Neural variability compresses with increasing belief precision during Bayesian inference

Alexander Skowron[1,2,3*], Julian Q. Kosciessa[1,2,4], Robert Lorenz[5,6], Ralph Hertwig[1,7], Wouter van den Bos[8], & Douglas D. Garrett[1,2]*

[1]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin/London; [2]Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany; [3]Humboldt Universität zu Berlin, Germany; [4]Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands; [5]Lise Meitner Group for Environmental Neuroscience, Max Planck Institute for Human Development, Berlin, Germany; [6]Max Planck Dahlem Campus of Cognition (MPDCC), Berlin, Germany; [7]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany; [8]University of Amsterdam, The Netherlands.

*corresponding authors (skowron@mpib-berlin.mpg.de; garrett@mpib-berlin.mpg.de)

To make optimal decisions, intelligent agents must learn latent environmental states from discrete observations. Bayesian frameworks argue that integration of evidence over time allows us to refine our state belief by reducing uncertainty about alternate possibilities. How is this increasing belief precision during learning reflected in the brain? We propose that moment-to-moment neural variability provides a signature that scales with the degree of reduction of uncertainty during learning. In a sample of 47 healthy adults, we found that BOLD signal variability (SD_{BOLD}, as measured with functional MRI) indeed compressed with successive exposure to decision-related evidence. Crucially, more accurate participants expressed greater SD_{BOLD} compression primarily in Default Mode Network regions, possibly reflecting the increasing precision of their latent state belief during more efficient learning. Further, computational modeling of behavior suggested that more accurate subjects held a more unbiased (flatter) prior belief over possible states that allowed for larger uncertainty reduction during learning, which was directly reflected in SD_{BOLD} changes. Our results provide first evidence that moment-to-moment neural variability compresses with increasing belief precision during effective learning, proposing a flexible mechanism for how we come to learn the probabilistic nature of the world around us.

To make optimal decisions, animals often need to learn about "states" (i.e., decision-relevant properties) of the environment, such as the availability of fish in a nearby pond based on previous catch successes. These states are often not directly observable, so optimal agents need to combine the uncertainty about their beliefs with the evidence available to them. Research suggests that humans and other animals often take into account uncertainty during learning and decision-making[2-4], while suboptimal use of uncertainty has been linked to maladaptive behavior observed in clinical and ageing populations[3]. Bayesian decision theory prescribes how agents should combine evidence with their internal beliefs under uncertainty to arrive at optimal estimates of external states and provides a benchmark to assess rational behavior in the context of learning [2,3,5]. In this framework, an agent must represent a probability distribution over possible states. The variance of this distribution reflects the uncertainty of the agent's state belief. As evidence accumulates and the agent refines its representation of the "correct" world state, the variance of the belief distribution is thought to reduce. How this increasing belief precision during learning is reflected in the brain remains unclear.

A potential candidate for tracking the degree of state uncertainty is the moment-to-moment variability of neural responses. At the level of sensory cortex, computational modeling and non-human animal work suggests that neuronal variability signals uncertainty about peripheral inputs, such that neural population activity reflects sampling from a probability distribution over potential stimulus features; the higher perceptual uncertainty, the higher the neural variability [6-10]. This idea is in line with theoretic work suggesting that neural systems should maintain an appropriate degree of instability in the face of uncertainty to permit the exploration of alternative causes of incoming sensory input [11]. While plausible, no study to date has investigated whether stimulus-evoked neural variability tracks changes in perceptual uncertainty in a learning task and how it relates to behavior. With regard to human data, a recent study by Kosciessa, et al. [12] showed that increasing uncertainty about the task-relevance of different perceptual stimulus features is accompanied by an increase in cortical "excitability" (heightened desynchronization of alpha rhythms and increased entropy in the EEG signal). However, it is not clear whether these perceptual accounts of uncertainty-based shifts in neural variability also translate to higher-order decision variables that arise while learning about latent environmental states.

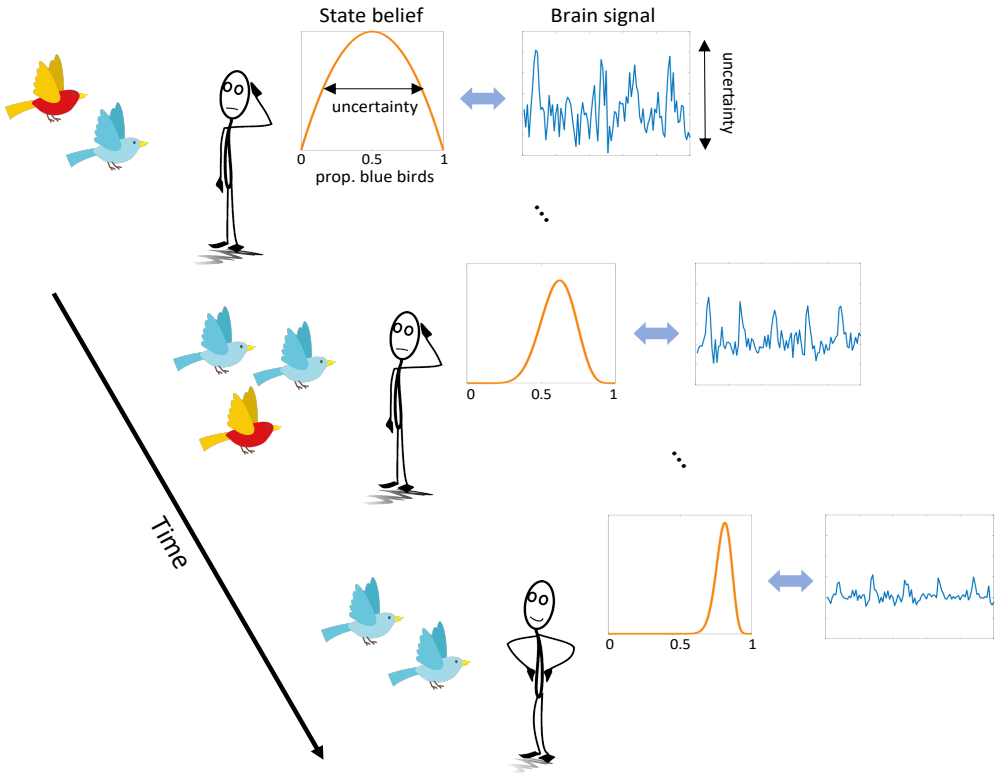

Figure 1. **Hypothetical model of uncertainty and brain dynamics.** In this example, the agent has to infer the proportion of blue birds in a world made up of red and blue birds by combining information across discrete observations over time. Under high state uncertainty early on in the learning process, more brain signal variability allows the brain to entertain a larger variety of possible states (i.e., proportions of blue birds in the world). With each new observation, the precision of the belief distribution increases and brain signal variability compresses.

In the context of learning, larger temporal variability under high uncertainty may not just represent exploration of potential world states but may allow more flexible updating of one's belief once new information becomes available. Indeed, previous human work suggests that higher brain signal variability affords larger cognitive flexibility [13,14]. For example, several studies have found that brain signal variability increases with increasing task demands (at least until processing limits are reached) and that the ability to upregulate variability predicts task performance [15-22]. These studies commonly argue that neural variability supports performance under increasing task demand by allowing the brain to maintain flexible responding to stimulus information. In line with this idea, Armbruster-Genç, et al. [16] observed better performance on a task switching paradigm under higher BOLD signal variability, whereas cognitive stability during distractor inhibition was related to lower brain signal variability. During learning, task demands are highest early on, when the brain needs to maintain high flexibility to incorporate incoming evidence for belief updating. This should allow it to converge on the "correct" state representation with learning. We therefore hypothesize that moment-to-moment brain signal variability compresses with increasing belief precision (i.e., decreasing uncertainty) during learning (Figure 1).

To test this hypothesis, we acquired functional MRI (fMRI) while participants performed a "marble task". In this task, participants had to learn the probability of drawing a blue marble from an unseen jar (i.e., urn) based on five samples (i.e., draws from the urn with replacement). In a Bayesian inference framework, the jar marble ratio can be considered a latent state that participants must infer. We hypothesized that (1) variability in the BOLD response (SD_{BOLD}) would compress over the sampling period, thus mirroring the reduction in state uncertainty, and that (2) subjects with greater SD_{BOLD} compression would show smaller estimation errors of the jars' marble ratios as an index of more efficient belief updating. A secondary aim of the current study was to directly compare the effect of uncertainty on SD_{BOLD} with a more standard General Linear Modelling (GLM) approach, which looks for correlations between average BOLD activity and uncertainty. This links our findings directly to previous investigations of neural uncertainty correlates, which disregarded the magnitude of BOLD variability [2,23-37]. We

hypothesized (3) that SD_{BOLD} would uniquely predict inference accuracy compared to these standard neural uncertainty correlates.

Our results showed that SD_{BOLD} closely tracked individual differences in the reduction of state uncertainty during learning which related to task accuracy in a unique manner that was not captured by a more standard GLM approach. As such, we identified a novel neural signal reflecting uncertainty during learning.

METHODS

Participants and procedure

51 healthy young adults (18 – 35 years) participated in the study. They were recruited from the participant database of the MPI for Human Development (Berlin, Germany) and gave written informed consent according to the guidelines of the German Psychological Society (DGPS) ethics board, who approved the study. Participants underwent fMRI while performing a perceptual gambling task. They received 10 Euros per hour in addition to a variable performance bonus of up to 2 Euros.

The marble task

The task was divided into four blocks of 18 trials each. Participants performed three blocks during fMRI acquisition, while the final block was performed outside the scanner. We will only consider the three MR blocks for our analyses. Participants received instructions and completed nine practice trials prior to the first scanning block. Each trial of the task consisted of three phases: a sampling, an estimation, and a gambling phase. To investigate brain signal variability changes during learning, we focused on the sampling and estimation phases in this study (referred to as the "marble task"). A description of the gambling phase, which was unrelated to the learning process, is provided in the Supplementary Methods. The task design also included a between-trial reward manipulation. Here, we collapsed trials across the two levels of the reward manipulation given that we assumed those would only affect the choice and not the learning part. The task was programmed using Presentation software (Version 14.9, Neurobehavioral Systems Inc., Albany, CA, USA).

In the marble task, participants were asked to estimate the proportion of blue marbles in an unseen "jar" (i.e., urn), containing a total of 100 blue and red marbles, based on five successive samples (i.e., draws from the urn with replacement) presented successively during the *sampling phase* (Figure 2A). In total, 18 different jars with different proportions of blue marbles (ranging from 0.1 to 0.91) were presented across all trials (Supplementary table S1). Each jar was presented four times across the four experimental blocks in random order. Each sample from a given jar could contain either one, five or nine marbles, which manipulated the informativeness of the samples (i.e., one marble was least informative for inferring the proportion of blue marbles whereas nine marbles was most informative). Samples were presented in 3x3 grids with grey marbles serving as placeholders to ensure similar visual inputs across samples. Additionally, within each sample the order of the marbles (i.e., the location in the 3x3 grid) was permuted. The samples for each jar were randomly drawn from a binomial distribution with the corresponding probability of drawing a blue marble. This draw was performed once so each jar was associated with a consistent set of five samples with varying sizes (note that the total number of marbles across samples thus varied between jars). This ensured that all participants received identical sample information. For each trial with the same unseen jar, the order of the associated samples was varied randomly. Each sample was presented for 1s followed by a fixation cross presented for 2 to 6s. Following the sampling phase, participants were asked to indicate their estimate of the proportion of blue marbles in the jar by adjusting the ratio of blue to red marbles in a 100-marble grid using an MR-compatible five-button box. The starting value of the grid was always set to 50 red and 50 blue marbles. With the upper and lower buttons, the participants were able to adjust the grid in steps of five marbles, and with the left and right button in a one marble step, respectively. With the middle button participants confirmed their estimation. The maximum reaction time for the estimation phase was set to 7s. Participants received a bonus payment based on the average accuracy of their marble ratio estimates (performance-dependent bonus).

Behavioral modelling

We modeled participants' responses using variants of a Bayesian observer model. The Bayesian observer represents the probability of drawing a blue marble from the unseen jar as a beta distribution with parameters

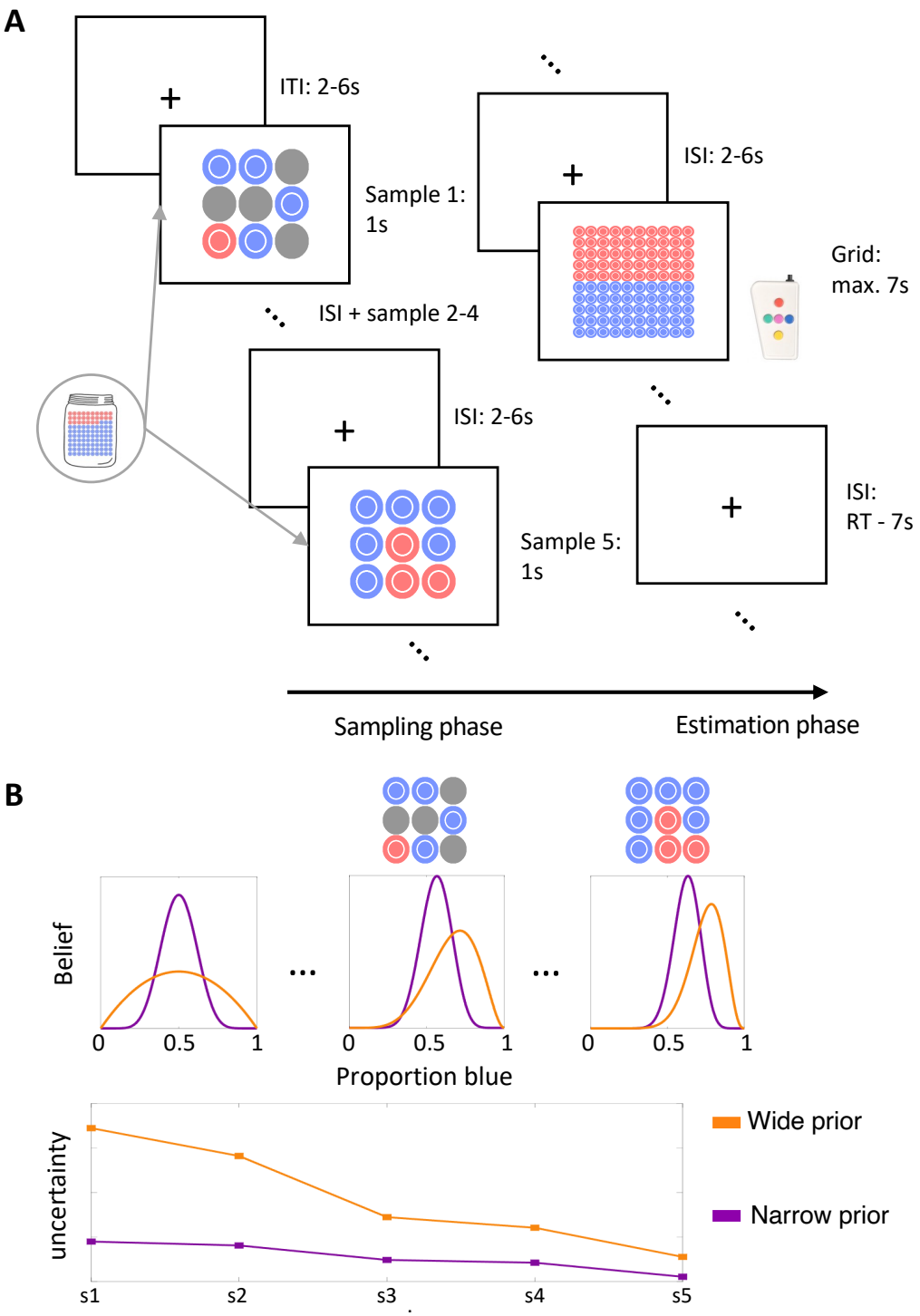

Figure 2. **Behavior in the marble task.** A: Task design. Participants observed five samples from an unseen jar (sampling phase) and then entered their estimate of the proportion of blue marbles in the jar into a 10x10 grid (estimation phase). B: Schematic illustration of the Bayesian observer model with varying priors. A narrower prior belief centered on a default 50:50 marble ratio predicts more misestimation of extreme jar proportions (top) and less uncertainty reduction during the sampling phase (bottom; s1 to s5 correspond to the five samples presented on a given trial).

α and β, which is updated after each draw according to Bayes' rule (Figure 2B). Because the beta distribution is a conjugate prior for the binomial distribution (the likelihood function in this task), the posterior belief distribution is also a beta distribution with $\alpha_{s+1} = \alpha_s + B_s$ and $\beta_{s+1} = \beta_s + R_s$, where B_s and R_s are the number of blue and red marbles for a sample s respectively. The prior for the first draw of each trial is given by $\alpha = \beta = 1$, representing a flat prior. The best estimate for the probability of drawing a blue marble is given by the expectation of the beta distribution $\mu = \frac{\alpha}{\alpha + \beta}$.

To model empirical response patterns affecting subjects' task accuracy, we considered two variants of this Bayesian observer model. In one model variant, we fit a parameter $\rho \geq 1$ for each subject, which described the initial setting of α and β on each trial prior to observing samples from a given jar (where $\rho = \alpha = \beta$; Figure 2B). A higher value for ρ describes a narrower prior around the default belief of 0.5. In a second Bayesian observer model variant, we fit a parameter $\delta \geq 0$, which exponentially weighs the evidence of each sample: $\alpha_{s+1} = \alpha_s + (B_s)^{\delta}$ and $\beta_{s+1} = \beta_s + (R_s)^{\delta}$. If $\delta < 1$, the evidence of larger draws is underweighted, while for $\delta > 1$ the evidence of larger draws is overweighted. Comparing these models allowed us to test whether deviations from the unbiased Bayesian observer model were related to prior representation or evidence weighting respectively.

To accommodate probabilistic responding of subjects, we compared these models with two alternative choice rules. One model family assumes that subjects' reported estimates represent draws from the final beta distribution on each trial. Another model family assumes that subjects' choices are draws from a truncated normal distribution ($0 \leq x \leq 1$) that is centered on the expectation of the final beta distributions: $N(E[beta(\alpha, \beta)],\ \sigma)$, where σ is a free model parameter that captures response noise around the model prediction. This parameter may reflect imprecision in entering one's marble ratio estimate into the response grid, but it can also capture any unmodeled sources of (biased and random) errors in jar estimates beyond those explained by the model parameters of interest. This parameter may thus also capture model misfit.

Bayesian observer models assume that participants track the full belief distribution over potential marble ratios and consider uncertainty when updating their estimate. To test whether subjects indeed behave in a Bayesian manner, we also fitted a simpler Rescorla-Wagner model that only updates a point estimate of the marble ratio belief: $\hat{R}_{s+1} = \hat{R}_s - \alpha(\hat{R}_s - \frac{B_s}{(B_s + R_s)})$, where $\hat{R}_s$ is the marble ratio point estimate after observing sample s, B_s and R_s are the number of blue and red marbles respectively, and $0 \leq \alpha \leq 1$ is the free learning rate parameter. To model subjects' responses, we included a free noise parameter σ modelling Gaussian response noise around the final point estimate, similar to the noisy choice rule described previously for the Bayesian observer models.

Model fitting and comparison. We fitted behavioral models to participants' responses on all trials of the three MR blocks by minimizing the joint negative log likelihoods under each model using the fmincon routine implemented in MATLAB (version 2017b). We ensured model convergence by fitting each model 10 times and used the fitting iteration with the minimal negative log likelihood across subjects. We compared model fits by computing the Bayesian Information Criterion (BIC): $BIC = k\ln(n) - 2\ln(L)$, where k is the number of parameters, n is the number of datapoints, and L is the maximized model likelihood. We report BICs for each subject and across all subjects.

Model simulations. We simulated behavior using the empirical parameter estimates and the trial sequences of each subject. This resulted in one synthetic dataset for each model with 51 simulated subjects each. We used these simulated datasets to perform model and parameter recovery checks. See the Supplementary Information for the results of these checks.

Image acquisition and pre-processing. Participants underwent functional MRI scanning at the Max Planck institute for Human Development (Berlin, Germany) in a 3 Tesla Siemens TrioTim MRI system (Erlangen, Germany) using a multi-band EPI sequence (factor 4; TR = 645 ms; TE = 30 ms; flip angle 60°; FoV = 222 mm; voxel size 3 × 3 × 3 mm; 40 transverse slices). This amounted to 1010 T2*-weighted functional images per scanner run per subject (except for one subject, who had 1020 acquired images per run). A T1-weighted structural scan was also acquired (MPRAGE: TR = 2500 ms; TE = 4.77 ms; flip angle 7°; FoV = 256 mm; voxel size 1 × 1 × 1 mm; 192 sagittal slices).

T1-weighted images were brain extracted using ANTs software (version 2.3.5, http://stnava.github.io/ANTs/) using population level templates from the OASIS dataset (https://figshare.com/articles/dataset/ANTs_ANTsR_Brain_Templates/915436). The functional T2*-weighted scans were pre-processed in FSL (version 5.0.11) FEAT separately for each run. The pipeline includes motion correction, brain extraction of the functional images, and spatial smoothing using a Gaussian 5mm kernel. Following FEAT, the functional images were first detrended using SPM's detrend function (at a 3rd order polynomial) and then high-pass filtered using a standard 8th order butterworth filter implemented in MATLAB (version 2017b) with a cut-off of 0.01 Hz. We then performed ICA on the resulting data using FSL MELODIC to identify residual artifacts. We manually labeled rejected components for 19 subjects (~37% of the total data; for details on the component rejection criteria see [38]) and then used these labels to train FSL's ICA classifier FIX to automatically label artifactual components for the remaining subjects. We tested different classification

thresholds and validated classification accuracy for several randomly selected subjects from the test set. We accepted a threshold of 20 after this manual inspection. Rejected components were then regressed out of the data using FSL's fsl_regfilt function. Lastly, functional images were first registered to the brain-masked T1-weighted images and then to MNI space (using the MNI152 template provided in FSL) via linear rigid-body transformation as implemented in FSL's flirt function.

Statistical analysis

Behavioral analysis. All behavioral analyses reported were run in SPSS (Version 25). To limit the effect of univariate outliers on the reported results, we winsorized the estimation error, extreme jar bias, and behavioral model parameter variables. Outliers were defined as datapoints that were 1.5 times the interquartile range beyond the 25th or 75th percentile respectively. We used the highest/lowest score of the non-outlying datapoints to impute outlying values. Between zero and three values were imputed for each variable.

Functional MRI analysis. Subject-level fMRI data were modelled using a mass-univariate GLM approach as implemented in SPM12 (Wellcome Department of Imaging Neuroscience, London, UK). To quantify BOLD response variability, we adapted a "least squares – single" (LS-S) approach as described by Mumford, et al. [39]. This approach is implemented in an in-house version of variability toolbox for SPM (http://www.douglasdgarrett.com/#software) developed by our research group. This toolbox takes as input a standard GLM design matrix. We included the following regressors in each subject's design matrix: Five regressors for the onsets of each successive sample presentation (with a duration of 1s), one regressor for the estimation phase onsets (with a duration of subjects' RTs), and one regressor for the gambling phase onsets (with a duration of subjects' RTs). All regressors were modelled as stick functions that were convolved with the canonical hemodynamic response function (HRF) and its first and second derivatives resulting in a total of 21 regressors per scanner run plus one constant regressors for each run. The toolbox then proceeds to iteratively fit GLMs which include one regressor modelling a single event and a second regressor modelling all other events of the same and all other conditions (this is done separately for each task run). To avoid issues with beta estimation close to the end of a timeseries (caused by truncation of the HRF), the toolbox discards any onsets that occur within the final 20s of the timeseries. Finally, the toolbox computes the standard deviation over the resulting (across-trial) beta estimates for each condition to yield the measure of SD_{BOLD} used in this study. Compared to other approaches to SD_{BOLD} estimation, this SD_{BOLD} quantification allows one to parse dynamics amongst neighboring events/time points, while accounting for hemodynamic delays.

To compare our SD_{BOLD} results to standard analysis approaches, we also obtained restricted maximum likelihood (ReML) GLM beta estimates for each task condition. The design matrix of this GLM included the same regressors as the one passed to variability toolbox for SD_{BOLD} estimation. We defined contrasts to summarize condition effects across task runs. However, this approach ignores the variance in uncertainty trajectories across trials that result from our sample size manipulation (see Figure S1B). To account for this variance, we ran another GLM modelling the modulation of the BOLD signal by the posterior variance (i.e., uncertainty) of the unbiased Bayesian observer model during the sampling process. The design matrix of this GLM included the following regressors for each scanner run: one regressor for the sample onsets (with a duration of 1s), one regressor for the parametric modulation of the sample onsets regressor by model-derived posterior variances, one regressor for the estimation phase onsets (with a duration of subjects' RTs), and one regressor for the gambling phase onsets (with a duration of subjects' RTs). We mean-centered the parametric modulation regressor prior to model estimation. All regressors were modelled as stick functions (except for the parametric modulation regressor) that were convolved with the canonical hemodynamic response function (HRF) and its first and second derivatives resulting in a total of 12 regressors per scanner run. We also included three additional constant regressors modelling the mean of each scanner run. To look at the effect of the parametric modulation regressor, we defined a contrast that picked out the beta estimates of this regressor for each scanner run (more precisely, the parametric modulation regressors convolved with the canonical HRF).

To relate within-person SD_{BOLD} or standard GLM beta estimates to the task design or behavior, we use a partial least-squares approach (PLS) [40]. PLS finds latent factors that express maximal covariance between brain and behavior/design matrices using singular value decomposition. Brain scores are defined as the product of the brain data matrix with their respective latent factor weights (saliences). We will refer to brain scores in our SD_{BOLD} analyses as "latent SD_{BOLD}" and to the brain scores of our parametric modulation analyses as "latent uncertainty modulation". To obtain a summary measure of the spatial expression of a latent variable, we can sum brain scores over all voxels. We performed 1000 permutation tests to assess the significance of the brain-

design/behavior relationship against a null model. We then divided each voxel's salience by its bootstrapped standard error, at +/- 3 (approximating a 99.9% confidence interval), which serves as a pseudo-normalized measure of voxel robustness. We determined clusters of voxels with robust saliences by applying a cluster mask with a minimum threshold of 25 voxels. Prior to all PLS analyses, brain measures images were grey matter masked using the tissue prior provided in FSL at a probability threshold of > 0.37, and restrained to voxels with non-zero values for the respective brain measure across subjects.

To account for univariate outliers and non-linear relationships, we ran all reported brain-behavior analyses on ranked brain and behavioral variables in line with previous research [41]. The results should thus be interpreted as monotonic rather than strictly linear relationships. Furthermore, we used a criterion of Cook's distance > 4/N to identify multivariate outliers in all PLS models, removing where present. We nevertheless report the results of all analyses for the full dataset (N=51) without outliers removed in the Supplementary Results.

Data availability

The code to perform all behavioral and neuroimaging analyses is available at: https://github.com/LNDG/Skowron_etal_2023.

RESULTS

Subjects' estimation errors are explained by individual differences in prior belief width

Participants performed the "marble task" during functional MRI scanning (Figure 2A). On each trial, they observed five sequential sample draws from a hypothetical "jar" containing a certain proportion of red and blue marbles (sampling phase). The marble ratio of the unseen jar varied from trial to trial and sample draws could contain one, five or nine marbles. Afterwards, participants indicated their estimate of the proportion of blue marbles in the jar by adjusting the number of blue marbles in a 10 by 10 response grid representing the unseen jar (estimation phase; see methods for further details). Our primary measure of interest was estimation error, defined as $\mathcal{E}_t = |\theta_t - \widehat{\theta}_t|$, where θ is the experienced proportion of blue marbles on trial t (i.e., the mean across samples; note that this quantity differs from the true proportion of the unseen jar but constitutes the best (unbiased) estimate given the available samples) and $\widehat{\theta}$ is the subject's reported estimate of the blue marble proportion of the jar on trial t (Figure 3A). We computed the median estimation error across trials as an individual difference measure of estimation accuracy. Subjects' median estimation error ranged from 0.01 to 0.20 (Median = 0.07, SD = 0.035). A Wilcoxon signed-ranks test revealed that estimation errors significantly differed from zero (Z = 6.215, one-sided p = $2.57 \cdot 10^{-10}$). This indicates differences between experienced and estimated marble ratios on the group level. Previous research suggests that people often underestimate probabilities of frequent events and overestimate probabilities of rare events in similar decision-from-experience tasks [42,43]. To see whether this effect contributed to estimation accuracy in our task, we examined whether subjects made more errors for jars with more extreme marble ratios. We ran a linear mixed model predicting trial-wise estimation error from the marble ratio distance from 50:50 for a given jar. To account for individual differences in intercepts, we entered subject ID as a covariate in the model. There was a significant positive relationship between trial-wise estimation-error and jar marble ratio distance from 50:50 (F(1,2702) = 154.608, p = $1.48 \cdot 10^{-34}$, semi-partial η^2 = 0.048). The more extreme the jar marble ratio, the larger subjects' estimation error on a given trial. In the following sections, we will refer to this effect as the "extreme jar bias". Next, we examined whether individual differences in this extreme jar bias explained individual differences in median estimation error. We computed individuals' extreme jar bias by running regression models predicting trial-wise estimation error from marble ratio distance from 50:50 for each subject separately. Individuals' extreme jar bias was defined as the fitted (unstandardized) slopes of these regression models (Mean = 0.198, SD = 0.230). We found that individual differences in extreme jar bias correlated strongly and positively with subjects' median estimation error (Pearson's r(49) = 0.63, p = $8.28 \cdot 10^{-7}$; Figure 3B). Biased estimation of extreme jar proportions thus contributed to individual differences in task accuracy.

We next fitted a set of Bayesian observer models that could account for this estimation bias in terms of inadequate uncertainty representations. The model with the overall best Bayesian Information Criterion (BIC) value was a Bayesian observer model which included a fitted prior parameter ρ, describing individual differences in the width of the belief distribution over marble ratios before observing any samples, and a response noise parameter σ, which captures variability of subjects' reported marble ratio estimates around model predictions (Table 1). The second-best overall model included an exponential evidence weight δ and a noisy choice rule

parameter σ , followed by an exponential evidence weight and fitted prior model assuming responses were sampled from the final trial posterior of the belief distribution. A Rescorla-Wagner model had the worst fit out of all models, which suggests that people indeed keep track of and leverage uncertainty estimates to update their marble ratio belief in a Bayesian manner. Parameters of the winning model were recoverable from simulations and model recovery was acceptable at the group level (see Supplementary Methods for details).

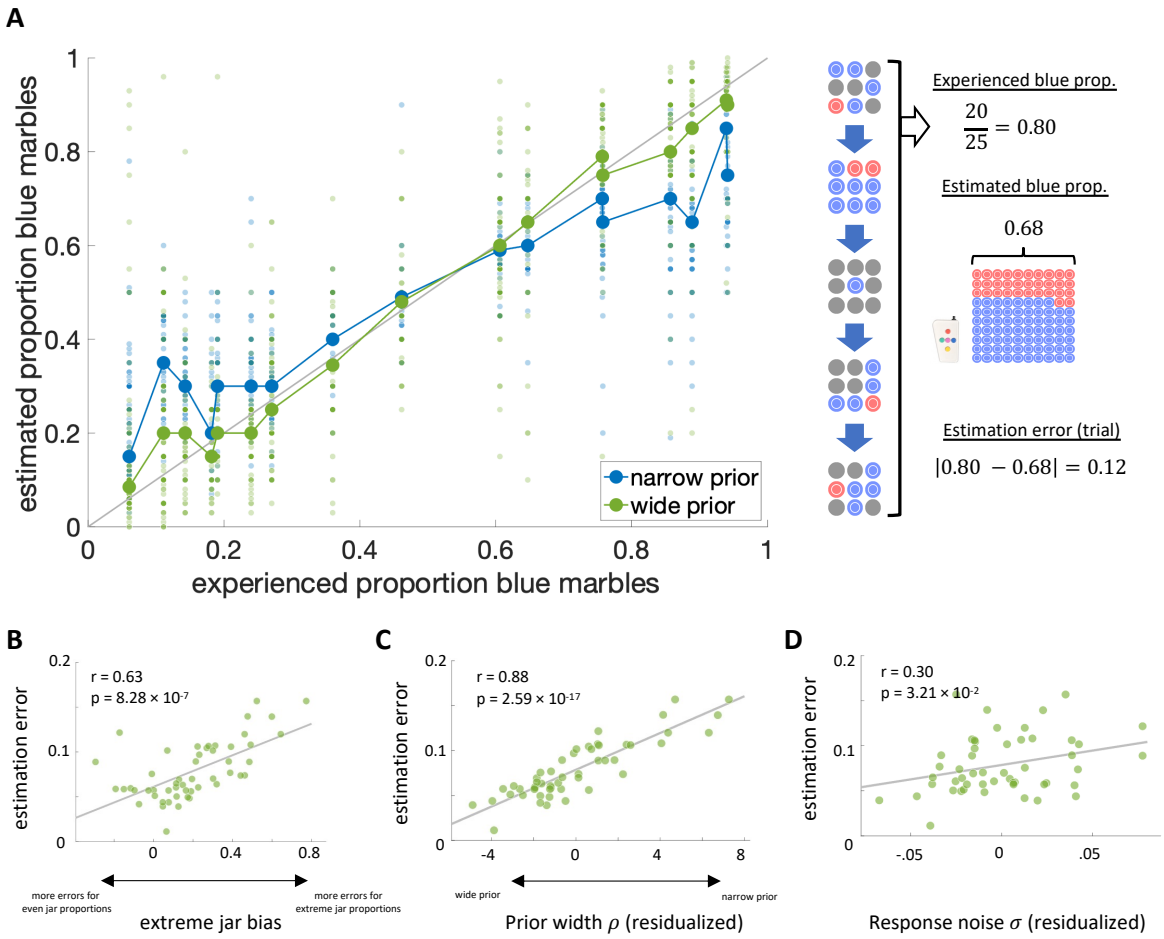

Figure 3. **Latent sources of individual estimation errors.** A: Left: Reported blue marble proportion estimates across trials and subjects against the experienced blue marble proportion across samples on a given trial. Groups are based on a median split on the fitted prior width parameter r of the Bayesian observer model (i.e., the narrow prior group had higher parameter estimates). Right: Schematic example of our trial-wise estimation error measure. B: Individual differences in median estimation error related to estimation accuracy for jars with extreme marble ratios. C-D: The fitted prior width and response noise parameters of the Bayesian observer model predicted differences in median estimation error. Plotted parameters are residualized by the other model parameter respectively.

We next tested whether the prior width parameter ρ of the winning Bayesian observer model captured individual differences in extreme jar bias that contributed to subjects' estimation error while controlling for other unmodelled sources of error captured by the response noise parameter σ. The fitted prior parameter ρ of the winning Bayesian observer model had a median of 2.45 (SD = 4.23), indicating that subjects on average had a narrower prior than the unbiased Bayesian observer ($\rho = 1$). Moreover, the fitted noise parameter σ had a median of 0.10 (SD = 0.05), highlighting that subjects' reported marble ratios deviated from model predictions by 10 marbles on average. We ran multiple regression analyses predicting median estimation error and extreme jar bias from the fitted prior parameter ρ and noise parameter σ respectively. In the first regression model, a narrower prior parameter ρ ($t(48) = 13.152$, $p = 1.57 \cdot 10^{-17}$, semi-partial $\eta^2 = 0.771$; Figure 3C) and a higher noise parameter σ ($t(48) = 4.501$, $p = 4.30 \cdot 10^{-5}$, semi-partial $\eta^2 = 0.090$; Figure 3D) uniquely predicted larger estimation errors (full model: ($F(2,48) = 88.135$, $p = 8.54 \cdot 10^{-17}$, $R^2 = 0.786$). Importantly, the prior width parameter ρ accounted for most of the explained variance in estimation errors suggesting that our model captured individual differences in subjects' response behavior well. A narrower parameter ρ reflects a stronger prior belief in a 50:50 marble ratio, which should bias estimates for more extreme jars towards this (incorrect)

default belief. In line with this expectation, our second regression model revealed that subjects' extreme jar bias was positively related to prior width ρ (t(48) = 11.780, p = $9.11 \cdot 10^{-16}$, semi-partial η^2 = 0.499) but also negatively to response noise σ (t(48) = -2.983, p = $4.48 \cdot 10^{-3}$, semi-partial η^2 = -0.032; full model: F(2,48) = 114.968, p = $4.956 \cdot 10^{-19}$, R^2 = 0.827). The significant (although smaller) effect of the noise parameter σ on extreme jar bias was driven by a small number of subjects with negative extreme jar bias slopes (i.e., subjects who made larger errors for jars with even rather than extreme marble ratios), which the Bayesian observer model could not account for (see Supplementary Results for further details). Taken together, these analyses indicate that the width of the prior belief substantially impacts individual task performance by capturing subjects' tendency to overestimate jars with blue marble proportions close to 0 and underestimate jars with blue marble proportions close to 1 (Figure 3A).

Table 1. Model comparison. BIC for each fitted behavioral model.

Model	Rescorla-Wagner	Bayesian observer			
Parameters	α, σ	ρ	δ	ρ, σ	δ, σ
BIC	$-3.38 \cdot 10^3$	$-3.68 \cdot 10^3$	$-3.82 \cdot 10^3$	$-4.59 \cdot 10^3$	$-4.16 \cdot 10^3$

BOLD signal variability tracks individual differences in prior uncertainty and compresses during learning

To investigate our hypothesis that BOLD signal variability (SD_{BOLD}) collapses with successive sample presentations, we ran a task PLS analysis [40] relating SD_{BOLD} to the five sample periods. This analysis returned a significant latent effect (permuted p = 0) showing that SD_{BOLD} reduced with successive exposure to marble samples across a distributed set of cortical brain regions spanning the parietal, prefrontal and temporal lobes (Figure 4A, see Figure S2A for full axial brain plots and Table S3 for peak voxel coordinates in robust clusters). We further investigated this trend by fitting orthogonal polynomial contrasts in a linear mixed model entering subject ID as covariate. The pattern of latent SD_{BOLD} change was best described by a linear contrast (F(1,201) = 40.444, p = $1.34 \cdot 10^{-9}$, semi-partial η^2 = $1.12 \cdot 10^{-2}$) over a quadratic (F(1,201)=3.066, p = $8.15 \cdot 10^{-2}$, semi-partial η^2 = $8.41 \cdot 10^{-4}$) and cubic (F(1,201)=3.416, p = $6.61 \cdot 10^{-2}$, semi-partial η^2 = $9.61 \cdot 10^{-4}$) one.

Next, we asked whether the ability to collapse SD_{BOLD} over the sampling period related to individual differences in task accuracy. To this end, we quantified SD_{BOLD} change (ΔSD_{BOLD}) by fitting linear regression slopes to the SD_{BOLD} estimates for the five sample presentations in each voxel. We ran behavioral PLS analysis relating these SD_{BOLD} slopes to subjects' median estimation error. Four Cook's d (i.e., multivariate) outliers were removed from this and all subsequent reported analyses resulting in N = 47 (we refer to the Supplementary Information for PLS results ran on the full dataset). We found a significant latent relationship (permuted p = $3.00 \cdot 10^{-3}$, Figure 4B, cf. Figure S3A for N = 51) revealing that subjects who decreased SD_{BOLD} more, especially in parietal, prefrontal (PFC) and temporal cortex, also produced smaller estimation errors (Spearman's r = 0.612). Notably, this set of brain regions largely overlapped with the canonical Default Mode Network (DMN) that is typically observed in resting-state fMRI (BSR > 3; Figure 4C, see Figure S2B for full axial brain plots and Table S4 for peak voxel coordinates in robust clusters) [1].

We further investigated our hypothesis that individual differences in state uncertainty representations could explain the relationship between ΔSD_{BOLD} and estimation error. Our winning behavioral model suggested that subjects with a narrower prior belief distribution over marble ratios (i.e., large ρ parameter) showed little decrease in state uncertainty during sampling, because they started out with less uncertainty to begin with, and made more errors because they remained inflexible about their (incorrect) prior belief of a 50:50 marble ratio. In contrast, more accurate subjects (i.e., ρ close to 1) represented maximal state uncertainty at the start of a trial and reduced uncertainty more with each presented sample (Figure 2B). We expected ΔSD_{BOLD} to mirror these individual differences in state uncertainty change during the sampling phase. To test this hypothesis, we ran multiple regression analysis predicting latent ΔSD_{BOLD} (i.e., the whole-brain pattern of ΔSD_{BOLD} which relates to median estimation error; see Figure 3b) from the prior width ρ parameter of the winning behavioral model. We also entered the response noise parameter σ as a control variable, which captures all unmodeled sources of estimation error and thus serves as a measure of model misfit (see Methods). This model explained a significant amount of variance (F(2,44) = 4.489, p = $1.68 \cdot 10^{-2}$, R^2 = 0.169). There was a significant main effect only for prior width ρ (t(44) = 2.703, p = $9.72 \cdot 10^{-3}$, semi-partial η^2 = 0.138) but not for noise parameter σ (t(44) = 0.425, p = 0.673). Therefore, the wider a participant's prior belief (i.e., smaller ρ), the more they collapsed

SD_{BOLD} during the sampling phase (Figure 4D). To ensure that that these effects were not due to the specific parameterization of the winning behavioral model, we replicated this analysis also for the parameters of the (second-best) evidence weight model (see Supplementary Results).

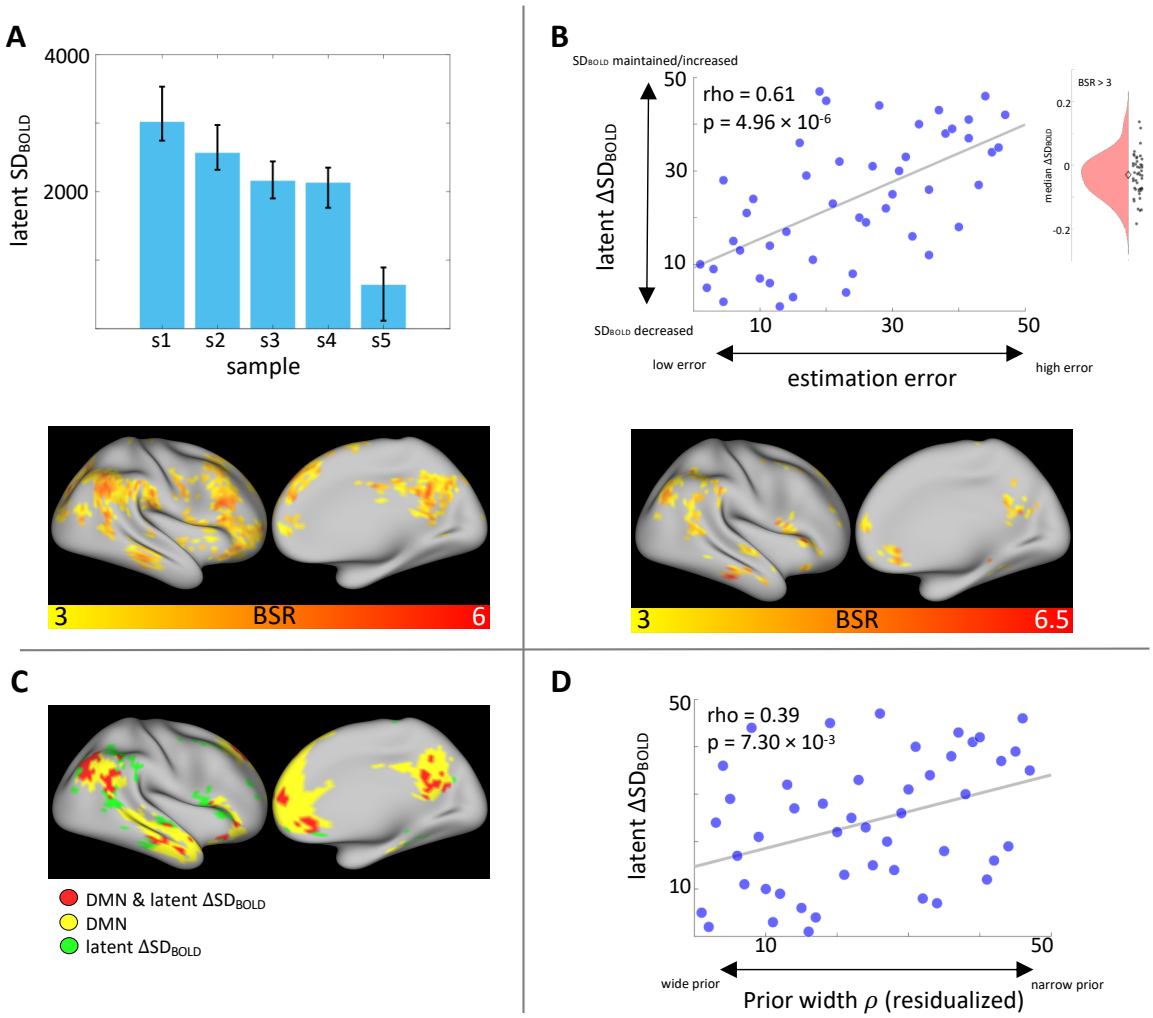

Figure 4. **SD_{BOLD} tracks idiosyncratic uncertainty representations and predicts inference accuracy.** A: Task PLS revealed average SD_{BOLD} compression over the sampling phase. Error bars represent bootstrapped 95% confidence intervals of the mean. For visualization, we also plot means for groups based on a median split on the fitted prior width parameter r of the Bayesian observer model. B: Change in SD_{BOLD} over the sampling phase correlated with estimation accuracy. Greater SD_{BOLD} collapse related to smaller median estimation errors. The inset shows the distribution of median SD_{BOLD} change over positive effect voxels (BSR > 3). The diamond marker indicates the distribution median. C: Regions showing performance-related SD_{BOLD} change (see panel B) largely overlapped with the canonical Default Mode Network (DMN)[1]. D: Fitted prior width parameter estimates (residualized by noise parameter σ) of the Bayesian observer model positively correlated with performance-related SD_{BOLD} change. Note: All variables are rank-transformed.

It is possible that our within-person effects of ΔSD_{BOLD} are not independent of between-subject differences in SD_{BOLD} levels, which have also been linked to task performance in previous studies [13]. To account for this, we computed a latent SD_{BOLD} control variable by matrix multiplying brain saliences from the behavioral PLS analysis relating ΔSD_{BOLD} to estimation error (which reflect the spatial expression of the observed relationship) with subjects' SD_{BOLD} level at the first sample presentation period. We found no significant correlation between latent ΔSD_{BOLD} and latent SD_{BOLD} at first sample exposure (Spearman's $r(45) = -0.073$, $p = 6.28 \cdot 10^{-1}$). Next, estimation error was significantly related to ΔSD_{BOLD} ($t(44) = 5.430$, $p = 2 \cdot 10^{-6}$, semi-partial $\eta^2 = 0.625$) when controlling for the effect of latent SD_{BOLD} at first sample exposure ($t(44) = 1.806$, $p = 7.78 \cdot 10^{-2}$, semi-partial $\eta^2 = 0.208$; full model: $F(2,44) = 15.747$, $p = 7 \cdot 10^{-6}$, $R^2 = 0.417$). Likewise, prior width ρ (residualized by noise parameter σ) of the winning Bayesian-observer model was significantly related to ΔSD_{BOLD} ($t(44) = 3.151$, $p = 2.92 \cdot 10^{-3}$, semi-partial $\eta^2 = 0.409$) when controlling for the effect of latent SD_{BOLD} at first sample exposure ($t(44) = 2.526$, $p = 1.52 \cdot 10^{-2}$, semi-partial $\eta^2 = 0.328$; full model: $F(2,44) = 7.618$, $p = 1.44 \cdot 10^{-3}$, $R^2 = 0.257$). In both models, the inclusion of the control variable also did not diminish the effect size of ΔSD_{BOLD} on these

behavioral measures. Thus, the effect of within-person SD_{BOLD} compression on task performance was independent of between-subject differences in initial SD_{BOLD} levels.

Overall, these results indicate that the degree to which subjects compress SD_{BOLD} with learning predicts individual differences in inference accuracy, which can be explained in terms of idiosyncratic uncertainty representations.

Standard GLM neural uncertainty correlates are distinct from BOLD variability effects

Finally, we compared our SD_{BOLD} findings to a more standard GLM analysis approach looking for brain areas whose (average) BOLD response magnitude is correlated with uncertainty. To this end, we obtained voxel-wise GLM beta estimates for each of the sample periods and entered them into behavioral PLS analysis relating them to estimation error. This multivariate approach makes the results directly comparable to our main finding showing that SD_{BOLD} compression during learning relates to individual differences in task performance. This analysis did not yield a significant latent relationship (permuted p = 0.161 for N = 47 and permuted p = 0.268 for N = 51), suggesting that SD_{BOLD} but not mean BOLD change predicted individual differences in inference accuracy. However, this change in average BOLD activity over the sampling period does not account for more nuanced modulation of the BOLD signal by uncertainty resulting from the varying sample sizes in our task design (see Figure S1B). In other words, an unbiased Bayesian observer model would predict more uncertainty reduction after observing a larger compared to a smaller sample in addition to the overall reduction in uncertainty with the total accumulated evidence (i.e., trial time). Our SD_{BOLD} metric required discrete uncertainty conditions (with sufficient trial counts) and we thus only considered the main effect of trial time on uncertainty. However, in a standard GLM approach we could model the parametric modulation of the BOLD signal by the posterior variance of the belief distribution (i.e., uncertainty) for the unbiased Bayesian observer model, which accounts for both trial time and sample size effects on uncertainty. Again, we used a multivariate PLS approach to relate individuals' voxel-wise GLM beta estimates (for a parametric uncertainty regressor) to

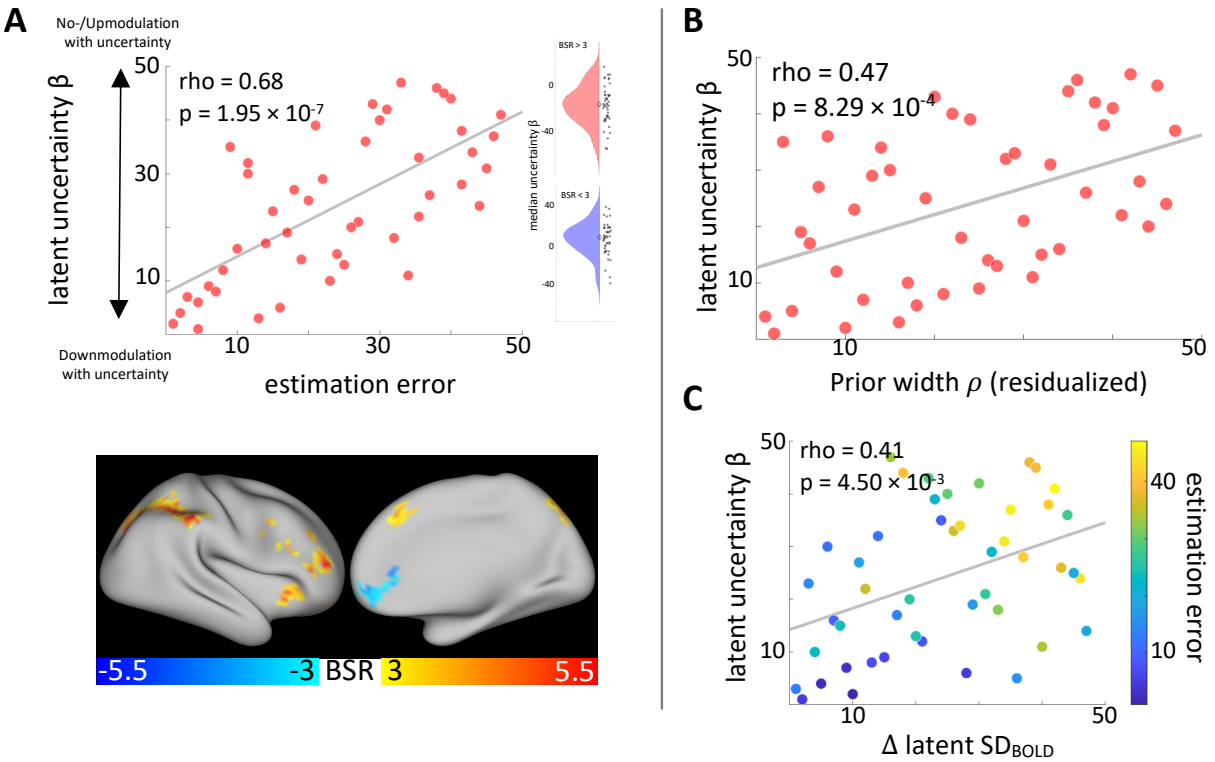

Figure 5. **Standard GLM neural uncertainty correlates are distinct from SD_{BOLD} effects.** A: BOLD signal modulation by uncertainty (GLM b estimates) correlated with average estimation accuracy. More downmodulation with increasing state uncertainty across the whole brain related to smaller estimation errors. The inset shows the distribution of b estimates over positive (BSR > 3, top) and negative (BSR < 3, bottom) effect voxels. The diamond marker indicates the distribution median. B: Fitted prior width parameter values (residualized by noise parameter σ) of the Bayesian observer model positively correlated with performance-related BOLD signal modulation by uncertainty (GLM b estimates). C: Changes in SD_{BOLD} over the sampling phase and BOLD signal modulation by uncertainty (GLM b estimates) uniquely predicted task performance. Good performers showed both more SD_{BOLD} compression and more downmodulation of the BOLD signal when state uncertainty was high. Note: All PLS results are presented for rank-transformed variables.

median estimation error. This revealed a significant latent association (permuted p = $3.40 \cdot 10^{-2}$, Spearman's r = 0.675, Figure 5A, cf. Figure S3B for N = 51). Robust positive clusters were mainly located in the lateral parietal cortex, PFC (dorsomedial, ventrolateral, and rostrolateral parts), and in the insula. Robust negative clusters were present in ventromedial PFC (vmPFC) and in the left hippocampus (see Figure S2C for full axial brain plots and Table S5 for peak voxel coordinates in robust clusters). This latent relationship was related to the prior width parameter ρ of the winning Bayesian observer model, similar to the effect of ΔSD_{BOLD} (Figure 5B). For further details on these results, please refer to the Supplementary Information.

Qualitatively, these effect regions seem to show little overlap with those showing an effect for ΔSD_{BOLD}. To test the spatial specificity of ΔSD_{BOLD} effects, we computed a new latent variable by extracting the brain saliences from the behavioral PLS analysis relating latent ΔSD_{BOLD} to estimation error and matrix-multiplied them with the uncertainty beta estimates in each voxel from our standard parametric GLM analysis. We then ran a multiple regression analysis predicting median estimation error from latent ΔSD_{BOLD} and this spatially-matched latent parametric uncertainty modulation variable. The regression model was significant ($F(2,44) = 13.901$, p = $2.10 \cdot 10^{-5}$) with an R^2 of 0.387. The main effect of latent ΔSD_{BOLD} was significant ($t(44) = 5.265$, p = $4 \cdot 10^{-6}$, semi-partial $\eta^2 = 0.386$) but not for the new latent control variable ($t(44) = 0.975$, p = 0.335). This indicates that ΔSD_{BOLD} effects and the parametric uncertainty modulation of the BOLD signal on performance are spatially distinct in the brain.

Furthermore, we found a significant positive correlation between the latent ΔSD_{BOLD} (Spearman's $r(45) = 0.407$, p = $4.55 \cdot 10^{-3}$) and latent uncertainty modulation effects on estimation accuracy indicating that more accurate performers show both types of uncertainty representation in the brain. We ran multiple regression analysis predicting estimation errors from both latent ΔSD_{BOLD} and latent uncertainty modulation to investigate whether they explained unique variance in estimation errors. The full model was significant ($F(2,44) = 31.883$, p = $2.76 \cdot 10^{-9}$, $R^2 = 0.592$; Figure 5C). We found a significant main effect of both latent ΔSD_{BOLD} ($t(44) = 4.844$, p = $1.60 \cdot 10^{-5}$, semi-partial $\eta^2 = 0.218$) and latent uncertainty modulation ($t(44) = 3.829$, p = $4.04 \cdot 10^{-4}$, semi-partial $\eta^2 = 0.136$). This suggests that both types of neural uncertainty correlates uniquely relate to task accuracy.

Overall, these analyses reveal that SD_{BOLD} effects are spatially distinct and uniquely relate to task accuracy compared to neural correlates revealed by a more standard GLM approach.

DISCUSSION

In this study we show that reductions in state uncertainty over the course of learning co-occur with a compression in BOLD signal variability. Moreover, individuals who decreased state uncertainty to a greater extent showed more pronounced SD_{BOLD} compression and made smaller estimation errors on average. Behavioral modeling suggested that better performers reduced state uncertainty to a greater extent because they began with a wider (more uncertain and flexible) prior belief distribution before observing any samples, resulting in more unbiased marble ratio estimates.

Our findings establish SD_{BOLD} as a novel within-person neural correlate of uncertainty, which reduces over the course of Bayesian inference/learning, and relates to inference accuracy. The current findings thus add to a growing literature demonstrating that within-person SD_{BOLD} modulation in the face of varying task demands facilitates adaptive behavior across different cognitive domains (see [13]). Our study is the first to directly show that SD_{BOLD} tracks reductions in uncertainty. The bulk of past empirical work arguing for such a link showed modulation of brain signal variability across disparate task conditions that varied along several dimensions beyond uncertainty (e.g., cognitive demand, bottom-up sensory input or processing requirements [17,18,44,45]). By employing a learning paradigm, we were able to operationalize uncertainty in a highly precise manner; specifically, one's uncertainty regarding the latent state of an environmental variable, a property that can be reduced through learning from observations. To our knowledge, the only other study to date that has directly investigated brain signal variability changes with uncertainty is a recent study by Kosciessa, et al. [12], who systematically manipulated uncertainty in a perceptual decision-making task by varying the number of visual features that could be probed in a subsequent decision phase. This paradigm allowed the researchers to equate bottom-up visual input across uncertainty conditions, similar to our own study. However, their study focused on irreducible uncertainty (sometimes referred to as risk or outcome uncertainty in the literature [2,3]) and only considered temporal variability in the EEG signal. In contrast, we show that reductions in state uncertainty (i.e., uncertainty about the probability of an outcome) due to learning are mirrored in the variability of the fMRI BOLD

signal. This supports the longstanding idea that brain signal variability enables cognitive flexibility [13,14], which is required early on in the learning process to update one's internal state belief.

Another innovation of our study over prior BOLD variability work more generally[13,14] is the use of an event-related method to quantify trial-to-trial SD_{BOLD}. This allowed us to track the evolution of SD_{BOLD} with changes in uncertainty on a short time scale (several seconds) [39]. This method is inspired by the machine-learning fMRI literature allowing for the estimation of GLM beta parameters for single events, over which we can compute the variance for each uncertainty condition.

These features of our study resolve various shortcomings of previous work and provide strong support for a link between brain signal variability and uncertainty.

Individual differences in uncertainty representations track inference accuracy

We found that low performing individuals made most estimation errors for extreme marble ratios, such that jars with blue marble proportions close to one were underestimated and those with proportions close to zero were overestimated. This effect has been reported previously for tasks where participants are asked to judge outcome probabilities from experience, sometimes referred to as under-extremity [42,43,46]. Notably, some studies found that this effect is abolished when participants were asked to give verbal judgements instead [47]. This may suggest that individual differences in extreme jar bias result from people's (in)ability to accurately enter their marble ratio estimates into the grid in the estimation phase of our task. However, we found that brain activity during sampling (*prior* to the estimation phase) predicted individual differences in peoples' (average) estimation error. It is thus likely that the observed bias reflects people's internal representations rather than being a mere consequence of our chosen response modality.

Our winning Bayesian observer model suggested that such estimation errors could best be explained by individual differences in the representation of the initial prior belief over blue marble proportions rather than in belief updating; low performers held a narrower prior belief distribution over the default belief of a 50:50 marble ratio. It is possible that this prior was induced by our task design given the default position of the response grid. The fitted prior could thus reflect individual differences in anchoring to this task feature [48]. Alternatively, this bias could reflect bona fide trait-like individual differences in expectations about extreme outcomes. In line with this explanation, Glaze, et al. [49] found that performance differences on a change-point inference task were best explained by participants' prior width over potential change rates, which cannot simply be explained by anchoring. Note that the distinction between prior beliefs and belief updating is also pertinent to other approximately-Bayesian and non-Bayesian cognitive models that have been used to model the inference process in similar tasks, which we did not consider here [43,50]. These models account for cognitive limitations in human information processing and may provide an even better fit to behavioral data. While not the main aim of the current study, future modelling and experimental work is required to pin down the source of individual differences in under-extremity observed in the current study.

We showed that less SD_{BOLD} compression (independent of initial SD_{BOLD} levels) in poorer performers could best be explained by more limited uncertainty reduction, based on parameter estimates for the winning behavioral model. Still, our results rely on the assumption that participants represent and utilize uncertainty akin to an (approximate) Bayesian observer. Future work should reduce reliance on model assumptions and obtain subjective uncertainty estimates. For example, one could regularly ask people to report their state belief together with their belief confidence throughout the learning process to directly track updates in their subjective belief distribution.

Performance-related SD_{BOLD} collapse in the Default Mode Network

We observed performance-relevant collapse in SD_{BOLD} in a network of brain regions that largely overlapped with regions commonly ascribed to the default mode network (DMN) in resting-state fMRI connectivity analyses [1,51-54], including parietal (precuneus, inferior parietal lobe including the angular gyrus), prefrontal (superior frontal gyrus, frontopolar cortex, orbitofrontal cortex, paracingulate gyrus) and temporal (middle temporal gyrus) cortices. Previous work has consistently reported deactivation of this brain network during externally-cued, demanding cognitive tasks compared to rest [51,52,54-57]. In contrast, internally-oriented tasks that call for self-referential processing and memory recollection demonstrate increased activity in key regions of the DMN [54,58-60]. With respect to brain signal variability, one study by Grady and Garrett [44] found *increased* SD_{BOLD} in DMN

regions in externally- as compared to internally-guided tasks. Given previous reports that higher local SD_{BOLD} is associated with lower whole-brain functional network dimensionality [61], this finding may reflect more crosstalk between the DMN and other brain networks during externally-guided tasks. Therefore, SD_{BOLD} compression during learning may reflect the gradual build-up of an internal world model; larger SD_{BOLD} under high state uncertainty early on may afford a more flexible incorporation of incoming information to update one's internal state belief. As state uncertainty reduces and SD_{BOLD} compresses, the brain arrives at a stable internal belief representation, which informs one's jar marble ratio estimate in the subsequent task phase.

Previous research has shown that internally-guided tasks evoke a coupling between the DMN and the frontoparietal-control network to support task performance [62]. In our task, we would expect such a coupling to emerge towards the end of the sampling period reflecting the utilization of one's internal state belief to prepare the upcoming response. Indeed, we observe a concomitant increase in BOLD activity in the fronto-parietal network and compression of SD_{BOLD} in the DMN in high performing subjects. How functional connectivity between different brain networks changes over the course of learning goes beyond the scope of the current study, but constitutes an interesting target for future research.

Standard analytic approaches reveal different neural uncertainty correlates compared to SD_{BOLD}

Our standard GLM analysis revealed that BOLD modulation by uncertainty also related to estimation accuracy in our task. On one hand, higher BOLD (in lateral parietal cortex, dorsomedial PFC, ventro- and rostrolateral PFC, and anterior insula) with lower uncertainty predicted higher estimation accuracy. In other words, high performers showed an increase in BOLD activity in these regions as they observed more samples and became more certain about the jar marble ratio on a given trial. These areas correspond to fronto-parietal control and dorsal attention brain networks, which support goal-directed behavior in externally-driven tasks [1,63-66]. Conversely, *positive* coupling between the BOLD signal and state uncertainty in mainly in the vmPFC (and a cluster in the left hippocampus) also predicted better performance in our task, which has previously been linked to confidence representations [67] (see Supplementary Information for detailed discussion of these findings).

Interestingly, we found that uncertainty representations in SD_{BOLD} (in DMN regions) and parametric BOLD signal modulation (in fronto-parietal control/attention regions) were spatially different and *uniquely* predicted estimation accuracy. Why would the brain track uncertainty in these different neural signals and across different networks? One possibility could be that SD_{BOLD} tracks uncertainty about the latent environmental state with or without immediate relevance for decision-making, while parametric BOLD modulation by uncertainty may be more relevant in the context of goal-directed decision formation and action selection. In support of this view, previous research has shown modulation of brain signal variability by stimulus features even when no action was required. For example, in the human neuroimaging literature, changes in neural variability have been observed in response to varying complexity of visual stimuli during passive viewing, which were linked to offline cognitive performance [41,45]. In a similar vein, it has recently been argued that early perceptual uncertainty, which has been the main focus of sampling accounts of neural variability, is tracked irrespective of task demands but can be flexibly utilized in higher-order decision-making [6-8,68]. Thus, changes in brain signal variability may track environmental variables without the need for immediate action. In contrast, parametric BOLD modulation by uncertainty has been reported for learning and decision-making tasks that require participants to make choices on every trial [2,23-37]. As such, these neural uncertainty correlates may be more tightly linked to decision formation and may in some cases even relate to metacognitive awareness of this decision-relevant variable [67,69,70]. Although speculative, this framework makes testable predictions for situations in which we would expect to see a dissociation of these two neural signals, which should be addressed in future work.

Limitations and next steps

One potential limitation of our task design was that our stimulus sample size manipulation introduced trial-to-trial variance in how much uncertainty could be reduced with each presented sample (Figure S1B). Because our SD_{BOLD} measure was computed for each sample period across trials, uncertainty levels were by definition not equivalent within each condition bin (although uncertainty always decreased across sample periods). However, our results showed that SD_{BOLD} effects could not be explained by parametric BOLD signal modulation by trial-to-trial differences in uncertainty trajectories in the same brain regions, which could have accounted for this variance. Ideally, future studies should keep sample size consistent between trials so that SD_{BOLD} can be computed over comparable uncertainty levels.

Another potential limitation is that uncertainty was collinear with within-trial time in our task; the more samples were presented, the more uncertainty could be reduced in general. However, this is an inevitable consequence of decision-making and learning in stable environments, in which one can simply accumulate evidence over time. Although within-trial time could have introduced arousal- or attention-related effects (which have previously been linked to human brain signal variability [13]), we observed performance-related interindividual variability in SD_{BOLD} compression over the sampling period. Such an effect cannot, by definition, be accounted for by the fixed factor of trial time. Nevertheless, future work could investigate the coupling between brain signal variability and uncertainty in non-stationary learning environments (see e.g. [3]), within which uncertainty can increase *or* decrease over time.

Conclusion

We provide first evidence that moment-to-moment brain signal variability compresses with increasing belief precision during learning. Whether brain signal compression is directly proportional to the informativeness of the available evidence is an important prediction that should be investigated in future work.

Acknowledgements and Funding

During the work on his dissertation, A.S. was a pre-doctoral fellow of the International Max Planck Research School on the Life Course (LIFE, www.imprs-life.mpg.de; participating institutions: Max Planck Institute for Human Development, Freie Universität Berlin, Humboldt-Universität zu Berlin, University of Michigan, University of Virginia, University of Zurich) and the International Max Planck Research School on Computational Methods in Psychiatry and Ageing Research (COMP2PSYCH, https://www.mps-ucl-centre.mpg.de/comp2psych). D.D.G. received an Emmy Noether Programme grant from the German Research Foundation. D.D.G. is affiliated with the Max Planck UCL Centre for Computational Psychiatry and Aging Research.

References

1 Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* **106**, 1125-1165 (2011). https://doi.org:10.1152/jn.00338.2011

2 Bach, D. R. & Dolan, R. J. Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat Rev Neurosci* **13**, 572-586 (2012). https://doi.org:10.1038/nrn3289

3 Bruckner, R., Heekeren, H. R. & Nassar, M. R. Understanding Learning Through Uncertainty and Bias. (2022). https://doi.org:10.31234/osf.io/xjkbg

4 Ma, W. J. & Jazayeri, M. Neural Coding of Uncertainty and Probability. *Annu Rev Neurosci* **37**, 205-220 (2014). https://doi.org:10.1146/annurev-neuro-071013-014017

5 Jacobs, R. A. & Kruschke, J. K. Bayesian learning theory applied to human cognition. *Wiley Interdiscip. Rev.: Cogn. Sci.* **2**, 8-21 (2011). https://doi.org:10.1002/wcs.80

6 Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci* **14**, 119-130 (2010). https://doi.org:10.1016/j.tics.2010.01.003

7 Orbán, G., Berkes, P., Fiser, J. & Lengyel, M. Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex. *Neuron* **92**, 530-543 (2016). https://doi.org:10.1016/j.neuron.2016.09.038

8 Berkes, P., Orbán, G., Lengyel, M. & Fiser, J. Spontaneous Cortical Activity Reveals Hallmarks of an Optimal Internal Model of the Environment. *Science* **331**, 83-87 (2011). https://doi.org:10.1126/science.1195870

9 Echeveste, R., Aitchison, L., Hennequin, G. & Lengyel, M. Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. *Nat Neurosci* **23**, 1138-1149 (2020). https://doi.org:10.1038/s41593-020-0671-1

10 Festa, D., Aschner, A., Davila, A., Kohn, A. & Coen-Cagli, R. Neuronal variability reflects probabilistic inference tuned to natural image statistics. *Nat Commun* **12**, 3635 (2021). https://doi.org:10.1038/s41467-021-23838-x

11 Friston, K., Breakspear, M. & Deco, G. Perception and self-organized instability. *Front Comput Neurosc* **6**, 44 (2012). https://doi.org:10.3389/fncom.2012.00044

12 Kosciessa, J. Q., Lindenberger, U. & Garrett, D. D. Thalamocortical excitability adjustments guide human perception under uncertainty. *Biorxiv*, 2020.2006.2022.165118 (2020). https://doi.org:10.1101/2020.06.22.165118

13 Waschke, L., Kloosterman, N. A., Obleser, J. & Garrett, D. D. Behavior needs neural variability. *Neuron* **109**, 751-766 (2021). https://doi.org:10.1016/j.neuron.2021.01.023

14 Garrett, D. D. *et al.* Moment-to-moment brain signal variability: A next frontier in human brain mapping? *Neurosci Biobehav Rev* **37**, 610-624 (2013). https://doi.org:10.1016/j.neubiorev.2013.02.015

15 Garrett, D. D. *et al.* Amphetamine modulates brain signal variability and working memory in younger and older adults. *Proc National Acad Sci* **112**, 7593-7598 (2015). https://doi.org:10.1073/pnas.1504090112

16 Armbruster-Genç, D. J. N., Ueltzhöffer, K. & Fiebach, C. J. Brain Signal Variability Differentially Affects Cognitive Flexibility and Cognitive Stability. *J Neurosci* **36**, 3978-3987 (2016). https://doi.org:10.1523/jneurosci.2517-14.2016

17 Garrett, D. D., Kovacevic, N., McIntosh, A. R. & Grady, C. L. The Modulation of BOLD Variability between Cognitive States Varies by Age and Processing Speed. *Cereb Cortex* **23**, 684-693 (2013). https://doi.org:10.1093/cercor/bhs055

18 Grady, C. L. & Garrett, D. D. Understanding variability in the BOLD signal and why it matters for aging. *Brain Imaging Behav* **8**, 274-283 (2014). https://doi.org:10.1007/s11682-013-9253-0

19 Palva, S., Kulashekhar, S., Hämäläinen, M. & Palva, J. M. Localization of Cortical Phase and Amplitude Dynamics during Visual Working Memory Encoding and Retention. *J Neurosci* **31**, 5013-5025 (2011). https://doi.org:10.1523/jneurosci.5592-10.2011

20 Rouhinen, S., Panula, J., Palva, J. M. & Palva, S. Load Dependence of β and γ Oscillations Predicts Individual Capacity of Visual Attention. *J Neurosci* **33**, 19023-19033 (2013). https://doi.org:10.1523/jneurosci.1666-13.2013

21 Grundy, J. G., Barker, R. M., Anderson, J. A. E. & Shedden, J. M. The relation between brain signal complexity and task difficulty on an executive function task. *Neuroimage* **198**, 104-113 (2019). https://doi.org:10.1016/j.neuroimage.2019.05.045

22 Honkanen, R., Rouhinen, S., Wang, S. H., Palva, J. M. & Palva, S. Gamma Oscillations Underlie the Maintenance of Feature-Specific Information and the Contents of Visual Working Memory. *Cereb Cortex* **25**, 3788-3801 (2015). https://doi.org:10.1093/cercor/bhu263

23 Walker, E. Y. *et al.* Studying the neural representations of uncertainty. *Arxiv* (2022). https://doi.org:10.48550/arxiv.2202.04324
24 Huettel, S. A., Song, A. W. & McCarthy, G. Decisions under Uncertainty: Probabilistic Context Influences Activation of Prefrontal and Parietal Cortices. *J Neurosci* **25**, 3304-3311 (2005). https://doi.org:10.1523/jneurosci.5070-04.2005
25 Vilares, I., Howard, James D., Fernandes, Hugo L., Gottfried, Jay A. & Kording, Konrad P. Differential Representations of Prior and Likelihood Uncertainty in the Human Brain. *Curr Biol* **22**, 1641-1648 (2012). https://doi.org:10.1016/j.cub.2012.07.010
26 Grinband, J., Hirsch, J. & Ferrera, V. P. A Neural Representation of Categorization Uncertainty in the Human Brain. *Neuron* **49**, 757-763 (2006). https://doi.org:10.1016/j.neuron.2006.01.032
27 Michael, E., Gardelle, V. d., Nevado-Holgado, A. & Summerfield, C. Unreliable Evidence: 2 Sources of Uncertainty During Perceptual Choice. *Cereb Cortex* **25**, 937-947 (2015). https://doi.org:10.1093/cercor/bht287
28 Trudel, N. *et al.* Polarity of uncertainty representation during exploration and exploitation in ventromedial prefrontal cortex. *Nat Hum Behav* **5**, 83-98 (2021). https://doi.org:10.1038/s41562-020-0929-3
29 Badre, D., Doll, Bradley B., Long, Nicole M. & Frank, Michael J. Rostrolateral Prefrontal Cortex and Individual Differences in Uncertainty-Driven Exploration. *Neuron* **73**, 595-607 (2012). https://doi.org:10.1016/j.neuron.2011.12.025
30 Tomov, M. S., Truong, V. Q., Hundia, R. A. & Gershman, S. J. Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nat Commun* **11**, 2371 (2020). https://doi.org:10.1038/s41467-020-15766-z
31 Meyniel, F. & Dehaene, S. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proc National Acad Sci* **114**, E3859-E3868 (2017). https://doi.org:10.1073/pnas.1615773114
32 Meyniel, F. Brain dynamics for confidence-weighted learning. *Plos Comput Biol* **16**, e1007935 (2020). https://doi.org:10.1371/journal.pcbi.1007935
33 Bach, D. R., Hulme, O., Penny, W. D. & Dolan, R. J. The Known Unknowns: Neural Representation of Second-Order Uncertainty, and Ambiguity. *J Neurosci* **31**, 4811-4820 (2011). https://doi.org:10.1523/jneurosci.1452-10.2011
34 Payzan-LeNestour, E., Dunne, S., Bossaerts, P. & O'Doherty, John P. The Neural Representation of Unexpected Uncertainty during Value-Based Decision Making. *Neuron* **79**, 191-201 (2013). https://doi.org:10.1016/j.neuron.2013.04.037
35 McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* **84**, 870-881 (2014). https://doi.org:10.1016/j.neuron.2014.10.013
36 Nassar, M. R., Bruckner, R. & Frank, M. J. Statistical context dictates the relationship between feedback-related EEG signals and learning. *Elife* **8**, e46975 (2019). https://doi.org:10.7554/eLife.46975
37 Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat Neurosci* **10**, 1214-1221 (2007). https://doi.org:10.1038/nn1954
38 Garrett, D. D., McIntosh, A. R. & Grady, C. L. Brain Signal Variability is Parametrically Modifiable. *Cereb Cortex* **24**, 2931-2940 (2014). https://doi.org:10.1093/cercor/bht150
39 Mumford, J. A., Turner, B. O., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* **59**, 2636-2643 (2012). https://doi.org:10.1016/j.neuroimage.2011.08.076
40 Krishnan, A., Williams, L. J., McIntosh, A. R. & Abdi, H. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *Neuroimage* **56**, 455-475 (2011). https://doi.org:10.1016/j.neuroimage.2010.07.034
41 Waschke, L. *et al.* Single-neuron spiking variability in hippocampus dynamically tracks sensory content during memory formation in humans. *Biorxiv*, 2023.2002.2023.529684 (2023). https://doi.org:10.1101/2023.02.23.529684
42 Spiliopoulos, L. & Hertwig, R. Variance, Skewness and Multiple Outcomes in Described and Experienced Prospects: Can One Descriptive Model Capture It All? *J Exp Psychology Gen* **152**, 1188-1222 (2022). https://doi.org:10.1037/xge0001323
43 Spiliopoulos, L. & Hertwig, R. Noisy retrieval of experienced probabilities underlies rational judgment of uncertain multiple events. *Manuscript in preparation* (2023).

44 Grady, C. L. & Garrett, D. D. Brain signal variability is modulated as a function of internal and external demand in younger and older adults. *Neuroimage* **169**, 510-523 (2018). https://doi.org:10.1016/j.neuroimage.2017.12.031
45 Garrett, D. D., Epp, S., Kleemeyer, M., Lindenberger, U. & Polk, T. A. Higher performers upregulate brain signal variability in response to more feature-rich visual input. *Neuroimage* **217**, 116836 (2020). https://doi.org:10.1016/j.neuroimage.2020.116836
46 Griffin, D. & Brenner, L. Perspectives on probability judgment calibration. *Blackwell handbook of judgment and decision making*, 177-199 (2004).
47 Gottlieb, D. A., Weiss, T. & Chapman, G. B. The Format in Which Uncertainty Information Is Presented Affects Decision Biases. *Psychol Sci* **18**, 240-246 (2006). https://doi.org:10.1111/j.1467-9280.2007.01883.x
48 Camilleri, A. R. & Newell, B. R. The role of representation in experience-based choice. *Judgm Decis Mak* **4**, 518-529 (2009). https://doi.org:10.1017/s1930297500001315
49 Glaze, C. M., Filipowicz, A. L. S., Kable, J. W., Balasubramanian, V. & Gold, J. I. A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nat Hum Behav* **2**, 213-224 (2018). https://doi.org:10.1038/s41562-018-0297-4
50 Olschewski, S. & Scheibehenne, B. What's in a Sample? Epistemic Uncertainty and Metacognitive Awareness in Risk Taking. (2023). https://doi.org:10.31234/osf.io/hf8dm
51 Greicius, M. D., Krasnow, B., Reiss, A. L. & Menon, V. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proc National Acad Sci* **100**, 253-258 (2003). https://doi.org:10.1073/pnas.0135058100
52 Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc National Acad Sci* **102**, 9673-9678 (2005). https://doi.org:10.1073/pnas.0504136102
53 Damoiseaux, J. S. *et al.* Consistent resting-state networks across healthy subjects. *Proc National Acad Sci* **103**, 13848-13853 (2006). https://doi.org:10.1073/pnas.0601417103
54 Raichle, M. E. The Brain's Default Mode Network. *Annu Rev Neurosci* **38**, 1-15 (2013). https://doi.org:10.1146/annurev-neuro-071013-014030
55 Shulman, G. L. *et al.* Common Blood Flow Changes across Visual Tasks: I. Increases in Subcortical Structures and Cerebellum but Not in Nonvisual Cortex. *J Cognitive Neurosci* **9**, 624-647 (1997). https://doi.org:10.1162/jocn.1997.9.5.624
56 Raichle, M. E. *et al.* A default mode of brain function. *Proc National Acad Sci* **98**, 676-682 (2001). https://doi.org:10.1073/pnas.98.2.676
57 Utevsky, A. V., Smith, D. V. & Huettel, S. A. Precuneus Is a Functional Core of the Default-Mode Network. *J Neurosci* **34**, 932-940 (2014). https://doi.org:10.1523/jneurosci.4227-13.2014
58 Buckner, R. L. & Carroll, D. C. Self-projection and the brain. *Trends Cogn Sci* **11**, 49-57 (2007). https://doi.org:10.1016/j.tics.2006.11.004
59 Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc National Acad Sci* **98**, 4259-4264 (2001). https://doi.org:10.1073/pnas.071043098
60 Spreng, R. N. & Grady, C. L. Patterns of Brain Activity Supporting Autobiographical Memory, Prospection, and Theory of Mind, and Their Relationship to the Default Mode Network. *J Cognitive Neurosci* **22**, 1112-1123 (2010). https://doi.org:10.1162/jocn.2009.21282
61 Garrett, D. D., Epp, S. M., Perry, A. & Lindenberger, U. Local temporal variability reflects functional integration in the human brain. *Neuroimage* **183** (2018). https://doi.org:10.1016/j.neuroimage.2018.08.019
62 Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W. & Schacter, D. L. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage* **53**, 303-317 (2010). https://doi.org:10.1016/j.neuroimage.2010.06.016
63 Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* **3**, 201-215 (2002). https://doi.org:10.1038/nrn755
64 Dosenbach, N. U. F. *et al.* Distinct brain networks for adaptive and stable task control in humans. *Proc National Acad Sci* **104**, 11073-11078 (2007). https://doi.org:10.1073/pnas.0704320104
65 Pessoa, L., Kastner, S. & Ungerleider, L. G. Neuroimaging Studies of Attention: From Modulation of Sensory Processing to Top-Down Control. *J Neurosci* **23**, 3990-3998 (2003). https://doi.org:10.1523/jneurosci.23-10-03990.2003
66 Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E. & Buckner, R. L. Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *J Neurophysiol* **100**, 3328-3342 (2008). https://doi.org:10.1152/jn.90355.2008

67 Vaccaro, A. G. & Fleming, S. M. Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain Neurosci Adv* **2**, 2398212818810591 (2018). https://doi.org:10.1177/2398212818810591

68 Koblinger, Á., Fiser, J. & Lengyel, M. Representations of uncertainty: where art thou? *Curr Opin Behav Sci* **38**, 150-162 (2021). https://doi.org:10.1016/j.cobeha.2021.03.009

69 Martino, B. D., Fleming, S. M., Garrett, N. & Dolan, R. J. Confidence in value-based choice. *Nat Neurosci* **16**, 105-110 (2012). https://doi.org:10.1038/nn.3279

70 Morales, J., Lau, H. & Fleming, S. M. Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. *J Neurosci* **38**, 3534-3546 (2018). https://doi.org:10.1523/jneurosci.2360-17.2018

SUPPLEMENTARY INFORMATION

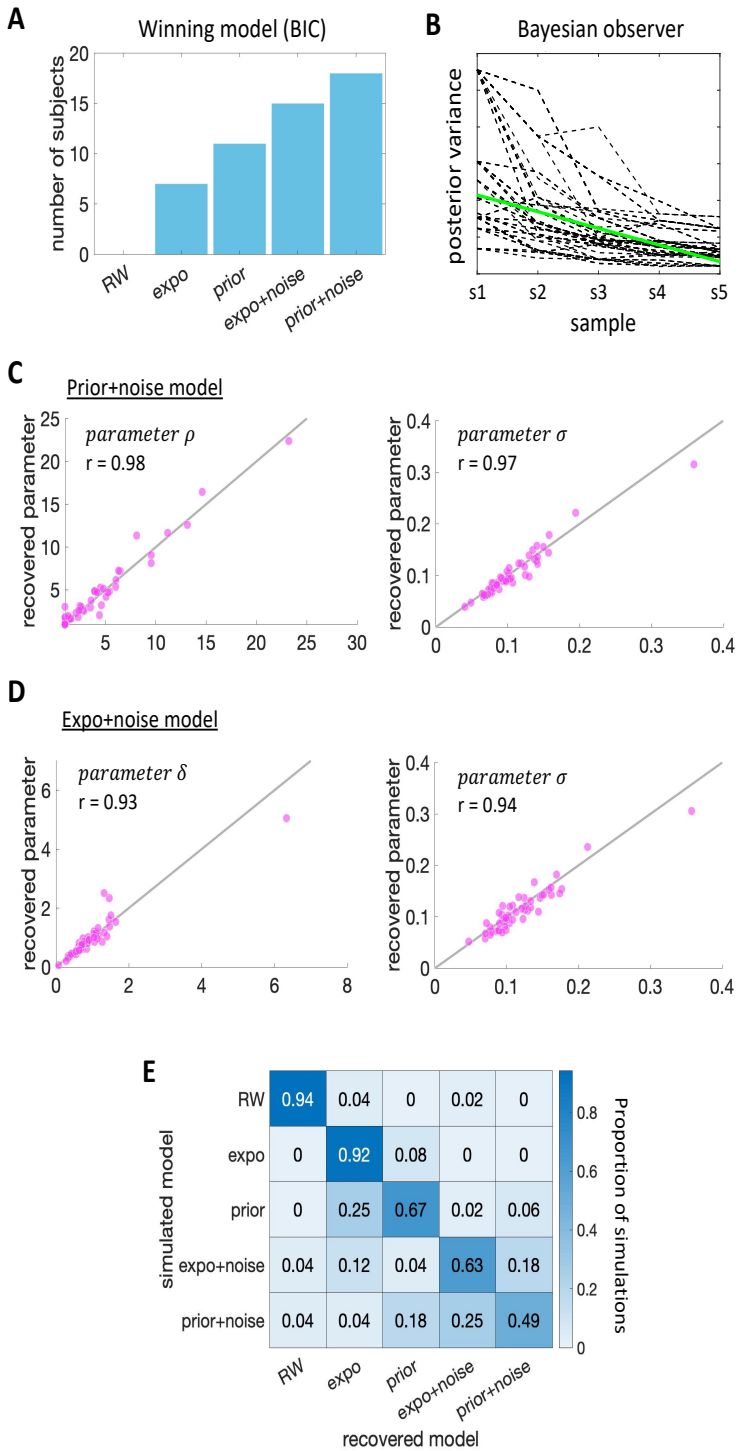

Figure S1. Model comparison and simulations. A: Counts of models with the lowest BIC across subjects. B: Trial-wise trajectories of the posterior variances of the marble ratio belief distribution for the unbiased Bayesian observer model. C: Parameter recovery for the Bayesian observer model with free prior width ρ and noisy choice σ. D: Parameter recovery for the Bayesian observer model with free evidence weight δ and noisy choice σ. E: Proportion of models recovered from simulated subjects for each model. *Note.* RW = Rescorla-Wagner model, expo = Bayesian observer with exponential evidence weight δ and sampling choice rule, prior = Bayesian observer model with fitted prior width ρ and sampling choice rule, expo+noise = same as expo model but with noisy choice σ, prior+noise = same as prior but with noisy choice σ

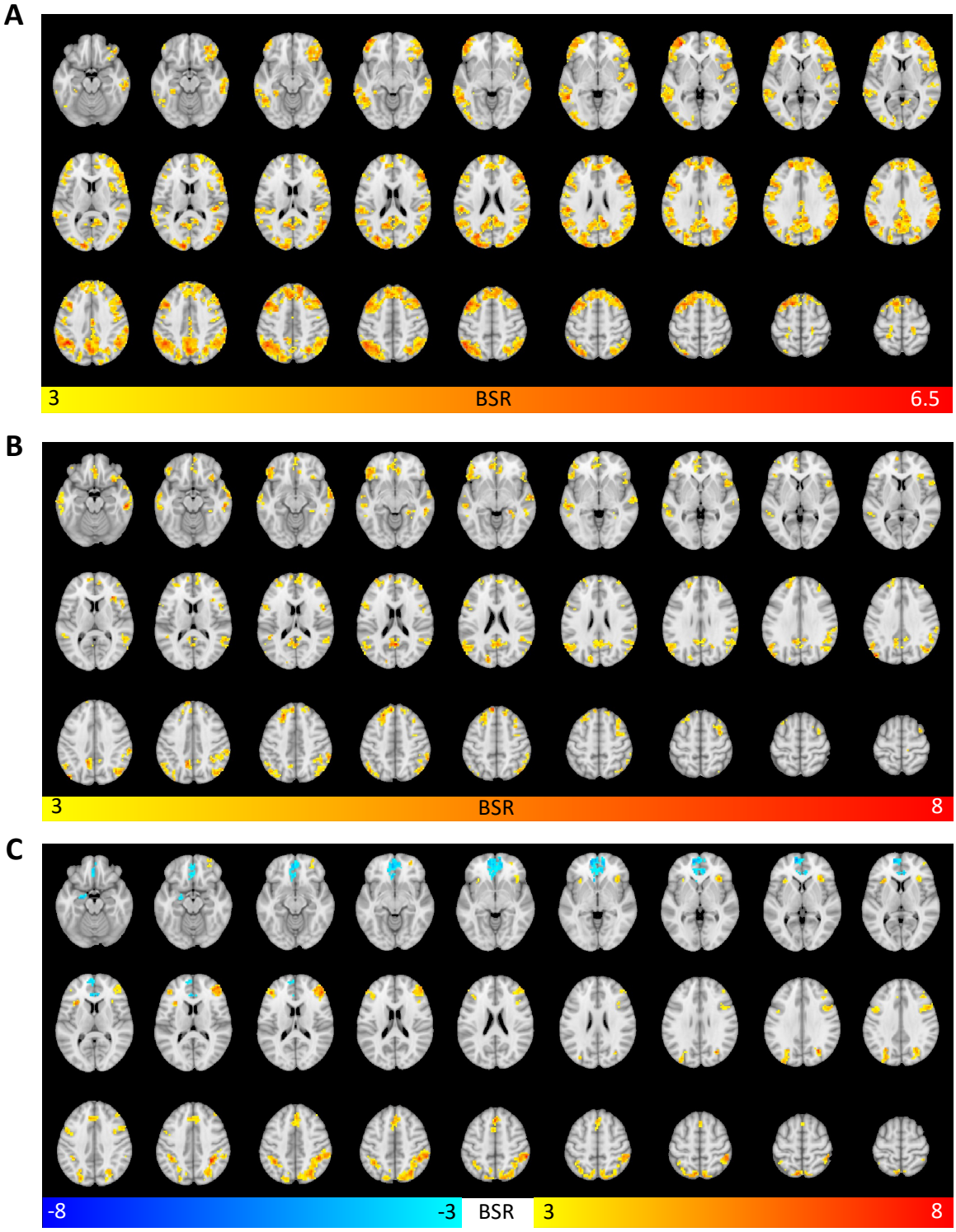

Figure S2. Axial brain plots for the PLS results. A: Task PLS model relating SD_{BOLD} to sample presentations. B: Behavioral PLS model relating ΔSD_{BOLD} to median estimation error (N = 47). C: Behavioral PLS model relating BOLD signal uncertainty modulation to median estimation error (N = 47). Bootstrap ratio = BSR.

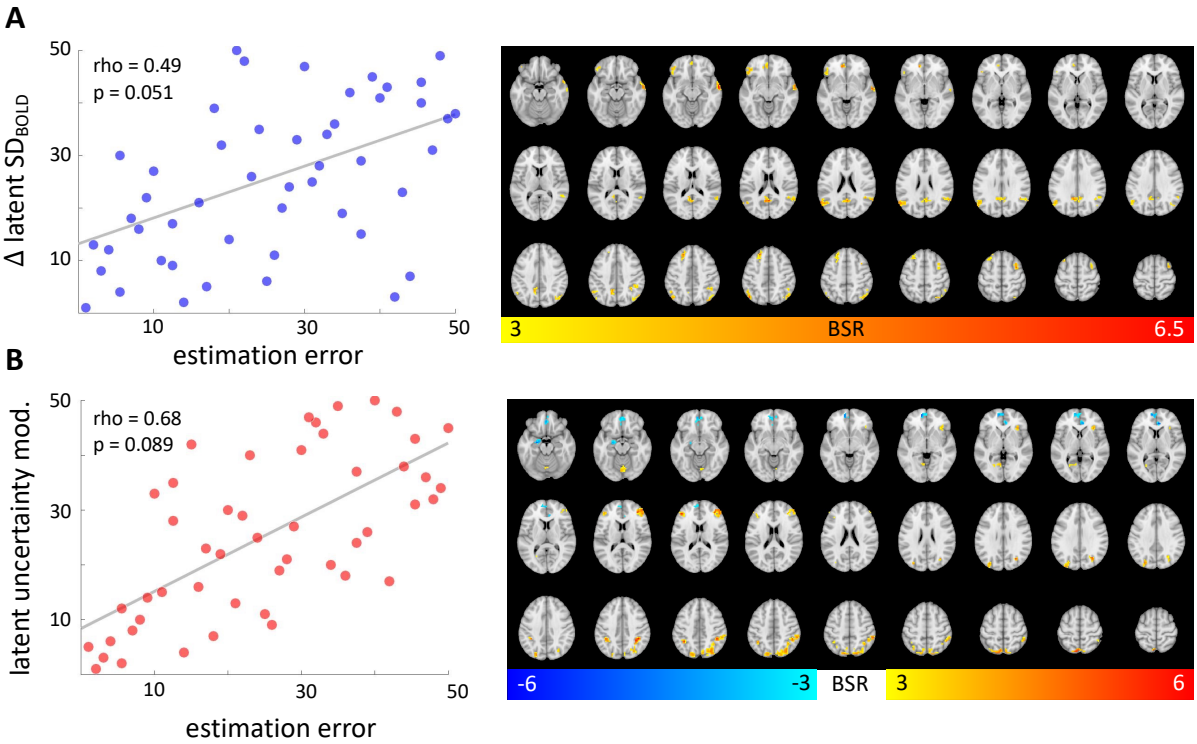

Figure S3. Behavioral PLS results for the full dataset (N = 51). A: Behavioral PLS model relating ΔSD_{BOLD} to average estimation error. B: Behavioral PLS model relating BOLD signal uncertainty modulation to median estimation error. Note: The results are presented for rank scored variables. P-values are based on nonparametric permutation tests.

Supplementary Table S1. Overview of the 18 marble jars. Each participant completed each jar four times (twice in the low reward and twice in the high reward condition).

Jar	Jar type	p(blue)	Payoff gamble	EV gamble	EV certain	EV difference
1	risk seeking	0.2	55	11	10	1
2	risk seeking	0.1	130	13	10	3
3	risk seeking	0.4	28	11.2	10	1.2
4	risk seeking	0.6	22	13.2	10	3.2
5	risk seeking	0.9	12	10.8	10	0.8
6	risk seeking	0.8	16	12.8	10	2.8
7	risk neutral	0.1	100	10	10	0
8	risk neutral	0.2	50	10	10	0
9	risk neutral	0.25	40	10	10	0
10	risk neutral	0.77	13	10	10	0
11	risk neutral	0.83	12	10	10	0
12	risk neutral	0.91	11	10	10	0
13	risk averse	0.1	92	9.2	10	-0.8
14	risk averse	0.2	35	7	10	-3
15	risk averse	0.6	15	9	10	-1
16	risk averse	0.4	17	6.8	10	-3.2
17	risk averse	0.8	11	8.8	10	-1.2
18	risk averse	0.9	8	7.2	10	-2.8

Note. EV = Expected value, p(blue) = blue marble probability

Supplementary Table S2. Model recovery. BIC of each model for each simulated dataset.

Recovered model:	**RW**	**expo**	**prior**	**expo+noise**	**prior+noise**
RW	-3508.163	-1984.135	1560.267	-2694.659	-2633.512
expo	-2828.659	-4926.453	-3708.676	-4223.386	-4018.178
prior	-3706.85	-5704.678	-6375.963	-5297.432	-5702.832
expo+noise	-2888.764	-3287.136	-2035.45	-4396.865	-4037.069
prior+noise	-3207.927	-3669.239	-3177.083	-4293.817	-4770.482

Supplementary Table S3: *Peak voxel coordinates in robust clusters (BSR > 3) for the significant latent variable of the task PLS analysis relating SD_{BOLD} to sample presentation.*

	MNI					
X	**Y**	**Z**	**BSR**	**p-value**	**Cluster size (voxels)**	**Anatomical region**
-39	-54	36	-6.4734	0.0000	4086	L Inferior Parietal Lobule
-33	51	0	-6.3604	0.0000	488	L Middle Frontal Gyrus
-45 (-45)	12 (9)	54 (54)	-6.2607	0.0000	3370	L Middle Frontal Gyrus*
-9	-96	12	-5.9355	0.0000	384	L Superior Occipital Gyrus
60	-24	-3	-5.2715	0.0000	189	R Middle Temporal Gyrus
3	-6	66	-4.5180	0.0000	36	R Posterior-Medial Frontal
-30	-39	-12	-4.5173	0.0000	25	L Fusiform Gyrus
15	48	0	-4.3964	0.0000	58	R Superior Medial Gyrus
45	-6	-3	-4.2319	0.0000	27	R Insula Lobe
-24	-39	66	-4.1320	0.0000	46	L Postcentral Gyrus
18 (22)	-24 (-24)	63 (63)	-4.0587	0.0000	52	R Precentral Gyrus*

Note. Abbreviations: MNI = Coordinates in Montreal Neurological Institute space, BSR = Bootstrap ratio. Anatomic labels were derived from the cytoarchitectonic atlas provided in the anatomy toolbox for SPM 8 (Eickhoff et al., 2005). We only report clusters of 25 voxels or more. *Peak voxel not assigned anatomical label in anatomy toolbox. Anatomical label corresponds to bracketed coordinates.

Supplementary Table S4: *Peak voxel coordinates in robust clusters (BSR > 3) for the significant latent variable of the behavioral PLS analysis relating ΔSD_{BOLD} to estimation error (N = 47).*

MNI						
X	**Y**	**Z**	**BSR**	**p-value**	**Cluster size (voxels)**	**Anatomical region**
57	3	-24	7.6886	0.0000	180	R Middle Temporal Gyrus
3	-57	18	7.5207	0.0000	257	R Precuneus
-45	-78	33	7.4781	0.0000	275	Area PGp (IPL)
-21	33	48	7.3272	0.0000	187	L Superior Frontal Gyrus
33	27	9	6.6384	0.0000	26	R Insula Lobe
48	-69	33	5.9344	0.0000	180	R Angular Gyrus
-9	66	18	5.8743	0.0000	178	L Superior Medial Gyrus
-39	42	-6	5.8357	0.0000	169	L IFG (p. Orbitalis)
60	-45	48	5.7765	0.0000	327	R Inferior Parietal Lobule
-15	-87	21	5.7644	0.0000	32	L Superior Occipital Gyrus
-54	-27	-3	5.7185	0.0000	51	L Middle Temporal Gyrus
36	3	57	5.5382	0.0000	84	R Middle Frontal Gyrus
51	6	15	5.5042	0.0000	107	R Rolandic Operculum
36	18	-18	5.4959	0.0000	42	R Insula Lobe
12	-24	69	5.2755	0.0000	39	R Paracentral Lobule
24	-36	-6	5.1858	0.0000	25	R ParaHippocampal Gyrus
-9	57	39	5.1066	0.0000	53	L Superior Frontal Gyrus
-45	-45	15	4.9372	0.0000	51	L Superior Temporal Gyrus
-51	12	18	4.9333	0.0000	43	L IFG (p. Opercularis)
-66 (-63)	-27 (-31)	-24 (-20)	4.8241	0.0000	133	L Inferior Temporal Gyrus*
15	36	51	4.8089	0.0000	45	R Superior Frontal Gyrus
-33	60	18	4.6589	0.0000	50	L Middle Frontal Gyrus
9	60	15	4.5953	0.0000	54	R Superior Medial Gyrus
42	57	15	4.1974	0.0000	70	R Middle Frontal Gyrus

Note. Abbreviations: MNI = Coordinates in Montreal Neurological Institute space, BSR = Bootstrap ratio. Anatomic labels were derived from the cytoarchitectonic atlas provided in the anatomy toolbox for SPM 8 (Eickhoff et al., 2005). We only report clusters of 25 voxels or more. *Peak voxel not assigned anatomical label in anatomy toolbox. Anatomical label corresponds to bracketed coordinates.

Supplementary Table S5: *Peak voxel coordinates in robust clusters (BSR > 3) for the significant latent variable of the behavioral PLS analysis relating uncertainty modulation to estimation error (N = 47).*

	MNI					
X	**Y**	**Z**	**BSR**	**p-value**	**Cluster size (voxels)**	**Anatomical region**
57	-39	51	7.5835	0.0000	636	R Inferior Parietal Lobule
0	27	51	6.7004	0.0000	172	L Superior Medial Gyrus
48	39	15	6.5347	0.0000	202	R IFG (p. Triangularis)
-36	-48	45	6.3783	0.0000	365	L Inferior Parietal Lobule
-42	39	15	6.1012	0.0000	49	L IFG (p. Triangularis)
-33	18	9	5.8193	0.0000	40	L Insula Lobe
36	24	0	5.3671	0.0000	59	R Insula Lobe
51	12	33	5.0933	0.0000	71	R IFG (p. Opercularis)
30	54	-9	4.5326	0.0000	27	R Middle Orbital Gyrus
-48	6	36	4.4879	0.0000	36	L Precentral Gyrus
-9	60	3	-5.4550	0.0000	391	L Superior Medial Gyrus
-18	-6	-21	-4.3491	0.0000	35	L Hippocampus

Note. Abbreviations: MNI = Coordinates in Montreal Neurological Institute space, BSR = Bootstrap ratio. Anatomic labels were derived from the cytoarchitectonic atlas provided in the anatomy toolbox for SPM 8 (Eickhoff et al., 2005). We only report clusters of 25 voxels or more.

SUPPLEMENTARY METHODS

Task design. During the final experimental block, performed outside the MR scanner, confidence ratings of the estimation of the blue-red marble ratio were collected. For this block of trials, a rating scale was included directly after completion of the estimation and prior to the gambling phase. Participants were instructed to indicate their confidence ranging from "very unsure" to "very sure" about the estimation of the blue-red marble ratio on a continuous confidence judgment scale [1]. Additionally, the amount of obtainable reward on each trial was manipulated. Each marble was marked with one ring (low reward) or two rings (high reward). In the high reward condition, the magnitude of the reward in the gamble was doubled compared to the low reward condition. Each participant completed each jar an equal number of times in the low and high reward conditions.

The gambling phase. At the end of each trial, participants had to make a risky decision between two options. One option represented a draw from the current urn. If the outcome had been a blue marble, the participant would have received a payoff with a specific magnitude (which is different in each trial, see details below). The other option represented a reference option with a certain payoff (i.e., reward probability was 100 percent). After the experiment finished, one of these gambles was picked and participants received a bonus payment based on the result of a draw from their chosen option. The options were presented in the left or right hemi-field of the screen and the order was pseudo-randomized across the experiment. Participants should choose one option via left or right button of the button box. Each trial ended following the decision and subsequently the next trial started with a different urn and reward probability.

The 18 different marble jars in the experiment differed in expected value in the following way: In six jars the expected value (probability of drawing a blue marble times the varying payoff) was larger than the reference option (ranging between 10 and 30 percent), in another six urns the expected value of both options were approximately equal and in the final six urns the expected value was lower than the reference option (ranging between -10 and -30 percent, Supplementary Table S1). The probability of drawing a blue marble varied between 0.1 and 0.91, whereas the payoff varied between 8 and 130 points. Theoretically, a decision-maker that always chooses the gambling option could expect the same summed expected value as a decision maker that always chooses the certain reference option.

SUPPLEMENTARY RESULTS

Model and parameter recovery. The ground truth model was recovered well from each simulated dataset. The respective ground truth model consistently had the lowest BIC overall (supplementary Table S2). Across the simulated subjects, the ground truth model was always recovered most frequently (Supplementary Figure S1E). However, there was some model confusion, particularly for our winning model (fitted prior with noisy choice rule), which was only recovered in about 49% of simulated subjects. This may also explain the variance in empirical model fit on the subject level, although the majority of subjects (35%) were still best fit by the overall winning Bayesian observer model (Supplementary Figure S1A). Some degree of model confusion is to be expected since all variants of our Bayesian observer model aim to account for the same behavioral pattern (biased estimation of extreme marble proportions). Certain parameterizations of the prior model family may thus yield simulated behavior that is also well accounted for by models from the exponential family. Furthermore, some confusion between nested models is not necessarily surprising. For example, certain ranges of the response noise parameter σ may result in similar behavioral predictions to that of the same model that samples responses from the final beta distribution. It is noteworthy, however, that in general models including a noisy response rule are well differentiable from models with a sampling response rule. This supports the inclusion of the noise parameter σ in our winning model. There was also good differentiation between the Rescorla-Wagner model and all Bayesian observer models on the subject level supporting our inference that subjects seem to use uncertainty to inform their marble ratio estimates.
We also checked the recoverability of model parameters. For our winning model, both prior (Pearson's r = 0.98) and noise (Pearson's r = 0.97) parameters estimated for each simulated subject were highly correlated with their ground truth values (Supplementary Figure S1C). Parameters of the second-best model were also recoverable (Supplementary Figure S1D). This reinforces the inferences made from individuals' parameter estimates in our results.

Explaining the effect of response noise parameter σ on extreme jar bias. The significant main effect of response noise σ in predicting subjects' extreme jar bias suggests that people with opposite extreme jar biases (i.e., more misestimation for jars with marble ratios close to 50:50), is not captured by the winning behavioral model. Indeed, if we remove subjects with opposite extreme jar bias (i.e., negative effect slopes, N = 9) from

this regression analysis, the main effect of noise parameter σ on extreme jar bias is not significant (t(39) = -0.113, p = 0.922). Thus, our behavioral model accounts well for the empirical response pattern we sought to capture (i.e., positive effect slopes).

Analysis of BOLD signal modulation by uncertainty. Our PLS model relating BOLD signal modulation by uncertainty (i.e., standard GLM beta estimates for the parametric uncertainty regressor) to subjects' median estimation error revealed a significant latent association (permuted p = $3.40 \cdot 10^{-2}$, Spearman's r = 0.675, Figure 5B, cf. Figure S3B for N = 51). Because brain saliences were both positive and negative, individual differences in parametric modulation effects related differently to estimation errors between regions (see Figure S2C for full axial brain plots and Table S5 for peak voxel coordinates in robust clusters). In robust clusters with positive saliences, subjects who down-modulated BOLD activity more with increasing state uncertainty made fewer estimation errors (lower latent uncertainty modulation rank scores reflecting more negative modulation effects) than subjects who did not track state uncertainty in these regions (higher latent uncertainty modulation rank scores reflecting BOLD signal modulation by state uncertainty effects close to zero). Robust clusters were mainly located in the lateral parietal cortex, PFC (dorsomedial, ventrolateral, and rostrolateral parts), and in the insula. In contrast, in robust clusters with negative brain saliences, subjects who showed more positive coupling between BOLD signal and state uncertainty made fewer estimation errors than subjects with a negative coupling. This relationship was mainly expressed in a cluster located in the ventromedial PFC (vmPFC) and a cluster in the left hippocampus.

We also investigated whether these performance-related individual differences in BOLD signal modulation by uncertainty could be explained by idiosyncratic uncertainty representations, as captured by our winning behavioral model. We ran multiple regression predicting latent uncertainty modulation of the BOLD response, which relates to task accuracy, from the prior width parameter ρ controlling for response noise σ. This model explained a significant amount of variance (F(2,44) = 6.244, p = $4.10 \cdot 10^{-3}$) with an R^2 of 0.221 (Figure 5C). There was a significant main effect for prior width ρ (t(44) = 3.508, p = $1.05 \cdot 10^{-3}$, semi-partial η^2 = 0.218) but not for parameter σ (t(44) = 1.590, p = 0.119). This finding suggests that people with suboptimal uncertainty representations, due to a narrow prior belief distribution (i.e., higher ρ), show less BOLD signal modulation by state uncertainty trajectories derived from the unbiased Bayesian observer model. We again ran this analysis also for the parameters of the (second-best) evidence weight model, which afforded similar inferences (see supplementary results).

Brain-behavior results for the evidence weight model. Our alternative model, which explains suboptimal responding by an exponential weighting of the incoming evidence, would similarly predict individual differences in state uncertainty reduction during the sampling phase: People who reduce uncertainty less for large samples (i.e., δ < 1) make more estimation errors than people who weight the information provided by large samples more optimally (i.e., δ = 1). Subjects' fitted prior parameter ρ of the winning model and fitted evidence weight δ for second-best model (both including a response noise parameter σ) were highly correlated (Pearson's r(49) = -0.863, p = $3.73 \cdot 10^{-16}$), supporting the notion that they capture similar behavioral response patterns. For this model, the fitted evidence weight parameter δ had a median of 0.85 (SD = 0.84) and the fitted noise parameter σ had a median of 0.10 (SD = 0.05).

Predicting ranked latent ΔSD_{BOLD} from the ranked evidence weight δ and response noise σ parameters in a multiple regression model explained a significant amount of variance (F(2,44) = 6.938, p = $2.41 \cdot 10^{-3}$) with an R^2 of 0.240. There was a significant main effect only for evidence weight δ (t(44) = -3.725, p = $5.53 \cdot 10^{-4}$, semi-partial η^2 = -0.240): People who underweight the evidence of large samples more (i.e. δ < 1) show less SD_{BOLD} collapse during the sampling phase. Furthermore, the regression model predicting latent uncertainty modulation of the BOLD response from evidence weight parameter δ and response noise σ explained a significant amount of variance (F(2,44) = 5.544, p = $7.12 \cdot 10^{-3}$) with an R^2 of 0.201. There was a significant main effect for evidence weight δ (t(44) = -3.257, p = $2.17 \cdot 10^{-3}$, semi-partial η^2 = -0.193) and also for parameter σ (t(44) = 2.147, p = $3.74 \cdot 10^{-2}$). This suggests that people who underweight the evidence of large samples (δ < 1) show less BOLD signal modulation by uncertainty trajectories of the unbiased Bayesian observer.

Overall, our interpretation that individual differences in uncertainty representations explain the observed brain-behavior relationships would not have changed had we chosen this alternative model. We frame our conclusions in terms of the fitted prior model because it provided a better account of behavior.

SUPPLEMENTARY DISCUSSION

Our standard analysis approach revealed that BOLD signal modulation by state uncertainty predicted task accuracy. On one hand, higher BOLD signal (in lateral parietal cortex, dorsomedial PFC, ventro- and rostrolateral PFC, and anterior insula) with lower uncertainty predicted higher estimation accuracy. In other words, higher performers showed an increase in BOLD activity in these regions as they observed more samples and became more certain about the jar marble ratio on a given trial. These areas correspond to frontoparietal control and dorsal attention brain networks, which support goal-directed behavior in externally-driven tasks [2-6]. These networks of brain areas have also been found to correlate with state uncertainty in other inference tasks [7,8]. However, these prior studies commonly report an opposite effect direction in which high state uncertainty reflected *more* activity in these brain regions. For example, an fMRI study by McGuire, et al. [7] investigated how uncertainty drives learning in a task that required participants to infer the position of an unseen helicopter based on bag drops that followed a normal distribution around the helicopter's true position. In their task, the helicopter location could change unannounced leading to learning rate adjustments due to change-point uncertainty and uncertainty about the helicopter location in a given environment (which the authors term relative uncertainty). The authors found BOLD signal modulation by relative uncertainty in lateral parietal cortex, dorsomedial PFC, ventrolateral PFC, and anterior insula, which match the regions we found in the current study. However, the direction of effect is reversed showing a positive coupling between relative uncertainty and the BOLD signa. Two key differences of our study are that we investigate state inference in a stable rather than dynamic environment and that decisions are only required after observing several evidence samples. In these respects, our study shares similarities with perceptual decision-making studies, which assume that noisy perceptual evidence is integrated over time to arrive at a decision [9,10]. Previous fMRI studies of perceptual decision-making also report the involvement of a similar set of higher-order brain regions, which have been linked to evidence accumulation, decision formation and response preparation [11-18]. Notably, some studies have reported an *increase* in average BOLD activity in these brain areas over the sampling period (i.e., with decreasing state uncertainty), which aligns with our findings [11,19].

Conversely, *positive* coupling between the BOLD signal and state uncertainty in vmPFC and left hippocampus also predicted better performance in our task. The previously mentioned study by McGuire, et al. [7] also reported an effect in the vmPFC, but again with opposite effect direction compared to the one we see in high performers. At first glance, our findings appear at odds with previous work reporting vmPFC tracking of subjective confidence. However, recent work by Trudel, et al. [20] suggests that representations of uncertainty in this brain region depend on the behavioral goal. In their task, participants had to learn the predictiveness of two choice options in determining a target location that was later revealed in the trial. Early on during learning, vmPFC BOLD signal positively tracked the uncertainty differences of the two choice options and correlated with uncertainty-guided exploration. Later on, vmPFC BOLD signal negatively tracked the uncertainty difference and correlated with uncertainty-avoidant exploitation. The positive effect we observe in our study is thus in line with the general idea of an "exploratory brain mode" that supports uncertainty-guided learning. Overall, the GLM results directly connect to previous studies across various domains of decision-making and reveal surprising discrepancies that require more attention in future work.

References

1 Hebart, M. N., Schriever, Y., Donner, T. H. & Haynes, J.-D. The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. *Cereb Cortex* **26**, 118-130 (2016). https://doi.org:10.1093/cercor/bhu181

2 Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* **3**, 201-215 (2002). https://doi.org:10.1038/nrn755

3 Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* **106**, 1125-1165 (2011). https://doi.org:10.1152/jn.00338.2011

4 Dosenbach, N. U. F. *et al.* Distinct brain networks for adaptive and stable task control in humans. *Proc National Acad Sci* **104**, 11073-11078 (2007). https://doi.org:10.1073/pnas.0704320104

5 Pessoa, L., Kastner, S. & Ungerleider, L. G. Neuroimaging Studies of Attention: From Modulation of Sensory Processing to Top-Down Control. *J Neurosci* **23**, 3990-3998 (2003). https://doi.org:10.1523/jneurosci.23-10-03990.2003

6 Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E. & Buckner, R. L. Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *J Neurophysiol* **100**, 3328-3342 (2008). https://doi.org:10.1152/jn.90355.2008

7 McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* **84**, 870-881 (2014). https://doi.org:10.1016/j.neuron.2014.10.013

8 Payzan-LeNestour, E., Dunne, S., Bossaerts, P. & O'Doherty, John P. The Neural Representation of Unexpected Uncertainty during Value-Based Decision Making. *Neuron* **79**, 191-201 (2013). https://doi.org:10.1016/j.neuron.2013.04.037

9 Gold, J. I. & Shadlen, M. N. The Neural Basis of Decision Making. *Annu Rev Neurosci* **30**, 535-574 (2007). https://doi.org:10.1146/annurev.neuro.29.051605.113038

10 Heekeren, H. R., Marrett, S. & Ungerleider, L. G. The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* **9**, 467-479 (2008). https://doi.org:10.1038/nrn2374

11 Ploran, E. J. *et al.* Evidence Accumulation and the Moment of Recognition: Dissociating Perceptual Recognition Processes Using fMRI. *J Neurosci* **27**, 11912-11924 (2007). https://doi.org:10.1523/jneurosci.3522-07.2007

12 Heekeren, H. R., Marrett, S., Bandettini, P. A. & Ungerleider, L. G. A general mechanism for perceptual decision-making in the human brain. *Nature* **431**, 859 862 (2004). https://doi.org:10.1038/nature02966

13 Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A. & Ward, B. D. Neural correlates of sensory and decision processes in auditory object identification. *Nat Neurosci* **7**, 295-301 (2004). https://doi.org:10.1038/nn1198

14 Thielscher, A. & Pessoa, L. Neural Correlates of Perceptual Choice and Decision Making during Fear–Disgust Discrimination. *J Neurosci* **27**, 2908-2917 (2007). https://doi.org:10.1523/jneurosci.3024-06.2007

15 Grinband, J., Hirsch, J. & Ferrera, V. P. A Neural Representation of Categorization Uncertainty in the Human Brain. *Neuron* **49**, 757-763 (2006). https://doi.org:10.1016/j.neuron.2006.01.032

16 Liu, T. & Pleskac, T. J. Neural correlates of evidence accumulation in a perceptual decision task. *J Neurophysiol* **106**, 2383-2398 (2011). https://doi.org:10.1152/jn.00413.2011

17 Ivanoff, J., Branning, P. & Marois, R. fMRI Evidence for a Dual Process Account of the Speed-Accuracy Tradeoff in Decision-Making. *Plos One* **3**, e2635 (2008). https://doi.org:10.1371/journal.pone.0002635

18 Krueger, P. M. *et al.* Evidence accumulation detected in BOLD signal using slow perceptual decision making. *J Neurosci Meth* **281**, 21-32 (2017). https://doi.org:10.1016/j.jneumeth.2017.01.012

19 Ploran, E. J., Tremel, J. J., Nelson, S. M. & Wheeler, M. E. High Quality but Limited Quantity Perceptual Evidence Produces Neural Accumulation in Frontal and Parietal Cortex. *Cereb Cortex* **21**, 2650-2662 (2011). https://doi.org:10.1093/cercor/bhr055

20 Trudel, N. *et al.* Polarity of uncertainty representation during exploration and exploitation in ventromedial prefrontal cortex. *Nat Hum Behav* **5**, 83-98 (2021). https://doi.org:10.1038/s41562-020-0929-3