Brain networks for confidence weighting and hierarchical inference during probabilistic learning

Florent Meyniel[a,1] and Stanislas Dehaene[a,b,1]

[a]Cognitive Neuroimaging Unit, NeuroSpin Center, Institute of Life Sciences Frédéric Joliot, Fundemental Research Division, Commissariat à l'Énergie Atomique et aux Énergies Alternatives, INSERM, Université Paris–Sud, Université Paris–Saclay, 91191 Gif/Yvette, France; and [b]Chair of Experimental Cognitive Psychology, Collège de France, 75005 Paris, France

Learning is difficult when the world fluctuates randomly and ceaselessly. Classical learning algorithms, such as the delta rule with constant learning rate, are not optimal. Mathematically, the optimal learning rule requires weighting prior knowledge and incoming evidence according to their respective reliabilities. This "confidence weighting" implies the maintenance of an accurate estimate of the reliability of what has been learned. Here, using fMRI and an ideal-observer analysis, we demonstrate that the brain's learning algorithm relies on confidence weighting. While in the fMRI scanner, human adults attempted to learn the transition probabilities underlying an auditory or visual sequence, and reported their confidence in those estimates. They knew that these transition probabilities could change simultaneously at unpredicted moments, and therefore that the learning problem was inherently hierarchical. Subjective confidence reports tightly followed the predictions derived from the ideal observer. In particular, subjects managed to attach distinct levels of confidence to each learned transition probability, as required by Bayes-optimal inference. Distinct brain areas tracked the likelihood of new observations given current predictions, and the confidence in those predictions. Both signals were combined in the right inferior frontal gyrus, where they operated in agreement with the confidence-weighting model. This brain region also presented signatures of a hierarchical process that disentangles distinct sources of uncertainty. Together, our results provide evidence that the sense of confidence is an essential ingredient of probabilistic learning in the human brain, and that the right inferior frontal gyrus hosts a confidence-based statistical learning algorithm for auditory and visual sequences.

confidence | functional MRI | learning | probabilistic inference | Bayesian

The sensory data that we receive from our environment are often captured by temporal regularities—for instance the colors of traffic lights change according to a predictable green–yellow–red pattern; thunder is often followed by rain, etc. Knowledge of those hidden regularities is often acquired through learning, by aggregating successive observations into summary estimates (e.g., the probability of the light turning red when it is currently yellow). When sensory data are received sequentially, learning can be described as an iterative process that updates the internal estimates each time a new observation is received. Learners must therefore constantly balance two sources of information: their current estimates and the new incoming observations.

Any learning algorithm must find a solution to this balancing act. Finding the correct balance is especially critical in a world that is both stochastic and changing (1), i.e., where observations are governed by probabilities that can change over time (a situation called volatility). An excessive reliance on incoming observations will make the learned estimates dominated by fluctuations instead of converging to the true underlying probabilities. Conversely, an excessive reliance on the previously acquired knowledge will slow down the learning process and impede a quick reset of the internal estimates, which is useful when the environment changes.

In the 1950s, learning algorithms were proposed to solve this weighting problem optimally under specific conditions (e.g., the Kalman filters and subsequent developments; ref. 2). A general

and normative solution to this problem requires weighting each source of information according to its reliability (3–12). According to this Bayes-optimal solution, any discrepancy between a new observation and a learned estimate should lead to an update of this internal estimate, but the size of this update should decrease as the prior confidence in this internal estimate increases. Furthermore, this prior confidence should depend on two factors: the precision of the current internal estimate and a discounting factor that takes into account the possibility that a change occurred. Indeed, a change in environmental parameters would render the current estimate useless for predicting future observations. An optimal algorithm should therefore maintain distinct types of uncertainty, organized in a hierarchical manner: future events are uncertain because they are governed by probabilities (level 1); these probabilities themselves are known only with a certain precision (level 2); and this precision is limited because a change may have occurred (level 3).

In summary, efficient learning requires the learner to maintain and constantly update an accurate representation of the confidence in what has been learned. In a previous behavioral study (13), by engaging human adults in a probability-learning task, we showed that they possess a sense of confidence in what has been learned that is remarkably similar to the optimal algorithm. Here, we propose that learning approaches optimality in humans because it shares two features of the optimal algorithm: (*i*) it relies on a sense of confidence that serves as a weighting factor to balance prior estimates and new observations; and (*ii*) confidence is organized hierarchically, taking into account higher-order factors such as volatility. We aim to provide behavioral and functional magnetic resonance imaging (fMRI) evidence on

the brain mechanisms that implement such a hierarchical confidence-weighted learning.

We submitted human subjects to a learning task that requires the tracking of transition probabilities (i.e., the conditional probability of observing the current item, given the identity of the immediately previous item), knowing that those probabilities fluctuate randomly over time (e.g., Fig. 1A). This task is similar to previous studies (1, 14–21), but with several additional features. First, although many studies resort to reinforcement-learning tasks that evaluated the learned probabilities indirectly and implicitly (1, 17, 19), we opted
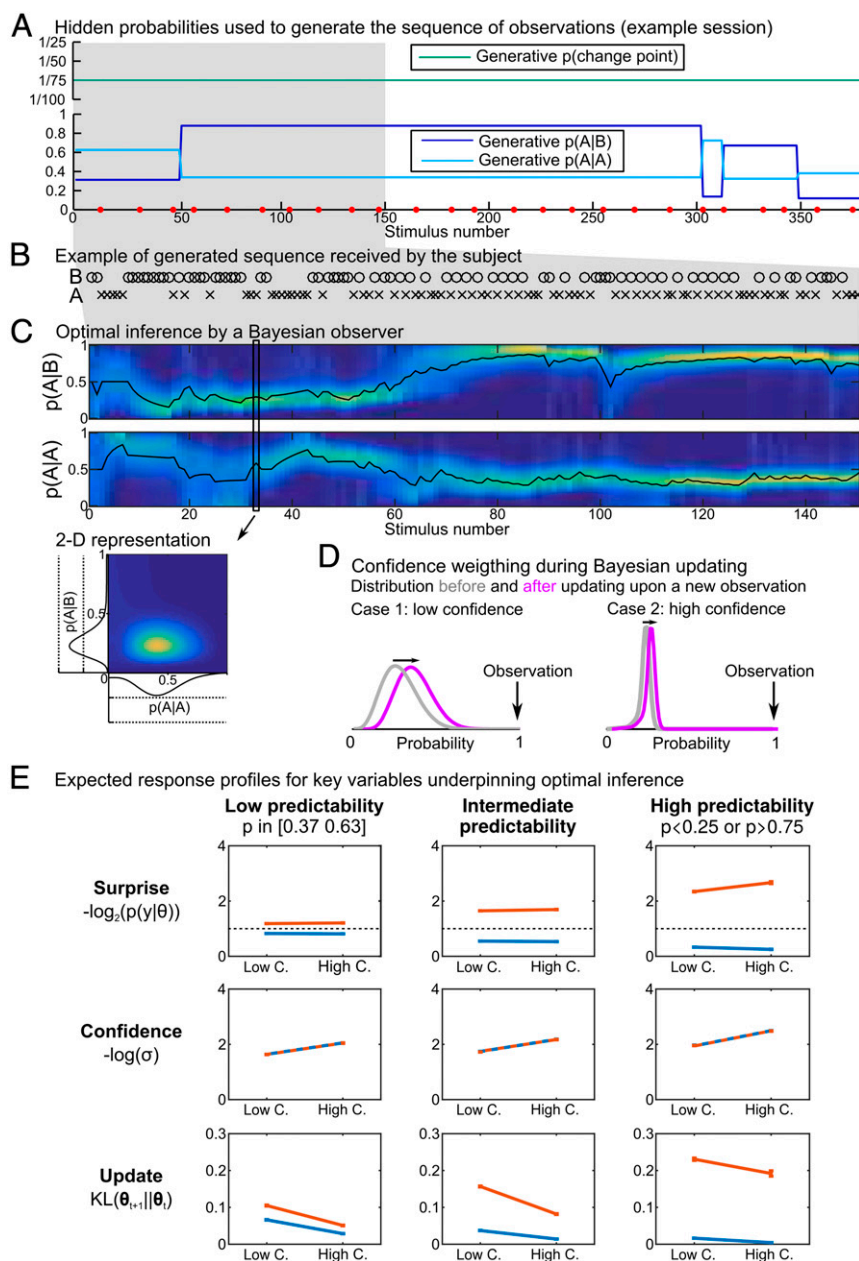

Fig. 1. A normative role for confidence during probabilistic inference. (*A*) Subjects were asked to continuously estimate the probabilities $p(A|B)$ and $p(A|A)$ governing the successions between two stimuli, while these probabilities underwent occasional stepwise changes. The probability of change point was held constant, but transition probabilities and change times varied randomly between sessions and participants. Subjective confidence levels were probed occasionally at trials marked by red marks. (*B*) Example sequence. Actual A and B stimuli corresponded to clearly distinguishable pairs of sounds or visual symbols, presented in distinct auditory and visual sessions. (*C*) The likelihood distribution of transition probabilities were estimated by a Bayesian observer, based on the observations received so far and the generative model. Likelihood distributions are presented as slices as a function of time. Each slice corresponds to a marginal distribution from the actual 2D space defined by combinations of $p(A|B)$ and $p(A|A)$. Confidence in the estimated transition probabilities corresponds to the (negative log) SD of those marginal distributions. (*D*) Bayesian theory predicts that confidence should serve as a weighting factor that modulates the incoming evidence during optimal inference: the update induced by a new observation should be less pronounced when confidence in the prior estimation is high. (*E*) Values of three key variables of Bayesian inference were sorted into bins and averaged across all observations presented to subjects. Notations: y is the sequence of observations received, θ the estimated transition probability, σ its SD, and KL (Kullback–Leibler divergence) quantifies the distance between two probability distributions, high/low C stands for high/low confidence. Blue/orange corresponds to expected (likelihood >0.5)/unexpected (likelihood <0.5) outcomes. Note that confidence in the current transition does not depend on the eventual outcome because it is not yet observed (the blue/orange curves overlap perfectly).

for an explicit statistical learning task in which participants overtly reported on a numerical scale their probability estimates and their subjective confidence levels in those estimates.

Second, we did not ask for a behavioral response on each trial, which could have perturbed the continuity of the learning process and interfered with fMRI measurements. Instead, we interrupted the stimulus sequence only occasionally to ask for subjective reports. We developed a mathematical model that solved the learning task optimally (an "ideal observer") to quantify the trial-by-trial fluctuations in the hidden inference process as a function of the observed sequence of stimuli (e.g., Fig. 1B). We derived quantitative and qualitative diagnostic predictions for the brain signatures of an optimal learning algorithm. We examined whether brain signals, measured between behavioral reports, conformed to the predictions of an ideal observer. We also checked that the subjects' behavior closely resembled the ideal observer's.

Third, we added a hierarchical component to our task by requiring the estimation of two distinct transition probabilities that changed suddenly and simultaneously, at unpredictable times. Subjects were fully informed about these properties of changes. This feature allowed us to probe the brain's ability to entertain a hierarchical model that requires the monitoring of two distinct but interdependent confidence levels, attached to each transition probability. In particular, we examined whether the confidence in both parameters collapsed simultaneously when it was likely that a change in probabilities had occurred. Such a finding would support the hierarchical model.

Results

Behavior During the Task.
During each fMRI session, subjects were presented with a sequence of two arbitrary stimuli, termed A and B. In distinct sessions, those stimuli could be either auditory or visual, thus allowing us to probe whether a given brain region operates in a modality-specific manner or in an abstract amodal manner. Sequences of stimuli A and B were generated according to a 2×2 transition-probability matrix that remained constant only for a limited period (Fig. 1 A and B). The entire matrix (i.e., two independent transition probabilities) was resampled whenever a change occurred. Subjects were fully informed about the task structure. They were asked to constantly keep track of the transition probabilities that were used to generate the observed sequence. During a training session, they were asked to detect occasional sudden changes in probabilities. Furthermore, every 15 stimuli on average, they evaluated, with a cursor, the probability of the next stimulus given the identity of the previous one, and their confidence about this estimate (Fig. S1).

In a previous behavioral study (13), we provided a detailed comparison between subjects' responses and a mathematical model of an ideal observer performing the same task. This study revealed a tight parallel between subjective and ideal estimates. Here, after one training session with this behavioral task design, fMRI signals were acquired while subjects performed four sessions of a trimmed-down and almost fully covert version of this learning task. In this fMRI version, the reports of changes and probability estimates were omitted, and subjects only occasionally reported, on a four-level scale, their confidence in their covert probability estimate at the moment of the question. In addition to keeping behavioral reports to a minimum, this method allowed us to verify that subjects engaged in the task and that their subjective confidence conformed to the normative model. As in our previous study (13), subjective confidence levels correlated linearly with the optimal confidence levels (linear regression, $t_{20} = 6.40$, $P = 3 \times 10^{-6}$). When optimal confidence in the relevant transition probability (the one corresponding to the stimulus preceding the question) and in the irrelevant one were both entered in the same regression to model subjective confidence levels, the regression weight was significantly higher for the relevant transition probability than for the irrelevant one

(paired difference, $t_{20} = 7.16$, $P = 6 \times 10^{-7}$). This indicates that subjects selectively monitored and reported the confidence attached to each transition probability.

We also reanalyzed the data from our previous behavioral study (13) to evidence the role of online confidence weighting during learning in this task. We compared the ideal-observer model, which implements a dynamic confidence weighting, with the delta-rule model with fixed learning rate (22), which implements a fixed weighting of the incoming information and is devoid of any representation of confidence. For the learning rate, we used the value that minimized the mean squared error (MSE) between the delta-rule estimates and the actual generative probabilities for the sequences presented to subjects. Nevertheless, subjects' estimates of probabilities were better captured by the ideal observer than by this optimized delta-rule model (difference in MSE across subjects: $t_{17} = 5.8$, $P = 2.0 \times 10^{-5}$). We also performed a Bayesian model comparison of the ideal observer and a delta rule whose optimal learning rate was separately fitted for each subject, instead of being fixed to the optimal value. This comparison revealed a 19:1 posterior probability ratio in favor of the hypothesis that the ideal observer is the more likely model of the subjects' probability estimates. Although the exact algorithm underlying human probabilistic learning remains unknown, our behavioral results indicate that this algorithm must implement some form of confidence weighting. In the following, we therefore used the ideal observer as a principled description that closely approximates the subjects' inference algorithm.

Predictions of the Ideal-Observer Model.
Given the hierarchical nature of our task, the ideal observer's inference must unfold over several levels (Fig. S2A). Starting from the observed sequence (level 1), it estimates the current transition probabilities underlying observations (level 2), knowing that those transition probabilities undergo stepwise changes from one trial to the next with a fixed probability that must itself be estimated (the "volatility," level 3).

Volatility could itself change with time, potentially adding a level 4 to our hierarchical design—and indeed, previous research has suggested that the human brain and even rodents can track fluctuations in volatility (1, 23). However, such fluctuations in volatility do not contribute to the present task. Volatility was held constant and furthermore, the ideal-observer model, when inferring volatility, quickly converged to a value very close to the constant generative value within a single training session, even in the absence of any prior information (Fig. S2B). Because subjects benefited from even more information (they were instructed that changes were rare), the ideal-observer analysis suggests that, in the present task, unlike in previous work (1), the assumption of fluctuations in volatility (level 4) was not required. In the remnant of this paper, volatility (level 3) can therefore be considered constant.

This constant, nonzero volatility is nevertheless a crucial factor in our experimental design: it is because changes occur at unpredictable times that a confidence-weighted learning algorithm must constantly endeavor to adjust the weights of incoming observations. We concentrate on the level at which the ideal observer predicts that inference should produce dynamic quantities (Fig. 1B): the two unknown transition probabilities $p(B|A)$ and $p(A|B)$ (level 2). These two probabilities suffice to characterize the 2×2 transition probability matrix, because $p(A|A) = 1 - p(B|A)$ and $p(B|B) = 1 - p(A|B)$.

A Normative Dissection of the Inference Process.
The ideal-observer analysis allows the key ingredients of the optimal confidence-weighted inference to be summarized by several unambiguous mathematical quantities (Fig. 1E). The confidence in the estimated statistic relates to the width of the estimated posterior distribution, which can be summarized as the negative log SD (12, 13, 24). We adopted a log space because it is the natural
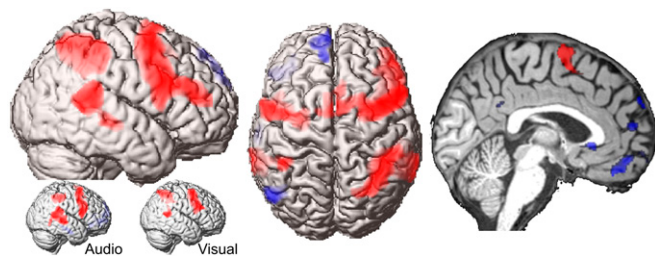
Fig. 2. Brain regions whose signals correlated with the internal updates predicted by the ideal-observer model. The statistical maps show the group-level significance of positive (red) and negative (blue) regression coefficients, when fMRI activity was regressed at every trial and all sessions on the optimal (Bayesian) amount of estimation update (GLM1). *Inset* shows the results of the analysis restricted to audio and visual sessions (see also Fig. S3). Maps are thresholded at the voxel level ($P < 0.001$ uncorrected) and the cluster level ($P < 0.05$ FWE).

space for variance and SD parameters (which are simply proportional in log space) (25); and because we previously found that subjective confidence relates linearly to the log SD in this task (13). However, this choice does not affect our conclusions. Following information theory, the surprise elicited by a given observation corresponds to the negative log likelihood of this observation (14, 18, 26, 27). The unpredictability of the outcome is formalized by the notion of entropy (the expected surprise). Finally, the magnitude of the update is the distance between the distributions estimated before and after each observation, which is technically a Kullback–Leibler divergence (18, 27). Fig. 1*E* summarizes the expected profiles of those distinct quantities during the task.

To probe the brain networks involved in probabilistic inference, we first regressed the fMRI time series on the optimal values of estimation update. Fig. 2 shows the whole-brain statistical maps [general linear model 1 (GLM1); *Materials and Methods*], with local maxima in the intraparietal sulcus, the posterior superior temporal sulcus, the inferior frontal gyrus, the frontal eye field, the supplementary motor area, which were observed irrespective of the sensory modality used in the task (Tables S1 and S2). However, optimal confidence-weighted updates depend on two factors: they are larger when surprise is higher and when confidence is lower (Fig. 1*E*). Therefore, positive correlations of brain signals with the optimal estimation update could reflect the update process itself or alternatively, one of those factors (confidence or surprise). To disentangle the variables underlying the inference process, we adopted a factorial analysis (GLM4): we first stratified trials by predictability levels, and we then further stratified by confidence and surprise levels. As shown in Fig. 1*E*, this sorting results in profiles for surprise, confidence, and update that were clearly distinguishable from one another in the ideal observer. The results of the factorial analysis are presented below. In Fig. S4 we report an alternative analysis, a multiple regression including confidence and surprise in the same model, which largely replicates the results of the factorial analysis.

Brain Correlates of Confidence. Using the factorial strategy, we looked for a main effect of confidence while controlling for the levels of predictability and surprise in GLM4. We found several clusters, notably in the right intraparietal sulcus [x, y, z: (32, −68, 59)] and the right inferior temporal gyrus [x, y, z: (56, −46, −14)], see Fig. 3*A* and Table S3. Plotting the fMRI signals in these regions for each category level (Fig. 3*B*) confirmed that they
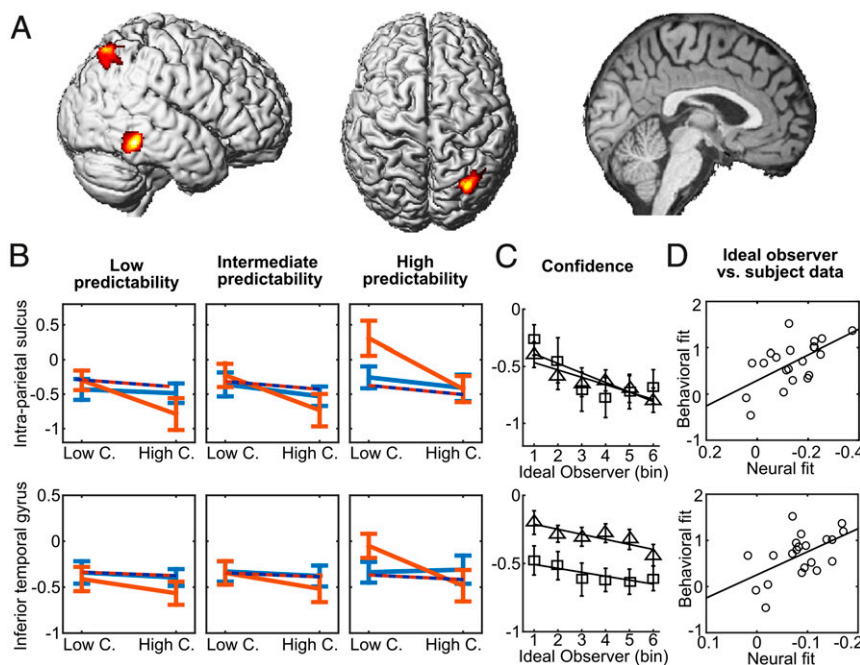

Fig. 3. Cortical correlates of confidence. (*A*) Main effect of confidence in an ANOVA also controlling for surprise and predictability levels (GLM 4). Maps are thresholded at the voxel level ($P < 0.001$ uncorrected) and the cluster level ($P < 0.05$ FWE). (*B*) The fMRI signals (plain lines) are plotted following the categorical approach presented in Fig. 1*E*. Signals were extracted with cross-validation: voxels were identified from ANOVA in one session type (auditory or visual) and tested in the other (*Materials and Methods*). The plot shows the average of both cross-validated extractions. To facilitate visual comparison, optimal confidence levels (dashed lines) are overlaid after adjusting for offset and scaling for each cluster (and not each subplot). Blue vs. orange correspond to expected vs. unexpected outcomes, following the ideal observer's estimates. (*C*) Cross-validated variations in fMRI signals as a function of the theoretical confidence predicted by the ideal-observer model (binned into 6 percentiles) (GLM5). Squares correspond to the signal extracted in auditory sessions, triangles to visual sessions. Fitted lines correspond to the average of fitted individual data. Error bars are SEM across subjects (*B* and *C*). (*D*) Interindividual variations in neural signals predict interindividual variations in behavior: A significant between-subjects correlation was observed between the neural fit (regression coefficients between fMRI signals and ideal-observer confidence levels—GLM2) and behavioral fit (regression coefficients between subjective confidence reports and ideal-observer confidence levels).

tracked the level of confidence. To avoid the frequent circularity inherent in plotting fMRI signals from selected voxels (28), we used a cross-validation strategy across sensory modalities. Within broadly defined anatomical regions of interest, we first selected the voxels ($n = 100$) showing the strongest effect of confidence when the ANOVA (GLM4) was restricted to the auditory sessions, and then extracted and plotted the signals from the same voxels in the independent data from the visual sessions. Doing the converse yielded similar results, and we report the average values in Fig. 3B. Plotting fMRI signals in percentiles of the predicted confidence levels (GLM5; Fig. 3C) showed that, in all of the above regions and in both modalities, activation decreased essentially monotonically as confidence increased. To quantify whether these effects of confidence were modality independent, we regressed activity in suprathreshold voxels on optimal confidence levels in auditory and visual sessions separately, and we tested for their equality with a Bayesian t test (29). The effect was similar across modalities in the intraparietal sulcus [Bayes factor (BF): 4.3] and to a lesser extent in inferior temporal gyrus (BF: 2.8).

We also examined whether interindividual differences in activity in these regions predicted interindividual differences in behavior. Because subjective confidence levels were occasionally sampled from subjects, we tested whether interindividual differences in the tightness of the fit between those fMRI signals and optimal confidence levels, measured in between the behavioral reports (GLM2), predicted interindividual behavioral differences in the tightness of the fit between subjective confidence ratings and optimal confidence levels at the moment of reports. We found significant correlations (intraparietal sulcus: $\rho_{21} = -0.60$, $P = 0.004$; inferior temporal gyrus: $\rho_{21} = -0.55$, $P = 0.010$). These correlations were negative because fMRI signals decreased with confidence, i.e., increased with uncertainty. Thus, the subjects whose brain activation best corresponded to the ideal-observer predictions were also those whose subjective confidence reports best tracked the optimal values.
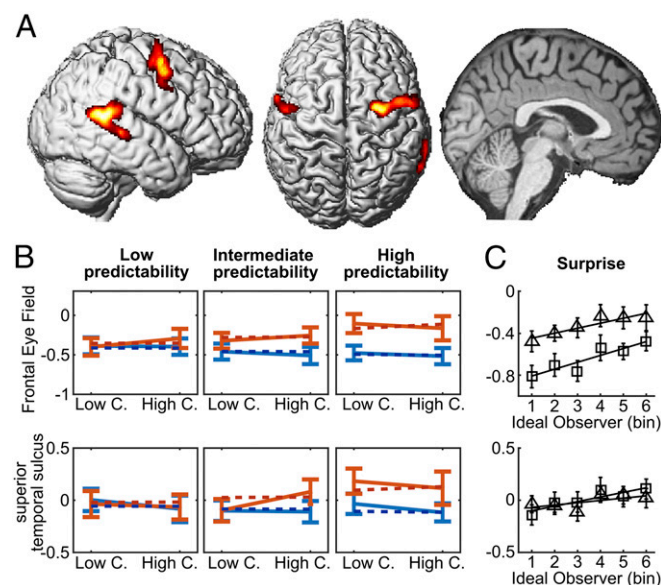

Fig. 4. Cortical correlates of surprise. (A) Main effect of surprise, in an ANOVA also controlling for confidence and predictability levels (GLM4). Maps are thresholded at the voxel level ($P < 0.001$ uncorrected) and the cluster level ($P < 0.05$ FWE). In B, dashed lines represent the optimal surprise signal and in C, bins correspond to 6 percentiles of the optimal surprise levels (GLM6). Same format as Fig. 3 (including cross-validation for B and C).

Brain Correlates of Surprise. Similarly, we looked for a main effect of surprise, while controlling for predictability and confidence levels in GLM4. Surprise signals were found in several clusters, notably in the right superior temporal sulcus [x, y, z: (50, −31, 0)] and the frontal eye field [on the right, x, y, z: (30, −6, 54), and left, x, y, z: (−60, 5, 35)], see Fig. 4A and Table S4. Using cross-validation between modalities and plotting fMRI signals for each category level (Fig. 4A) revealed that they conformed to surprise signals. We checked that variations of cross-validated fMRI signals in these regions parametrically followed the optimal surprise levels by plotting fMRI signals against percentiles of surprise (GLM6; Fig. 4C). Regression coefficients for optimal surprise levels (GLM3) were all positive, and coefficients were significantly similar in auditory and visual sessions in right frontal eye field (respective BF: 3.6) and to a lesser extent in the right superior temporal sulcus (BF: 1.6).

Brain Correlates of Estimation Update, Reflecting Confidence-Weighted Surprise. The central property of a confidence-weighted updating mechanism is that it should reflect both confidence and surprise. The theoretical analysis suggests that the combination may be nearly additive (Fig. 1E), and we therefore searched the whole brain for a conjunction of the main effects of optimal confidence and optimal surprise (GLM4). This conjunction was maximally significant in the right inferior frontal gyrus [x, y, z: (44, 6, 46), both main effects were significant at $P = 0.001$, uncorrected]; see Fig. 5A. Using cross-validation between modalities to plot the inferior frontal gyrus response in each cell of the factorial analysis (Fig. 5B) revealed that it conformed to an update signal. We checked that variations of cross-validated fMRI signals in this region parametrically followed the optimal update levels by plotting fMRI signals against percentiles of update (GLM7; Fig. 5C). Regression coefficients were positive and significantly similar in auditory and visual sessions (BF: 4.2). Overall, these results indicate that the right inferior frontal gyrus conforms to an amodal representation of the amount of update needed to optimally revise the internal model of learned transition probabilities, given the latest sensory observation received.

We also reasoned that, if the inferior frontal gyrus integrates both confidence and surprise, it may have significant functional connectivity with the brain regions whose activity reflects either confidence or surprise. We therefore looked for regions in which the fMRI signals correlated simultaneously with those of seed regions corresponding to the best confidence and surprise signals reported above. For confidence, signals were taken from the intraparietal sulcus and for surprise, from the frontal eye field (we discarded the superior temporal sulcus and inferior temporal gyrus because they showed only moderate evidence of an amodal response). Conjunction analysis of functional connectivity from the intraparietal sulcus and frontal eye field seeds (GLM8) was maximally significant in a brain site located in the right inferior and middle frontal gyrus [x, y, z: (32, 6, 59), at $p_{FWE} < 0.05$], see Fig. 5D and Table S5.

The Estimation Update Reflects a Hierarchical Inference. The use of two distinct transition probabilities in our experiment allowed us to test specific predictions of the hierarchical inference process, which is thought to underlie the update. We tested those predictions on inferior frontal gyrus signals because this region was identified above as a putative "update region." We used the inferior frontal gyrus cluster (that extends in the middle gyrus) identified with functional connectivity analysis (GLM8) rather than a direct contrast of task factors to avoid circularity. However, the results were similar when using the conjunction of contrasts reported above.

A first prediction is that, because the task involves two transition probabilities, $p(A|B)$ and $p(B|A)$, two confidence levels have to be constantly monitored. However, their relevance for
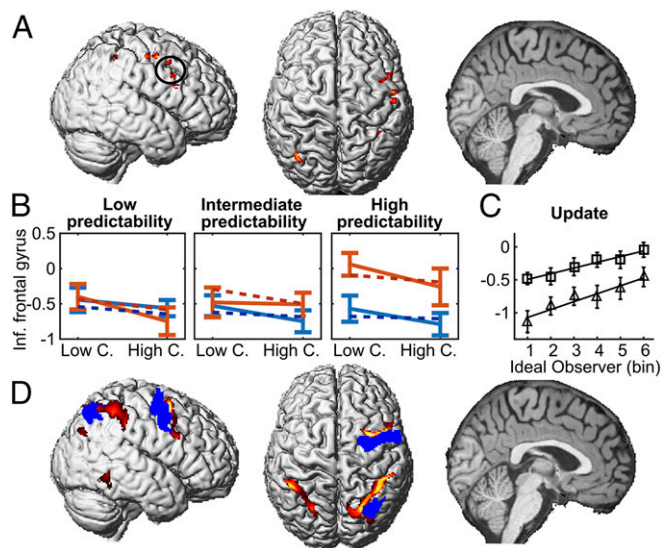
Fig. 5. Cortical correlates of confidence-weighted update. (*A*) Conjunction of the main effects of surprise and confidence in an ANOVA also controlling for predictability levels (GLM4), shown at $P < 0.005$ uncorrected. In *B* dashed lines represent the optimal update signal and in *C*, bins correspond to 6 percentiles of optimal update levels (GLM7), same format as Fig. 3 (including cross-validation). (*D*) Conjunction of functional connectivity with the intraparietal sulcus (signaling confidence; Fig. 3*A*) and the frontal eye field (signaling surprise; Fig. 4*A*) [results are shown at a voxel level $P < 0.05$ FWE (GLM8)]. The two seed clusters are shown in blue.

learning changes across trials, depending on the identity of the preceding stimulus. Only the confidence in the transition probability corresponding to the preceding stimulus should be used to weight the incoming evidence. To test this first prediction, we examined whether fMRI signals in the inferior frontal gyrus correlated specifically with the confidence that was relevant on the current trial. Brain activity conformed to those expectations: when confidence levels in the relevant and irrelevant transitions probabilities were both included in the fMRI model (GLM9), significant coefficients were found only for the relevant transition ($P = 2 \times 10^{-7}$) but not for the irrelevant one ($P = 0.14$), and this dissociation was significant ($P = 0.0016$).

A second prediction is that, although the optimal confidence levels attached the two transition types probabilities are largely independent, they should show parallel fluctuations at very specific moments in the task; namely, whenever a change is suspected. Changes play such an overarching role in our design because, whenever a change occurs, both transition probabilities are simultaneously reset to new values (and participants were explicitly informed of this fact). Therefore, given only evidence in favor of a change in one transition probability, one can infer that the other transition probability is also likely to have changed. For instance, when an observed succession (e.g., A→A) is so surprising as to arouse a suspicion of a change, confidence in both this transition probability $p(A|A)$ and the other transition probability $p(A|B)$ should be reduced, even if the other type of succession (from B) was not observed. On the contrary, when there is little surprise and hence no evidence for a global change, confidence in the transition probability corresponding to the current stimulus can increase and leave the estimate of the other transition probability essentially unchanged.

Streaks of repeated A's or B's (AAA... or BBB...) offer an opportunity to test this normative property. When such a streak is finally discontinued, the succession type that was not observed during the streak becomes relevant for learning again, and the confidence attached to it should therefore be seen in the inferior

frontal gyrus. The model predicts that the confidence associated with this transition probability should be lower after the streak than before it, but only when the streak aroused a suspicion of a change (Fig. 6*A*). We therefore screened the sequences that were presented to subjects for streaks of three or more repeated elements, and sorted into those in which optimal confidence levels in the observed succession type increased steadily from one observation to the next (thus plausibly licensing the inference that no change occurred) and those in which it decreased at least once (thus licensing the inference that a change occurred). In the ideal observer, confidence levels in the succession type not observed during the streak dropped as explained above when a change was suspected (post/pre streak change: -0.199 ± 0.011 in log-SD unit), otherwise they remained stable (difference in post/pre changes between the two streak types: -0.12 ± 0.02). As predicted, we observed a similar dissociation in the inferior frontal gyrus (GLM10), see Fig. 6*B*: the fMRI signals measured before and after the streaks were significantly different for streaks within which the optimal confidence levels decreased (post/pre streak change: 0.64 ± 0.35 in arbitrary fMRI signals, $P = 0.04$); and this difference was much larger than the one observed for the other streak type (difference in post/pre changes between the two streak types: 1.00 ± 0.55, $P = 0.04$). Thus, the inferior/middle frontal gyrus signal accurately reflects the subtle changes in confidence that are predicted by a hierarchical ideal-observer model.

Discussion

Confidence weighting of surprise signals is a normative property of Bayesian learning algorithms. Our results provide evidence for confidence weighting of surprise in the human brain, independently of the sensory modality tested, and human subjects were engaged in the covert estimation of the time-varying probabilities generating
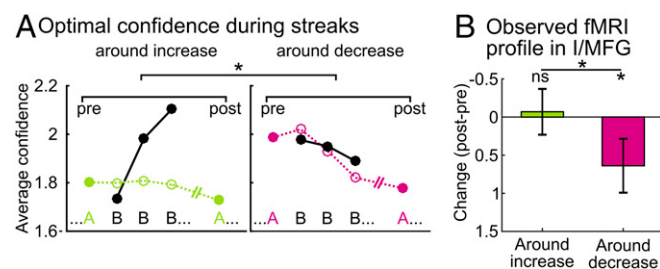

Fig. 6. Evidence for a hierarchical representation of confidence. (*A*) In the ideal-observer model, two transition probabilities are simultaneously monitored, each with its own confidence level (Fig. 1*C*). However, changes in the generative process are global, impacting both transition probabilities at once. Therefore, when a change is suspected, confidence in all transition probabilities should drop, even if this suspicion arises from the observation of a single succession type. To provide evidence of this effect with the actual sequences presented to subjects, optimal confidence levels were averaged during streaks with three or more repetitions (ABBBA, BAAAB, ABBBBA, etc.; note that A and B play symmetric roles). The plot shows confidence in the transition probability that is relevant during the streak (black) and confidence in the other transition probability (colors), separately for streaks within which confidence increases (no change is suspected, green) and within the others (in which a change is suspected, purple). When confidence in the observed succession type increases, the confidence for the unobserved succession type remains stable, but when confidence in the observed succession type decreases, the other one also drops: there is an interaction between streak type and the post- versus pre-streak optimal confidence levels. (*B*) The predicted interaction was observed in the inferior–middle frontal gyrus (I/MFG). The plot shows the difference in fMRI responses on trials that preceded and followed a streak of repeated items (filled, colored circles in *A*). Note that the activity in this region relates negatively to confidence, we therefore reverted the y axis to facilitate the visual comparison with the ideal observer. Error bars correspond to SEM across subjects; they were extremely small in *A* and thus omitted. *$P < 0.05$.

sequences of stimuli. The areas observed included the inferior frontal gyrus, intraparietal sulcus, and frontal eye field. By focusing on particular trials, we showed that each region exhibited a unique profile that was characteristic of either confidence, surprise, or update signals. We also showed that update signals in the inferior frontal gyrus conform to a hierarchical inference.

A Computational Definition of Confidence. Much confusion still surrounds the formalization of subjective confidence (12, 30, 31). In studies that investigate decision confidence (32–35), confidence formally corresponds to the probability of the decision being correct. Here, however, we studied confidence in the context of a learning task, where participants infer the value of hidden variables and use them to predict future outcomes. In this case, confidence in the continuous variable that is learned should be captured by the estimated SD of this variable, and confidence in the outcome (the upcoming stimulus) should be captured by the estimated probability of the outcome. It is essential to avoid conflating those two notions because they are normatively distinct components of the updating process.

Here, confidence depends on the evidence conveyed by observations, which can be quantified by an ideal observer: this origin is "external" (36). We did not explore "internal" or subject-specific fluctuations in confidence that may arise, for instance, due to imperfect computations or to temporary distraction. However, confidence should produce the same effect irrespective of its origins: in either case, lower confidence should trigger more learning from new observations. We identified the brain mechanisms of confidence by correlating brain activity with the levels of confidence predicted by the ideal observer. It might have been more efficient to collect trial-by-trial reports of subjective confidence, and use them as predictors of brain activity. In practice, however, frequent reports could also perturb the fMRI signal and hinder the inference process itself. Instead, we showed here and in our previous study (13) that the optimal externally driven confidence is a significant determinant of subjective confidence in our task. Thus, we used the ideal-observer algorithm as a starting point for modeling brain activity common to all subjects. Future work should investigate deviations from optimality and interindividual differences.

The ideal-observer approach also affords a theory-driven mapping of computational variables onto brain signals, as in previous studies (1, 14, 16–18, 21). Such model-based approach typically involves regressing the brain activity on explanatory variables. We acknowledge that this regression approach may not fully capture all of the brain signals involved in coding for the underlying computational variables (37), a problem that is aggravated by the low spatial and temporal resolution of fMRI. We therefore complemented standard regression analyses with a principled approach (38) in which we checked the ordering of fMRI signals in bins of predicted values (Figs. 3C, 4C, and 5C) and also designed categorical contrasts focusing on particular trials to identify response profiles associated with each computational variable (Figs. 3B, 4B, and 5B). We also tested the predictions of hierarchical Bayesian inference in the form of predicted differences and interaction (Fig. 6B).

The Role of Confidence in Learning. Confidence not only has a clear definition here, but also a precise computational role: it should control the weight of the incoming evidence, i.e., how much is learned from a new observation. Past research on the brain's learning algorithms has primarily focused on how surprising observations, i.e., prediction errors, drive the learning process (22, 39). It is only more recently that researchers have realized that the learning rate should also depend on confidence (10, 40).

The fact that the human learning algorithm includes an adjustable learning rate was first demonstrated in the context in which such an adjustment is most crucial, i.e., in unstable environments (1, 15, 17, 19, 20, 41, 42). For instance, Behrens and

colleagues (1) showed that the apparent learning rate increases with environmental volatility, i.e., when changes in generative characteristics are more frequent. This effect was paralleled by stronger activity in the anterior cingulate cortex. Indeed, a full hierarchical Bayesian analysis shows that, in the ideal observer, higher estimates of volatility decrease the confidence in current estimates (Fig. S2; Eq. 1). However, the present paper kept volatility constant and capitalized on a different effect: confidence should also decrease whenever an environmental change is suspected, even if the frequency of such change (i.e., volatility) is kept constant (19, 20). Accordingly, several studies have shown that a drop in confidence boosts learning (15, 20) and may even reset the learning process altogether (43, 44). The present study presents a detailed analysis of the brain mechanisms underlying this rational behavior, and leads to the conclusion that the human brain closely approximates the confidence-weighted hierarchical learning algorithm, which is optimal in the present circumstances. Because our results demonstrate a strong influence of confidence on brain activity and learning rates, they are incompatible with two general classes of alternative models of the learning process: classical learning algorithms with fixed learning rate, such as the Rescorla–Wagner or delta-rule models (which cannot account for the modulation of surprise signals by confidence), and nonhierarchical learning models (which cannot account for the overarching effect of change detection on both transition probabilities). Our findings do not preclude that the brain may resort to these simpler alternatives in different situations, or within specific brain circuits. However, they do show that the human brain performs better than these classical learning algorithms predict, and indeed makes near-optimal use of all of the available evidence when updating its internal model. An important issue for future research is whether such high-level performance is typical only of the adult educated human brain performing an explicit learning task, as studied in the present work, or whether confidence-weighted learning may also be observed during implicit tasks, or in nonverbal organisms.

Implementation of Confidence Weighting in the Brain. Our results suggest a mechanism by which adjustable learning rates may be implemented in the brain. The brain appears to independently track (*i*) the discrepancy between observations and predictions, as manifested by pure surprise signals in the frontal eye field and in the sensory cortices; and (*ii*) the confidence in these probabilistic forecasts, as manifested by pure confidence signals in the intraparietal sulcus. Both signals could then be combined into a confidence-weighted surprise signal in the inferior frontal gyrus. Anatomical segregation of different learning signals has already been reported for simple prediction errors and weighted prediction errors (45). Note that we focused primarily on the most significant loci for surprise, confidence and update signals, but our results suggest the existence of distributed brain networks interacting during learning.

In our view, confidence may serve as a gate for incoming information. This mechanism is similar, in computational terms, to the role traditionally ascribed to selective attention in the regulation of learning. Following an earlier proposal by Dayan et al. (4), information gating by a frontoparietal network could explain why this network is found here during learning but also in complex problem solving both in human and nonhuman primates (46, 47) and in visuospatial attention (48). Indeed, these tasks involve a similar notion of filtering that may be implemented by frontoparietal networks: at any given moment, some stimuli, features, or thoughts, provided either simultaneously or sequentially, are selected and given more weight for further processing.

It is interesting to note that the clusters identified here related to confidence were mostly lateralized to the right hemisphere. This finding accords well with the known lateralization of the attention system (48). Such a lateralization is also reported in the metacognition literature, where interindividual differences in

metacognitive abilities correlate with differences in white matter structure (49), fMRI signals (50), and even lesions (51) that are mostly observed in the right hemisphere, although not always (52).

Beyond confidence weighting, our results illustrate another property of optimal inference in the brain: its hierarchical nature. Indeed, not only was there an overarching effect of change detection on both transition probabilities, but confidence weighting in the inferior frontal gyrus was specific to the context provided by the preceding item. Contextual control of information could be a distinctive feature of the lateral prefrontal cortex, as reported in other studies (53, 54). Confidence weighting in this region was controlled by a higher-order statistic: the rate of the overarching changes. Other parts of the prefrontal cortex, such as the anterior cingulate cortex (1) could be involved in monitoring this higher-order statistic and its changes.

One limitation of our study is that the brain networks reported here could be specific to our learning task (although we do show that they are amodal, being similarly activated in distinct auditory and visual sessions). However, other studies of hierarchical probabilistic inference found similar networks despite important differences in the task used. For instance, several such studies involved a binary choice on every trial, which requires collapsing the estimated probability distribution (1, 16, 17) or a single value (15, 20). In some studies, the estimated probability was also further used in the task for the valuation of outcomes (1, 20) or to orient visuospatial attention (18, 21).

Although our study reveals the brain networks engaged in hierarchical probabilistic inference, the details of their internal computations remain open. Where is the learned probabilistic model of sequences stored? Our results are certainly compatible with a locus within the inferior frontal gyrus itself, but this locus could also be a mere node in the information-processing pathway that implements confidence-weighting of surprise. Furthermore, what is the neural code underlying the type of sophisticated Bayesian computations that the present ideal-observer model requires? The brain may compute with full probability distributions (8, 11), or with scalar estimates of sufficient parameters as in learning rules with adjustable learning rates. In the future, a major step forward will be to investigate how confidence is represented at the neuronal level, and thereby how it should translate in fMRI signals. Our results do not speak directly to the representational format by which confidence is encoded, but merely to its use: they uncover the functional brain-scale consequences of fluctuating confidence levels in the regulation of learning. Future work should aim to clarify the format for these probabilistic computations and their orchestration in large brain-scale networks.

Materials and Methods

Participants and Task. The study was approved by the local Ethics Committee (CPP 08–021 Ile de France VII), and participants gave their informed written consent before participating. Twenty-one participants (12 females), aged between 20 and 33 (mean 25.3, SEM: 0.73) were recruited by public advertisement.

The task was run using Octave (Version 3.4.2) and PsychToolBox (Version 3.0.11). The experiment was divided into one training session, performed outside the scanner, and then four sessions performed in the MRI scanner. Each session presented a sequence of 380 stimuli, denoted A and B, which were perceived without ambiguity. In the scanner, A and B were either auditory or visual stimuli presented on alternated sessions. The modality of the first session was counterbalanced across participants. Due to technical problems, one participant only had visual sessions and another participant had three visual sessions and one auditory session (instead of two and two).

Fig. S1 depicts the task and timing. A fixation dot separated the visual stimuli and remained present during the auditory blocks. The sequences were generated according to the same process as in ref. 13. We summarize the key points here. A and B stimuli were drawn randomly based on predefined transition probabilities between stimuli, e.g., $p(A|B) = 0.8$ and $p(A|A) = 0.5$. Transition probabilities were constant only for a limited number of stimuli. The length of stable periods was itself randomly sampled from a geometric distribution with average length of 75 stimuli, truncated

at 300 stimuli to avoid overly long stable periods. In each stable period, transition probabilities were sampled independently and uniformly in the 0.1–0.9 interval, with the constraint that, for at least one of the two transition probabilities, the change in odd ratio $p/(1 - p)$ between consecutive stable periods should be at least fourfold. With these constraints, the actual values covered the 2D range 0.1–0.9 × 0.1–0.9 uniformly. In particular, there was no correlation between generative transition probabilities (Pearson correlation $\rho = 0.006$), even when restricted to values that follow the first change-point of a session ($\rho = 0.015$).

The sequence was paused every 15 stimuli on average, with a jitter of ± 1, 2, or 3 stimuli, to probe subjects about their inference. During the training session, subjects were asked to report their estimate of the transition probability, $p(A|A)$ or $p(A|B)$ depending on whether the stimulus preceding the question was A or B, and confidence in this estimate. They indicated their answers using continuous sliders. They were also asked to report when they detected a change in the transition probability. In the MRI scanner, subjects were only asked about their confidence, which they reported on a four-step scale with dedicated push button (Fig. S1).

Before the experiment, subjects were fully informed about the task structure, and notably the process generating the sequences. An interactive display made intuitive the notions of randomness, transition probabilities, and changes in these probabilities.

The Ideal Observer: An Optimal Bayesian Model. The ideal observer "inverts" the generative process underlying the sequences: it optimally estimates the likelihood of the current hidden transition probabilities (θ_t) given the observations received so far ($y_{1:t}$). The observer assumes that these probabilities are volatile: they can change from one stimulus to the next, with probability $\nu = 1/75$ (which is the generative value). It also assumes that when a change occurs, both transition probabilities are resampled randomly and independently from a prior distribution (π) that is uniform here.

The generative process has one key property that makes the estimation computationally tractable: the Markov property (55). The value of θ at time t must be the same as at time $t + 1$ if no change occurred. In case a change occurred, which happens with probability ν, the new value of θ is drawn from the prior distribution π, and it determines the likelihood of the stimulus observed at time $t + 1$. Therefore, to estimate θ at time $t + 1$, all that one needs to know is π, ν, θ at time t, and the new observation y_{t+1}. Crucially, past observations are no longer needed. Thanks to this Markov property, θ can be estimated iteratively, by going forward. At stimulus $t + 1$, the ideal observer updates optimally its estimation using Bayes rule:

$$p(\theta_{t+1}|y_{1:t+1}, \nu, \pi) \propto p(y_{t+1}|\theta_{t+1}, y_t) \int p(\theta_{t+1}|\theta_t, \nu, \pi) p(\theta_t|y_{1:t}, \nu, \pi) d\theta_t. \quad [1]$$

In Eq. **1**, $p(\theta_{t+1}|\theta_t, \nu, \pi)$ captures that θ may change, with probability ν, and be resampled from the prior distribution π. Eq. **1** provides the likelihood distribution, up to a scaling factor, over any value of θ. We computed this distribution and the scaling factor by numeric integration on a grid.

The probability of the next stimulus can be read as the mean of this likelihood distribution (25):

$$p(y_t|y_{1:t-1}, \nu, \pi) = \int p(y_t|\theta_t) p(\theta_t|y_{1:t-1}, \nu, \pi) d\theta_t = \int \theta_t p(\theta_t|y_{1:t-1}, \nu, \pi) d\theta_t \quad [2]$$

Confidence in the probability of the next stimulus can be read as log SD; that is, the negative log of the SD of the distribution. Further details on the algorithm and its implementation can be found in ref. 13.

For all analyses but one, the volatility ν was not learned by the ideal observer, but fixed to the generative value. We relaxed this constraint only once, to provide in Fig. S1*B* the trial-by-trial estimation of the posterior probability distribution of the volatility (ν). To this end, we assumed that all volatility levels are equally probable a priori, meaning that $p(\nu)$ is a constant and thus that the posterior probability is proportional to the likelihood:

$$p(\nu|y_{1:t}, \pi) \propto p(y_{1:t}|\nu, \pi) \propto \prod_{i=1}^{t} p(y_i|y_{1:i-1}, \nu, \pi). \quad [3]$$

The second line is derived using the chain rule, and the Markov property of the generative process highlighted above. The probabilities appearing in the product can be computed with Eq. **2**.

Delta-Rule Model with Constant Weighting and Bayesian Model Comparison. The delta rule with fixed learning rate α updates the transition probabilities θ_t as follows (56):

$$\begin{cases} \theta_{t+1}(A|y_t) = \theta_t(A|y_t) + \alpha(1 - \theta_t(A|y_t)) & \text{if } y_{t+1} = A \\ \theta_{t+1}(A|y_t) = \theta_t(A|y_t) - \alpha\theta_t(A|y_t) & \text{if } y_{t+1} = B \\ \theta_{t+1}(A|\neg y_t) = \theta_t(A|\neg y_t) & \text{for the stimulus not} \\ & \text{observed at time } t \end{cases} \quad \text{[4]}$$

Note that only one stimulus is observed at a given trial, and that the transition probability from the stimulus that was not observed is kept unchanged. We initialized the transition probabilities without bias (equal to 0.5).

First, we compared the best possible delta rule and the ideal observer. Using a grid search (resolution 0.025) and the actual sequences presented to subjects, we estimated the value of α that minimized the mean square error between generative values and fitted probabilities. The optimal, best fitting value was α = 0.10. Predictions from both models were then compared against subjects' reports.

We also compared the ideal observer and the delta rule, when its learning rate was let free to fit the data of each participant. This introduces a free parameter in one model; we therefore performed a Bayesian model comparison with balances optimally the goodness-of-fit and model complexity. The likelihood of a subject's probability estimate was computed with the ideal observer and the delta-rule models, as one minus the distance between the subject's and model's probability estimates. For the delta rule, this likelihood depends on the value of α. We computed for each subject the model evidence (i.e., the marginal likelihood):

$$p(r_{1:N}|M) = \int p(\alpha|M) \prod_i^N p(r_i|\alpha, M) d\alpha, \quad \text{[5]}$$

where M denotes the model and $r_{1:N}$ the subject's probability estimates. We considered a uniform prior probability for α in the delta-rule model; note that there is no free parameter for the ideal observer. We used the random effect analysis developed by Stephan et al. (57) and implemented in the SPM toolbox (function spm_BMS.m) to compute the "exceedance probability" of models, based on log model evidence values in each subject. This statistic quantifies the belief that a given model is more likely than the other in the general population, based on the group data.

MRI Data Collection and Preprocessing. MRI data were acquired on a 3 Tesla scanner (Siemens Trio) with a 32-channel coil. Functional echo planar images (EPI) were acquired with a T2*-weighted contrast. The first four scans of each session were discarded to allow for equilibration effects. We used a multiband acquisition (58, 59), with acceleration factor = 3, GRAPPA 2, TE = 30.4 ms, flip angle = 74°, 84 interleaved slices, to cover the whole brain with a repetition time of 2 s and an isotropic 1.5-mm resolution. The encoding phase direction was from posterior to anterior (occipital to frontal) within sessions. To estimate distortions, we acquired two slices with opposite phase encoding direction: one slice in the anterior to posterior direction (AP) and one slice in the other direction (PA), with TR = 7.860 ms, TE = 54 ms. Structural T1-weighted images were also acquired (1.0 × 1.0 × 1.1 mm, 160 slices) and coregistered with the mean EPI for each subject.

All preprocessing steps (expect the TOPUP correction) were performed using the SPM12 software (Wellcome Trust Center for Neuroimaging, University College London). EPIs were corrected for slice timing and realigned, using affine rigid body transformations, on the AP slice. The multiband acquisition produced distortions in particular in the occipital and frontopolar regions, which we corrected by estimating the susceptibility field with the AP/PA slices and unwrapping the EPIs using the TOPUP software (FSL, fMRIB).

Anatomical images were segmented into gray matter, white matter, and cerebrospinal fluid, bias corrected and spatially normalized to the standard SPM template in the Montreal Neurological Institute (MNI) space. The segmented, normalized anatomy of the first subject served to render group-level statistical maps on a representative cortical surface for display purpose. EPIs were spatially normalized using the same transformation as for anatomical images and smoothed with a 5-mm full width at half maximum (FWHM) Gaussian kernel. MRI images showed focal folding artifacts in the cerebellum of several subjects; data from this anatomical region were therefore excluded from further analysis.

MRI Data Analysis. Statistical analyses of MRI data were performed using SPM12 and general linear models (GLM) that included realignment parameters as covariates of no interest. The other regressors were convolved with the canonical hemodynamic response function (HRF) and its first temporal derivative to allow for temporal adjustment. GLMs were estimated using restricted maximum likelihood (the classical estimator in SPM). Individual coefficient maps corresponding to regressors convolved with the HRF were

smoothed with a 6-mm FWHM Gaussian kernel and masked for gray matter and were taken for the group level F contrasts (GLM4) and T contrasts (otherwise). Gaussian random field theory was used to compute cluster statistics (with a cluster-defining threshold of $P < 0.001$, unless stated otherwise) and peak statistics, and their significance levels corrected for multiple comparison over the entire brain with family-wise error (FWE). We report results at a cluster-level threshold $p_{FWE} < 0.05$, unless stated otherwise.

GLM1–3 included one regressor modeling the onsets of stimuli, and another for parametric modulation by the optimal estimation update induced by the current observation (GLM1) or the optimal confidence in the transition probability leading to the current observation (GLM2) or the optimal surprise elicited by the current observation (GLM3); they also included regressors modeling the onsets of motor responses, their modulation by reaction times and subjective confidence level.

GLM4 was used for the analysis of variance. It comprised 12 categorical regressors of stimulus onsets, corresponding to combinations of three predictability levels, two confidence levels, and two surprise levels, all computed from the ideal observer. Predictability levels corresponded to $\log_2(p) < 0.816$ and $\log_2(p) > 0.953$, these values were determined to form equally filled bins of trials across subjects. Confidence levels were determined by median split of trials within each predictability level, across all subjects. Surprise levels corresponded to expected ($P > 0.5$) and unexpected ($P < 0.5$) outcomes. The GLM also included as covariates, regressors for the onset of the first stimulus, questions onsets, and response onsets (and the modulation by reaction times and confidence levels). Individual categorical regression maps entered a group-level ANOVA, with subjects as independent factors, and predictability, surprise, confidence, and predictability x surprise (the interaction) levels as dependent factors. F contrasts for main effects of confidence and surprise are shown in Figs. 3A and 4A, respectively. The same ANOVA without the predictability x surprise interaction yields almost the exact same maps as Figs. 3A and 4A.

GLM5–7 were used to estimate fMRI signals in six bins of trials defined based on the ideal-observer confidence (GLM5), surprise (GLM6), and estimation update (GLM7). They also included the response onsets, and their modulation by the reaction times and subjective confidence.

GLM8 modeled the functional connectivity of the intraparietal sulcus and frontal eye field. Time series of fMRI signals in each region of interest (ROI) were extracted as the first eigenvariate and adjusted for linear effects of no interest (movement parameters, the onsets of stimuli, the onset of questions, the onset of responses, and their modulation by reaction times). The GLM included the fMRI time series from each ROI and the white matter, gray matter, and cerebrospinal fluid signals, defined with the standard SPM templates, to capture global physiological variations. The conjunction of functional connectivity profiles identified voxels showing positive correlations with both seed regions at a threshold $p_{FWE} < 0.05$.

GLM9 comprised one regressor modeling the onsets of stimuli, and its modulation by the relevant confidence (i.e., confidence associated to the transition probability related to the identity of the preceding stimulus) and one regressor modeling the onsets of stimuli, and its modulation by the irrelevant confidence (i.e., confidence associated to the transition probability related to the stimulus not observed at the preceding trial). The same onset regressor is included twice but with different parametric modulations, so that the parametric modulations are not serially orthogonalized by SPM (which could have favored one over the other).

GLM10 focused on the observations preceding ("pre") and following ("post") a streak of at least three repeated stimuli. These stimuli were modeled with four categorical regressors, distinguishing streaks within which confidence in the observed succession type always increased, and the others (all based on the ideal observer). To avoid potential confounds, we also included: a regressor modeling the onsets of all stimuli not modeled by the four categorical regressors, its modulation by the optimal confidence level, the onsets of questions and the onsets of responses, and their modulation by reaction times and subjective confidence levels.

GLM11 included one regressor modeling the onsets of stimuli and its modulation by the optimal confidence level; one regressor modeling the onsets of stimuli and its modulation by the optimal surprise level. Optimal confidence and surprise levels were not orthogonalized with respect to one another. It also included regressors modeling the onsets of motor responses, and their modulation by reaction times and subjective confidence level.

Definition of ROIs. For the regression analyses and comparison between modalities with a Bayesian t test, the clusters were defined with suprathreshold voxels as they appear in Fig. 3A for confidence regions (ANOVA analysis), Fig. 4A for surprise regions (ANOVA analysis), and as the conjunction of both main effects at $P < 0.005$ for the update region. For generating the plots Figs. 3 B and C, 4 B and C, and 5 B and C with cross-validation between sensory modalities, we took the 100 most significant voxels (in one modality)

within broadly defined anatomical regions from the Anatomy toolbox (60). More precisely, we used the atlas MacroLabel.img and the label file Macro.mat that can be found in the SPM12 software. All regions were defined bilaterally; we report the region names as they appear in the atlas. For Fig. 3 *B* and *C* and region intraparietal sulcus, we took the superior and inferior parietal lobules, the postcentral gyri; for the region inferior temporal gyrus, we took the inferior temporal and occipital gyri. For Fig. 4 *B* and *C* and region superior temporal sulcus, we took the superior and middle temporal gyri and the medial temporal pole; for the frontal eye field we took the pre- and postcentral gyri. For tests about hierarchical processing in the inferior/middle frontal gyrus, we took the suprathreshold voxels of the cluster at that location as it appears in Fig. 5*A* in the conjunction of functional connectivity analyses.

Availability of data. Nonthresholded whole brain maps are available for display and download on the public repository neurovault.org/collections/1181/.

Bayesian *t* Test. Bayesian *t* tests allow one to assess the significance of a mean being zero or different from zero. We computed paired differences at the subject level of the coefficients obtained from regression of fMRI signals on the ideal-observer confidence (or its surprise, its estimation update) in the two sensory modalities. A zero mean would denote no difference between modalities. We computed the Bayes factor in favor of zero mean using the Bayes Factor R package described in ref. 29. Note that the maximum possible value in favor of the null hypothesis (mean being zero) is 4.3 with our group size.

1. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.
2. Sutton R (1992) Gain adaptation beats least squares? *Proceedings of the 7th Yale Workshop on Adaptive and Learning Systems* (Yale University, New Haven, CT), pp 161–166.
3. Pearl J (1997) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, San Francisco).
4. Dayan P, Kakade S, Montague PR (2000) Learning and selective attention. *Nat Neurosci* 3:1218–1223.
5. Doya K (2002) Metalearning and neuromodulation. *Neural Netw* 15:495–506.
6. Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433.
7. Jaynes ET (2003) *Probability Theory: The Logic of Science* (Cambridge Univ Press, Cambridge, UK).
8. Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27:712–719.
9. Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. *Nature* 427:244–247.
10. Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 10:294–300.
11. Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. *Nat Neurosci* 9:1432–1438.
12. Meyniel F, Sigman M, Mainen ZF (2015) Confidence as Bayesian probability: From neural origins to behavior. *Neuron* 88:78–92.
13. Meyniel F, Schlunegger D, Dehaene S (2015) The sense of confidence during probabilistic learning: A normative account. *PLOS Comput Biol* 11:e1004305.
14. Strange BA, Duggins A, Penny W, Dolan RJ, Friston KJ (2005) Information theory, novelty and hippocampal responses: Unpredicted or unpredictable? *Neural Netw* 18:225–230.
15. Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30:12366–12378.
16. Summerfield C, Behrens TE, Koechlin E (2011) Perceptual classification in a rapidly changing environment. *Neuron* 71:725–736.
17. Iglesias S, et al. (2013) Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 80:519–530.
18. O'Reilly JX (2013) Making predictions in a changing world-inference, uncertainty, and learning. *Front Neurosci* 7:105.
19. Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP (2013) The neural representation of unexpected uncertainty during value-based decision making. *Neuron* 79:191–201.
20. McGuire JT, Nassar MR, Gold JI, Kable JW (2014) Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 84:870–881.
21. Vossel S, Mathys C, Stephan KE, Friston KJ (2015) Cortical coupling reflects Bayesian belief updating in the deployment of spatial attention. *J Neurosci* 35:11532–11542.
22. 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.
23. Gallistel CR, Mark TA, King AP, Latham PE (2001) The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *J Exp Psychol Anim Behav Process* 27:354–372.
24. Yeung N, Summerfield C (2012) Metacognition in human decision-making: Confidence and error monitoring. *Philos Trans R Soc Lond B Biol Sci* 367:1310–1321.
25. Gelman A, et al. (2013) *Bayesian Data Analysis* (CRC Press, Boca Raton, FL), 3rd Ed.
26. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423.
27. Friston K (2010) The free-energy principle: A unified brain theory? *Nat Rev Neurosci* 11:127–138.
28. Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci* 4:274–290.
29. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev* 16:225–237.
30. Kepecs A, Mainen ZF (2012) A computational framework for the study of confidence in humans and animals. *Philos Trans R Soc Lond B Biol Sci* 367:1322–1337.
31. Pouget A, Drugowitsch J, Kepecs A (2016) Confidence and certainty: Distinct probabilistic quantities for different goals. *Nat Neurosci* 19:366–374.
32. Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioral impact of decision confidence. *Nature* 455:227–231.
33. Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324:759–764.
34. De Martino B, Fleming SM, Garrett N, Dolan RJ (2013) Confidence in value-based choice. *Nat Neurosci* 16:105–110.
35. Boldt A, Yeung N (2015) Shared neural markers of decision confidence and error detection. *J Neurosci* 35:3478–3484.
36. Juslin P, Winman A, Hansson P (2007) The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals. *Psychol Rev* 114:678–703.
37. Wilson RC, Niv Y (2015) Is model fitting necessary for model-based fMRI? *PLOS Comput Biol* 11:e1004237.
38. Rutledge RB, Dean M, Caplin A, Glimcher PW (2010) Testing the reward prediction error hypothesis with an axiomatic model. *J Neurosci* 30:13525–13536.
39. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behavior in humans. *Nature* 442:1042–1045.
40. Preuschoff K, Bossaerts P (2007) Adding prediction risk to the theory of reward learning. *Ann N Y Acad Sci* 1104:135–146.
41. Diaconescu AO, et al. (2014) Inferring on the intentions of others by hierarchical Bayesian learning. *PLOS Comput Biol* 10:e1003810.
42. Tan H, Wade C, Brown P (2016) Post-movement beta activity in sensorimotor cortex indexes confidence in the estimations from internal models. *J Neurosci* 36:1516–1528.
43. Gallistel CR, Krishan M, Liu Y, Miller R, Latham PE (2014) The perception of probability. *Psychol Rev* 121:96–123.
44. Kheifets A, Gallistel CR (2012) Mice take calculated risks. *Proc Natl Acad Sci USA* 109:8776–8779.
45. Roesch MR, Esber GR, Li J, Daw ND, Schoenbaum G (2012) Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *Eur J Neurosci* 35:1190–1200.
46. Duncan J (2013) The structure of cognition: Attentional episodes in mind and brain. *Neuron* 80:35–50.
47. Wang L, Uhrig L, Jarraya B, Dehaene S (2015) Representation of numerical and sequential patterns in macaque and human brains. *Curr Biol* 25:1966–1974.
48. Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215.
49. Baird B, Cieslak M, Smallwood J, Grafton ST, Schooler JW (2015) Regional white matter variation associated with domain-specific metacognitive accuracy. *J Cogn Neurosci* 27:440–452.
50. Yokoyama O, et al. (2010) Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neurosci Res* 68:199–206.
51. Schmitz TW, Rowley HA, Kawahara TN, Johnson SC (2006) Neural correlates of self-evaluative accuracy after traumatic brain injury. *Neuropsychologia* 44:762–773.
52. Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G (2010) Relating introspective accuracy to individual differences in brain structure. *Science* 329:1541–1543.
53. Koechlin E, Ody C, Kouneiher F (2003) The architecture of cognitive control in the human prefrontal cortex. *Science* 302:1181–1185.
54. Azuar C, et al. (2014) Testing the model of caudo-rostral organization of cognitive control in the human with frontal lesions. *Neuroimage* 84:1053–1060.
55. Sutton RS, Barto AG (1998) *Introduction to Reinforcement Learning* (MIT Press, Cambridge, MA), 1st Ed.
56. Bornstein AM, Daw ND (2012) Dissociating hippocampal and striatal contributions to sequential prediction learning. *Eur J Neurosci* 35:1011–1023.
57. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46:1004–1017.
58. Feinberg DA, et al. (2010) Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS One* 5:e15710.
59. Moeller S, et al. (2010) Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med* 63:1144–1153.
60. Eickhoff S, et al. (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 25:1325–1335.