and D_{KL}, as a confirmatory analysis, we also analyzed the effects of different trial types: update trials + one-off trials (both types evoke surprise/behavioral reorienting) and update trials − one-off trials (to isolate behavioral reorienting). The maps obtained are very similar to those obtained using our parametric regressors (Fig. S6). Region of interest (ROI) effects using the trial type and parametric regressors are compared in Fig. S7.

To characterize these effects further, we extracted and analyzed the fMRI signal from anatomically defined ROIs as follows.

ACC: Updating. The presence of updating-specific activity in the ACC was of particular interest because the ACC has previously been implicated in the updating of expectations about the reward environment (10, 29, 30) but not in spatial updating. At the same time, the ACC has been implicated in studies of spatial reorienting (31), but the nature of its specific contribution has remained unclear so far.

To characterize the role of the ACC in our task, we extracted the fMRI signal from an anatomical ROI, the rCMA, using a mask derived from a diffusion-imaging parcellation of the human
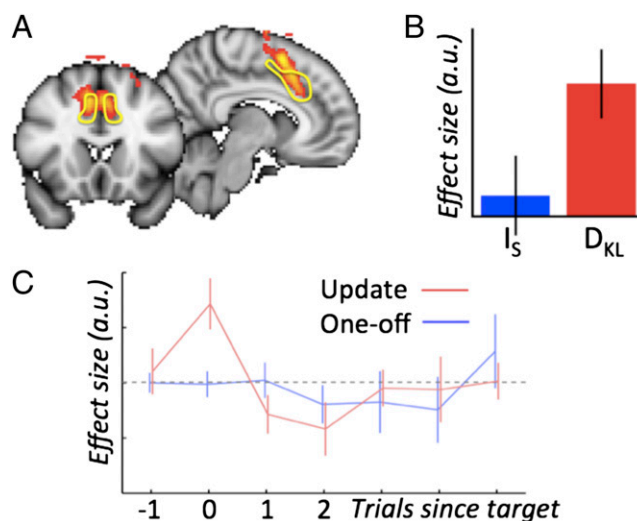

Fig. 5. Effect of updating in the ACC. (*A*) Results of whole-brain fMRI analysis. This region in the ACC and pre-SMA was the only area in which there was a significant effect of updating (contrast shows all voxels with a parametric effect of D_{KL} as defined in *Methods*), using cluster size-based multiple comparisons correction. The color scale is $2.3 < Z < 3$; the peak Z-score is 3.1 at MNI coordinates (6, 10, 54). The ROI denoted by the yellow line is the ACC ROI, a region described as the rCMA zone in a diffusion-weighted parcellation of the cingulate cortex (32). This was the ROI used in the analyses shown in *B* and *C*. (*B*) Effect size for I_S and D_{KL} in the ACC ROI, where bar height is the mean effect across the group of participants and error bars are the SEM. (*C*) Raw activity in the ACC ROI plotted as a function of trial-in-run; plotting conventions are as in Fig. 2.
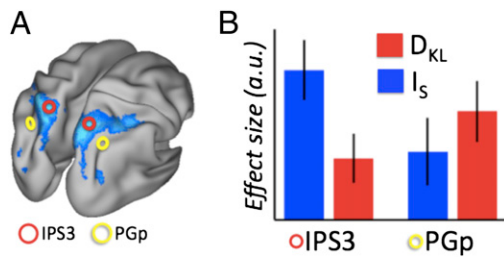
Fig. 6. Effect of surprise in the posterior parietal cortex. (*A*) Results of whole-brain fMRI analysis. This region in the posterior parietal cortex was the only area in which there was a significant effect of surprise (contrast shows all voxels with a parametric effect of I_S as defined in *Methods*), using cluster size-based multiple comparisons correction. The color scale is $2.3 < Z < 3$, and the peak Z-score is 4.8 at MNI coordinates (-18, -60, 58). The red circle is the IPS3 ROI (with coordinates ±26, -54, 60), and the yellow circle is the PGp ROI (with coordinates ±32, -64, 44) (33). (*B*) Effect size for the I_S and D_{KL} in each ROI; the bar height is the mean effect across the group of participants, and the error bars are the SEM. There is a significant ROI × condition interaction.

cingulate cortex (32). This ROI corresponded to the more ventral portion of the updating activity observed in the whole-brain analysis (Fig. 5). Fig. 5*B* shows the effect size in this ROI for updating (D_{KL}, red) and surprise (I_S, blue); as expected, in this ROI, there was a significant effect of updating ($P = 0.015$, two-tailed *t* test against zero), but note also the lack of effect for surprise ($P = 0.61$).

To characterize the pattern of ACC activity across a run of trials, we plotted raw activity (averaged across a time bin 6–7 s after the target, timed to coincide with the peak of the hemodynamic response function) in the rCMA ROI by trial number within a run for runs of trials around both update and one-off trials (using the same plotting conventions as in Fig. 2). The rCMA showed a pattern of activity that suggested it was uniquely activated on the first trial of a run: The rCMA was activated on update trials ($t_{16} = 2.7$, $P = 0.0075$, two-tailed *t* test of raw activity in the rCMA across a time bin 6–7 s after the target on update trials vs. mean activity at the same time point across all other trials) but not on subsequent trials or on one-off trials.

Posterior parietal cortex: Action reprogramming evoked by surprise. Surprise (I_S) was correlated with activity in the posterior parietal cortex. The activity extended along the SPL and into the IPS. Its peak fell in a region on the posterior medial bank of the IPS and adjacent SPL [Montreal Neurological Institute (MNI) coordinates -18, -60, 58]. This region, which we will call IPS3 in agreement with the terminology used by Mars et al. (33), has been postulated to be the human homolog of the monkey LIP because, like macaque LIP, it has strong connections with the frontal eye fields (33) and superior colliculus (34), and is retinotopically organized (35, 36). Given this motoric connectivity profile, the presence of the surprise (I_S)-evoked activity in the IPS may well reflect reprogramming of the current saccade to account for the new (surprising) target location.

In macaque parietal cortex, LIP is one of two regions that are especially linked to covert attention and overt eye movements; the other is area 7a (37). We were intrigued by the possibility that the two areas could play different roles in behavioral reorienting and updating in our task, because in contrast to LIP, area 7a is less strongly connected with oculomotor regions but more strongly connected with the ACC, in which we observed update-specific activity (38, 39). Additionally, as mentioned in the Introduction, there is a suggestion from the patient literature that lesions of the IPS (including human LIP homolog IPS3) and IPL (including human 7a homolog PGp) differently affect motoric and nonmotoric aspects of orienting (16, 17).

To address this hypothesis, we defined anatomical ROIs in the two functional areas. The ROIs were based on the diffusion-imaging parcellation of Mars et al. (33) and were defined based on the two connectivity-defined regions labeled IPS3 (LIP homolog) and PGp [a region on the posterior lateral bank of the IPS and adjacent posterior IPL that has been identified as a putative homolog of monkey area 7a (33, 40)]; the peak voxels were at MNI coordinates (±26, -54, 60) and (±32, -64, 44), respectively.

There was indeed a significant difference between ROIs in their responsiveness to updating (D_{KL}), as opposed to surprise (I_S). We extracted the effect size from each of these ROIs in each participant. A within-subjects ANOVA with the factors ROI (IPS3 vs. PGp) and effect (I_S vs. D_{KL}) indicated a significant interaction between ROI and condition ($F = 12.5$, $P = 0.03$); in Fig. 6*B*, it can be seen that the IPS3 region was more active for surprise than updating, whereas the PGp region was more active for updating than surprise.

Parietal activity correlates with within-trial reprogramming costs. Throughout this paper, we have interpreted the neural activity associated with I_S in terms of the behavioral reprogramming of saccades, necessitated when a low-probability stimulus location occurs. In support of this hypothesis, activity in both the IPS and PGp was significantly related to RTs (higher activity on longer RT trials), whereas there was no significant relationship between RTs and ACC activity. We conducted a multiple regression in which RTs (taken separately for one-off trials and update trials) were modeled in terms of activity in the IPS3, PGp, and ACC. For both one-off and update trials, there was a significant effect of activity in the PGp and IPS3, but not in the ACC, on RT. Results are shown in Fig. 7. The *P* values for a group *t* test of effect size (the effect of ROI activity on RT in the multiple regression) were as follows: for one-off trials, $t_{16} = -1.6$, $P = 0.934$; $t_{16} > 4.0$, $P < 0.0005$; and $t_{16} > 4.0$, $P < 0.0005$ for the ACC, PGp, and IPS3, respectively, and for update trials, $t_{16} = 1.4$, $P = 0.096$; $t_{16} = 3.8$, $P = 0.0008$; and $t_{16} = 4.3$, $P = 0.0003$ for the ACC, PGp, and IPS3, respectively (two-tailed *t* test against null hypothesis of zero effect).

Discussion

We observed specific neural signals associated with the updating of an internal model (D_{KL}) and the parametric effect of surprise (I_S). Updating was associated with activity in the rCMA zone of the ACC and ventral pre-SMA. On the other hand, activity in the posterior parietal region IPS3, a putative homolog of monkey saccade planning area LIP (33), was correlated with the parametric effect of surprise (I_S).

Activity has been reported in both the posterior parietal cortex and the ACC when stimuli are presented that would be surprising given participants' expectations (in other words, their internal models) of the environment (41, 42). Lesions in both regions have been linked to failures of reorienting (43, 44). However, the
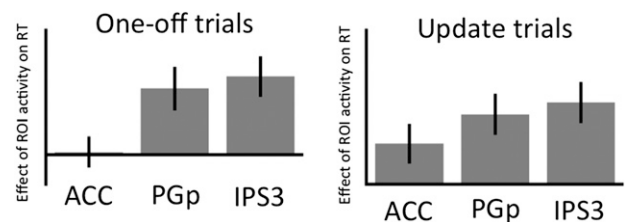

Fig. 7. Parietal activity, but not ACC activity, predicts RT on both one-off and update trials. Bars show the effect size (beta value) ± SEM from a multiple regression in which activity in the three ROIs (ACC, PGp, and IPS3) was used to predict RT. The effects of PGp and IPS3 activity on RT are significant for both one-off trials and update trials, and the effect of ACC activity is not significant in either case. A breakdown of this effect into onset, arrival, and dwell times is shown in Fig. S8.

specific roles of the parietal and cingulate components of the "reorienting network" have so far been unclear (31). Furthermore, there has been little attempt to relate the role of the two regions in responding to surprising stimuli with their other reported functions, such as the role of the ACC in reward prediction (29, 30) and error monitoring (45, 46). By identifying computational functions for the two regions, we hope the present results cast some light on this debate.

IPS, Surprise, and Saccadic Reprogramming. We observed activity in the IPS (area IPS3) that was parametrically related to the degree of surprise elicited by a stimulus (I_S); hence, stimuli that occurred in low-probability locations (given an optimal model of the environment) were associated with high levels of activity in the SPL.

This activity can be understood in terms of the need to reprogram planned eye movements in response to surprising target locations. Both surprise and posterior parietal activity were correlated with longer RTs, suggesting a link between behavioral reprogramming or spatial remapping (41, 42) and parietal activity. This remapping may be described as a within-trial effect because it is concerned with producing or reprogramming a behavioral response to the unexpected stimulus itself, whether or not predictions are updated for future trials.

Saccadic Reprogramming and Updating Expectations in Parietal Cortex. Although it is widely agreed that the posterior parietal cortex is active when either movements or covert attention is redirected or reoriented from one location to another (47–52), there is disagreement about which regions within the parietal cortex are most important for redirecting attention and eye movements (44). One possibility is that different parietal regions are concerned with different aspects of redirection (53).

A closer analysis of anatomically defined ROIs in the parietal cortex suggests that the computational functions defined in this study may differentiate two regions of the parietal cortex, both of which have been linked to saccadic planning. Region IPS3, a putative homolog of macaque region LIP, which has a generally motoric connection profile, was activated when surprise necessitated saccadic reprogramming. In contrast, posterior IPS/IPL region PGp, a putative homolog of macaque region 7a, was more strongly activated during updating. Human PGp and macaque area 7a are connected to the ACC, the main region activated during updating in the present study (38, 39).

This functional specialization within the parietal saccadic system is reminiscent of a long-standing observation from the neurological literature (16, 17) that lesions of the IPS and IPL produce different deficits in patients. Lesions of the IPS are associated with impairments in the ability to redirect the eyes, as in Balint syndrome (18), or the covert focus of attention, as in extinction (4). In contrast, spatial neglect is associated with damage to the IPL (16, 19). Neglect, unlike Balint syndrome and extinction, cannot be characterized as an inability to reorient to surprising events; rather, it seems to reflect a persistent inability to orient to the neglected field in the first place, which could be characterized as a distorted spatial prior.

ACC Is Specifically Involved in Updating. We observed activity in the ACC that was specific to trials on which the internal model was updated. ACC activity was not modulated by surprising stimuli that did not cause updating.

Although the ACC has previously been associated with the reorienting of attention and eye movements, and deficits therein, its specific role has remained unclear (31, 54). At the same time, little in the way of theory has been offered to relate the role of the ACC in reorienting to the role it clearly has in error monitoring and reward prediction.

The current findings suggest a specific computational function for the ACC: It is involved in updating internal models to facilitate future information processing. This mechanistic theory of ACC function suggests a possible framework in which observations that the ACC is involved in reorienting and attention may be reconciled with the abundant evidence for it having a role in reward processing, learning, exploration, and foraging.

Errors and Updating in the ACC. A major line of research on the ACC has focused on observations that the ACC is active in situations when participants make errors. It has long been known that a negative-going potential, the error- or feedback-related negativity, is evoked from the region of the ACC (45) when participants make errors in psychological tasks or are presented with feedback on their performance (46). However, these findings are not incompatible with a role for the ACC in updating of internal models; indeed, it could be argued that the functional role of error- or feedback-related signals is to facilitate the updating of internal models from which future action is generated (55). This interpretation is supported by the findings that activity in ACC cells is particularly strong in instrumental tasks (56) and, furthermore, that the magnitude of the error related negativity is proportional to the degree to which participants modify their behavior on future trials (45, 57).

Exploration, Updating, and Estimation Uncertainty. A separate line of research has linked ACC activity to exploratory behavior. The ACC is more active on experimental trials when participants explore alternative options rather than exploiting a known source of reward (58, 59). Furthermore, the ACC is active during foraging, when participants decide to "forage" (seek alternative options) as opposed to choosing from among the options immediately presented to them (60). Note that this role of the ACC in exploring and engaging in alternative actions is quite distinct from that of another frontal area, the orbitofrontal cortex, in updating the value or significance of specific stimuli (61, 62).

Although exploration and foraging may appear to be only loosely related to learning and updating, there is a computational concept that relates the two functions: control of estimation uncertainty. Estimation uncertainty (8) is uncertainty about the parameters of the environment. It is distinct from expected uncertainty (or risk), which characterizes the uncertainty or stochasticity that is inherent within a state of the environment and would persist even if the observer were certain about the parameters of the environment. In contrast, estimation uncertainty is not inherent in the environment but reflects the agent's knowledge of the environment and can be reduced if the agent has the opportunity to make further observations of the environment. Computationally, the level of estimation uncertainty affects the ability of an internal model to learn (or be updated), because a model that is totally certain about the state of the world should not learn anything new (9, 26); this idea is embodied in learning theory by the concept of "associability" (63).

Although estimation uncertainty is driven up by unexpected observations as in the present study (9), it can also be argued that estimation uncertainty should be elevated during foraging and exploration, because in exploring, the observer deliberately seeks new information; therefore, his neural networks should be in a state to accept new learning (11).

The pattern of activity of the ACC in learning tasks could support an interpretation of its role in terms of controlling estimation uncertainty. In a one-armed bandit task (10), it was observed that the ACC was more active in response to new observations when the environment was volatile (and estimation uncertainty was hence high). Furthermore, recent recordings from rat ACC (59) support the hypothesis that the ACC is active when estimation uncertainty should be elevated. In a two-armed bandit task, after the reward probability associated with different

actions changed, there was a radical shift in the pattern of activity across ACC neurons; this shift occurred at the start of a period of exploratory behavior rather than during the acquisition of new information per se, suggesting that the ACC was active at the point at which estimation uncertainty increased (when a model of the environment was abandoned in favor of "knowing nothing"). This result is particularly closely related to the present study, in which participants were instructed to disregard their previous experience as soon as the onset of a new experimental run was signaled; this was represented in our model by the application of a uniform probability distribution over all possible target locations at the start of each run.

Conclusions and Implications. Although updating and surprise are closely linked and are often correlated in learning paradigms, they are conceptually distinct processes. We observed distinct neural and behavioral responses correlated with surprise (I_S) and updating (D_{KL}). Our results suggest that although the SPL may play a particular role in spatial remapping and the reprogramming of actions within a trial, the ACC is only active when internal models are to be updated, in the context of change in the underlying environment. The results define distinct computational roles for two regions that have previously been shown to be involved in reorienting and reprogramming of saccades. Hence, they may inform models of attention and motor control, as well as models of neurological phenomena linked with failure to reorient, such as extinction and Balint syndrome. Conversely, the computational functions of the ACC and posterior parietal cortex in the present task may offer a framework in which the involvement of these regions in a variety of other paradigms can be explained.

Methods

Participants. Participants were 17 healthy volunteers age range 21–32 years, 7 females). All participants gave written informed consent in accordance with the National Health Service Office for Research Ethics Committees (ref 07/Q1603/11).

Model. The model was a normative Bayesian learner implemented numerically in MATLAB (MathWorks). It estimated the probability density distribution from which each target location, α_t, was drawn, based on the current and previous observations (i.e., $\alpha_{1:t}$), by assigning probabilities to putative parameters of the underlying Gaussian distribution $p(\mu_t, \sigma_t \mid \alpha_{1:t})$.

The posterior probabilities assigned to each parameter pair $p(\mu_t, \sigma_t \mid \alpha_{1:t})$ were calculated as follows. If the trial type is one-off, no updating occurs; that is,

$$p(\mu_t, \sigma_t | \alpha_{1:t}, \text{one}-\text{off}_t) = p(\mu_t, \sigma_t \mid \alpha_{1:t-1}). \qquad [4]$$

Otherwise, probabilities are updated using Bayes' rule:

$$p(\mu_t, \sigma_t | \alpha_{1:t}) \propto p(\alpha_t | \mu_t, \sigma_t) p(\mu_t, \sigma_t \mid \alpha_{1:t-1}, \varphi), \qquad [5]$$

where the variable φ denotes dependence of the prior on trial type as follows. If the trial type is update, a uniform prior over (μ_t, σ_t) is used:

$$p(\mu_t, \sigma_t \mid \alpha_{1:t-1}, \text{update}_t) = \text{1}/_{360 \times 50} \qquad [6]$$

where the denominator 360×50 represents the number of possible values for (μ_t, σ_t) probed in the numerical implementation of the model. If the trial type is expected (i.e., any trial that is not one-off or update), then the prior on trial t is obtained without modification from the posterior on trial $t − 1$:

$$p(\mu_t, \sigma_{t,} \mid \alpha_{1:t-1}, \text{expected}_t) = p(\mu_{t-1}, \sigma_{t-1} | \alpha_{1:t-1}). \qquad [7]$$

Full details are given in SI Methods.

A prior probability density function for the subsequent trial, t+1, was obtained from the posterior on trial t, and expressed as a function of target location (rather than parameters $\mu_t, \sigma_t,$) as follows:

$$p(\alpha_{t+1}|\alpha_{1:t}) = p(\text{expected}_{t+1}) \sum_{\mu_{t+1}} \sum_{\sigma_{t+1}} p(\alpha_{t+1}|\alpha_{t+1} \sim \mathcal{N}(\mu_{t+1}, \sigma_{t+1}))$$
$$\times p(\mu_{t+1}, \sigma_{t+1}|\alpha_{1:t}) + p(\text{one}-\text{off}_{t+1} \cup \text{update}_{t+1})$$
$$p(\alpha_{t+1}|\text{uniform}(0°, 360°)) \qquad [8]$$

The probabilities of each trial type occurring on trial t + 1, p(one-off$_{t+1}$), p(update$_{t+1}$), and p(expected$_{t+1}$) were simply set to the true probability of each trial type in the generative model [i.e., p(one-off$_{t+1}$) = 1/4, p(update$_{t+1}$) = 1/15, and p(expected$_{t+1}$) = 1 − (p(update$_{t+1}$) + p(one-off$_{t+1}$))].

Output of the model is illustrated in Figs. S9 and S10.

Information Theoretic Regressors. For the analyses presented in Fig. 3 (analysis of saccadic RTs), Fig. 4 (analysis of pupil area), and the fMRI whole-brain analysis and ROI analysis presented in Figs. 5 and 6, respectively, we used a GLM analysis with information theoretic regressors based on the learning model.

The four task regressors were main effect of task, entropy of the model, I_S, and D_{KL}. Each regressor used in fMRI analysis was defined as a series of 300 delta functions (events of 0.1 s in duration) at the times of target appearance (for 300 trials), weighted by task parameters and convolved with the hemodynamic response function. The weighting of each regressor on each trial was defined as follows.

All trials took value 1 as the main effect of task. Entropy of the model on trial t was expressed as

$$H(t) = \sum_{\alpha=1}^{360} p(\alpha_t \mid \alpha_{1:t-1}) \log\left(\frac{1}{p(\alpha_t \mid \alpha_{1:t-1})}\right), \qquad [9]$$

where $p(\alpha_{t+1}|\alpha_{1:t})$ is defined analogously as in Eq. **8**. The I_S of trial t is defined as

$$I_S(\alpha) = -\log p(\alpha_t|\alpha_{1:t-1}), \qquad [10]$$

where $p(\alpha_t|\alpha_{1:t-1})$ defined analogously as in Eq. **8**. The D_{KL} on trial t is defined as

$$D_{KL}(t) = \sum_{\alpha=1}^{360} p(\alpha_t|\alpha_{1:t-1}) \log \frac{p(\alpha_t|\alpha_{1:t-1})}{p(\alpha_t|\alpha_{1:t})}, \qquad [11]$$

where $p(\alpha_t|\alpha_{1:t-1})$ is defined analogously as in Eq. **8** and $p(\alpha_t|\alpha_{1:t})$ is defined analogously to the definition of $p(\alpha_{t+1}|\alpha_{1:t})$ in Eq. **8**:

$$p(\alpha_t|\alpha_{1:t}) = p(\text{expected}_t) \sum_{\mu_{t+1}} \sum_{\sigma_{t+1}} p(\alpha_t|\alpha_t \sim \mathcal{N}(\mu_t, \sigma_t)) p(\mu_t, \sigma_t|\alpha_{1:t})$$
$$+ p(\text{one}-\text{off}_t \cup \text{update}_t) p(\alpha_t|\text{uniform}(0°, 360°)). \qquad [12]$$

The output of the model and resulting regressors are illustrated in Figs. S9 and S10.

The shared variance between each pair of regressors was relatively low; the R^2 value for each pair was as follows: prior entropy and I_S/R^2 = 0.14; prior entropy and D_{KL}/R^2 = 0.14; and I_S and D_{KL}/R^2 = 0.15.

1. Friston K (2010) The free-energy principle: A unified brain theory? *Nat Rev Neurosci* 11(2):127–138.
2. Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B Biol Sci* 364(1521):1211–1221.
3. Posner MI, Snyder CRR, Davidson BJ (1980) Attention and the detection of signals. *J Exp Psychol* 109(2):160–174.
4. Posner MI, Walker JA, Friedrich FJ, Rafal RD (1984) Effects of parietal injury on covert orienting of attention. *J Neurosci* 4(7):1863–1874.

5. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
6. Baldi P, Itti L (2010) Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Netw* 23(5):649–666.
7. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, eds Black AH, Prokasy WF (Appleton–Century Crofts, New York), pp 64–99.

8. Knight FH (1921) *Risk, Uncertainty and Profit* (Hart, Schaffner & Marx, Boston, MA).
9. Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 10(7):294–300.
10. Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10(9):1214–1221.
11. Dayan P (2012) Twenty-five lessons from computational neuromodulation. *Neuron* 76(1):240–256.
12. Colby CL, Duhamel JR, Goldberg ME (1996) Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J Neurophysiol* 76(5):2841–2852.
13. Gold JI, Shadlen MN (2000) Representation of a perceptual decision in developing oculomotor commands. *Nature* 404(6776):390–394.
14. Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400(6741):233–238.
15. Yang T, Shadlen MN (2007) Probabilistic reasoning by neurons. *Nature* 447(7148): 1075–1080.
16. Milner A (1996) Neglect, extinction and the cortical streams of visual processing. *Parietal Lobe Contributions to Orientation in 3D Space*, Experimental Brain Research Series, eds Thier P, Karnath H-O (Springer, Berlin), Vol 25.
17. Milner A, Goodale M (1995) *The Visual Brain in Action* (Oxford Univ Press, Oxford).
18. Pierrot-Deseilligny C, Muri R (1996) Posterior parietal cortex control of saccades in man. *Parietal Lobe Contributions to Orientation in 3D Space*, Experimental Brain Research Series, eds Thier P, Karnath H-O (Springer, Berlin), Vol 25.
19. Mort DJ, et al. (2003) The anatomy of visual neglect. *Brain* 126(Pt 9):1986–1997.
20. Cox RT (1946) Probability, frequency and reasonable expectation. *American Journal of Physics* 14:1–13.
21. Vossel S, et al. (2013) Spatial Attention, Precision, and Bayesian Inference: A Study of Saccadic Response Speed. *Cereb Cortex*, 10.1093/cercor/bhs418.
22. Nassar MR, et al. (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat Neurosci* 15(7):1040–1046.
23. Preuschoff K, 't Hart BM, Einhäuser W (2011) Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Front Neurosci* 5:115.
24. Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu Rev Neurosci* 28:403–450.
25. Bouret S, Sara SJ (2005) Network reset: A simplified overarching theory of locus coeruleus noradrenaline function. *Trends Neurosci* 28(11):574–582.
26. Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46(4): 681–692.
27. Sara SJ (2009) The locus coeruleus and noradrenergic modulation of cognition. *Nat Rev Neurosci* 10(3):211–223.
28. Smith SM, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl 1):S208–S219.
29. Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M (2009) Adaptive coding of action values in the human rostral cingulate zone. *J Neurosci* 29(23):7489–7496.
30. Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9(7):940–947.
31. Mesulam MM, Nobre AC, Kim YH, Parrish TB, Gitelman DR (2001) Heterogeneity of cingulate contributions to spatial attention. *Neuroimage* 13(6 Pt 1):1065–1072.
32. Beckmann M, Johansen-Berg H, Rushworth MF (2009) Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci* 29(4):1175–1190.
33. Mars RB, et al. (2011) Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *J Neurosci* 31(11):4087–4100.
34. Rushworth MF, Behrens TE, Johansen-Berg H (2006) Connection patterns distinguish 3 regions of human parietal cortex. *Cereb Cortex* 16(10):1418–1430.
35. Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. *J Neurosci* 27(20):5326–5337.
36. Sereno MI, Pitzalis S, Martinez A (2001) Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science* 294(5545):1350–1354.
37. Gottlieb J, Snyder LH (2010) Spatial and non-spatial functions of the parietal cortex. *Curr Opin Neurobiol* 20(6):731–740.
38. Cavada C, Goldman-Rakic PS (1989) Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J Comp Neurol* 287(4):422–445.
39. Cavada C, Goldman-Rakic PS (1989) Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *J Comp Neurol* 287(4):393–421.
40. Caspers S, et al. (2008) The human inferior parietal lobule in stereotaxic space. *Brain Struct Funct* 212(6):481–495.
41. Molenberghs P, Mesulam MM, Peeters R, Vandenberghe RR (2007) Remapping attentional priorities: Differential contribution of superior parietal lobule and intraparietal sulcus. *Cereb Cortex* 17(11):2703–2712.
42. Vandenberghe R, Gitelman DR, Parrish TB, Mesulam MM (2001) Functional specificity of superior parietal mediation of spatial shifting. *Neuroimage* 14(3):661–673.
43. Molenberghs P, Gillebert CR, Peeters R, Vandenberghe R (2008) Convergence between lesion-symptom mapping and functional magnetic resonance imaging of spatially selective attention in the intact brain. *J Neurosci* 28(13):3359–3373.
44. Vandenberghe R, Molenberghs P, Gillebert CR (2012) Spatial attention deficits in humans: The critical role of superior compared to inferior parietal lesions. *Neuropsychologia* 50(6): 1092–1103.
45. Debener S, et al. (2005) Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J Neurosci* 25(50):11730–11737.
46. Holroyd CB, et al. (2004) Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat Neurosci* 7(5):497–498.
47. Thiel CM, Zilles K, Fink GR (2004) Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: An event-related fMRI study. *Neuroimage* 21(1): 318–328.
48. Mort DJ, et al. (2003) Differential cortical activation during voluntary and reflexive saccades in man. *Neuroimage* 18(2):231–246.
49. Hartwigsen G, et al. (2012) Left dorsal premotor cortex and supramarginal gyrus complement each other during rapid action reprogramming. *J Neurosci* 32(46): 16162–71a.
50. Mars RB, Piekema C, Coles MG, Hulstijn W, Toni I (2007) On the programming and reprogramming of actions. *Cereb Cortex* 17(12):2972–2979.
51. Rushworth MF, Ellison A, Walsh V (2001) Complementary localization and lateralization of orienting and motor attention. *Nat Neurosci* 4(6):656–661.
52. Rushworth MF, Paus T, Sipila PK (2001) Attention systems and the organization of the human parietal cortex. *J Neurosci* 21(14):5262–5271.
53. Gillebert CR, Mantini D, Peeters R, Dupont P, Vandenberghe R (2013) Cytoarchitectonic mapping of attentional selection and reorienting in parietal cortex. *Neuroimage* 67: 257–272.
54. Rushworth MF, Mars RB, Summerfield C (2009) General mechanisms for making decisions? *Curr Opin Neurobiol* 19(1):75–83.
55. Rushworth MF, Behrens TE (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* 11(4):389–397.
56. Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10(5):647–656.
57. Cohen MX, Ranganath C (2007) Reinforcement learning signals predict future decisions. *J Neurosci* 27(2):371–378.
58. Quilodran R, Rothé M, Procyk E (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57(2):314–325.
59. Karlsson MP, Tervo DG, Karpova AY (2012) Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* 338(6103):135–139.
60. Kolling N, Behrens TE, Mars RB, Rushworth MF (2012) Neural mechanisms of foraging. *Science* 336(6077):95–98.
61. Noonan MP, Mars RB, Rushworth MF (2011) Distinct roles of three frontal cortical areas in reward-guided behavior. *J Neurosci* 31(40):14399–14412.
62. Nobre AC, Coull JT, Frith CD, Mesulam MM (1999) Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat Neurosci* 2(1):11–12.
63. Dayan P, Kakade S, Montague PR (2000) Learning and selective attention. *Nat Neurosci* 3(11, Suppl):1218–1223.