Sophistication of Human Adaptive Probability Learning in Dynamic Environments

Wai Ying Chung (waiying.chung.0@gmail.com)[1,2]

Florent Meyniel (florent.meyniel@gmail.com)[1,2]

[1]Cognitive Neuroimaging Unit, NeuroSpin (INSERM-CEA), University of Paris-Saclay, Gif-sur-Yvette, France
[2]Institut de Neuromodulation, GHU Paris, Psychiatrie et Neurosciences, Hôpital Saint-Anne, Paris, France

Abstract

Maintaining accurate beliefs in a changing and noisy environment is a challenging computational problem. Previous studies have shown that humans adapt their learning dynamically, especially in the face of change. This conclusion is mostly supported in the context of magnitude learning (e.g., tracking a reward amount, an object position), and currently remains more uncertain in the case of probability learning (e.g., tracking the probability of an event occurring). Here, we initiate an open benchmarking approach to uncover the computations humans use for probability learning. We compared a wide range of models—including optimal Bayesian models, suboptimal variants, and simple prediction error-based update rules, using several datasets in which participants provided trial-by-trial probability estimates. Bayesian inference often outperformed simple prediction error-based models, despite being more computationally demanding and often considered less biologically plausible. Furthermore, inference strategies appear to depend on environmental volatility: under moderate volatility, an optimal Bayesian model best explains behavior, whereas in more stable environments, a simpler Bayesian approximation is better. These results so far highlight the sophistication of human adaptive learning for probability and suggest that humans can adapt their inference strategies based on environmental context. We invite others to contribute models and datasets to this benchmark to refine these conclusions.

Keywords: Adaptive learning; Bayesian inference; Model comparison; Probability learning

Introduction

How do humans update their beliefs in an abruptly changing, noisy environment? When faced with uncertain and dynamic conditions, learners should in theory continuously adjust their expectations based on new observations, balancing past knowledge with incoming evidence in an adaptive (i.e. dynamic) manner (Bruckner et al., 2025). What computational mechanisms support this adaptive learning process? Numerous models have been proposed to solve inference problems in dynamic, noisy environments. These models range from fully Bayesian (a.k.a. probabilistic), to the simple non-adaptive delta rule. The empirical evidence supporting the existence of adaptive learning in human is clear in magnitude learning, such as when tracking the latent mean of symbolic numbers (Nassar et al., 2012), positions (McGuire et al., 2014; Nassar et al., 2016), or angular directions (Vaghi et al., 2017)displayed in a sequence. These studies showed that human subjects weigh prior expectations and information from each observation according to their uncertainty, in a way that aligns with Bayesian inference. However, the biological plausibility of Bayesian inference is highly debated (Bowers and Davis, 2012; Findling et al., 2021; Fiser et al., 2010; Knill and Pouget, 2004), due to the algorithmic complexity of Bayesian inference in general (Cooper, 1990).

These findings on magnitude learning may not generalize to probability learning (such as tracking the probability of a stimulus in a sequence) since both types of learning have computational differences. In magnitude learning, the trial-by-trial rate of learning is primarily driven by the change-point probability, while in probability learning, it is primarily driven by prior uncertainty about the learned estimate (Foucault and Meyniel, 2024). This difference arises because the information conveyed by each observation about the latent estimate is larger for magnitude learning than for probability learning.

Experimental findings on adaptive probability learning remain limited. Several studies on reward probability have found evidence for an adaptive adjustment of learning based on environmental volatility (Behrens et al., 2007; Meder et al., 2017). However, these studies only compared average learning rates across blocks of trials, rather than examining trial-level adjustments in learning rates. Some decision-making studies manipulating probabilities used trial-by-trial modeling (Browning et al., 2015; Iglesias et al., 2013), but with the limit that the learning process may be partially confounded by decision-making processes. (By decision-making processes, we refer to processes that involve response selection, such as choosing between different actions or options that are associated with reward probabilities.) It is commonly found in decision-making research that subjects overestimate and underestimate extreme probabilities (Oprea and Vieider, 2024; Wulff et al., 2018), which is not the case in probability learning tasks devoid of choices and rewards (Gallistel et al., 2014).

In this study, our goal is to characterize the sophistication of human probability learning in the face of abrupt changes. To this end, we compared a broad range of existing models, ranging from the optimal Bayesian model to coarser and coarser approximations of it, as well as models that use simple update rules based on prediction errors. We used three openly available datasets corresponding to a probability learning task in which subjects continuously reported their estimate of the

abruptly changing hidden probability of stimuli presented in a sequence. We focused on "pure" probability learning tasks because they are less confounded by decision-making processes.

Methods

We aim to initiate an open benchmarking for probability learning in changing environments, with Python code for all models and openly available datasets. The code and the datasets are available on GitHub at `https://github.com/TheComputationalBrain/adaptive_prob_learning`. In the following, we describe the 10 models and 3 openly available datasets already included, used in this study. We hope that others will contribute new models and datasets.

Datasets: Probability learning in a changing environment

Common features. The three datasets used here shared similar experimental procedures (Fig 1). Participants viewed sequences made of two visual stimuli on a computer screen. On each trial, the stimulus was sampled from a Bernoulli distribution, whose parameter remained fixed between abrupt change points. After a change point, the new stimulus probability was resampled uniformly in a range that satisfied a change in the odds ratio (p/(1-p)) of at least 4 around the change point. A change point could occur independently on each trial with a small, fixed probability, and the changes were not signaled to the participants. Subjects were instructed about the generative process of sequences, and asked to estimate the latent probability of stimuli on each trial, which they report by moving a cursor on a slider.

Specific features of Gallistel et al. (2014) and Khaw et al. (2017). In Gallistel et al. (2014), stimuli were red and green rings, drawn in a self-paced sequence on a computer by participants. 10 subjects participated, each completing 10 sessions of 1,000 trials. The generative probability range was from 0 to 1 and the change-point probability was fixed at .005. In Khaw et al. (2017), the experimental procedure was identical to Gallistel et al. (2014), except that subjects were rewarded based on their performance. This dataset included 11 subjects, each completing 10 sessions of 1,000 trials (see supplementary Fig.1A for task illustration).

Specific features of Foucault and Meyniel (2024) The two visual stimuli were a blue circle and a yellow circle, presented in a sequence with 1.5 s interval between each observation. 96 participants each completed 15 sessions of 75 stimuli. Subjects were partly rewarded based on their performance, displayed as feedback at the end of each session. Performance was calculated based on the mean absolute error between subjective probability estimates and generative probabilities. The range of generative probability was from 0.1 to 0.9, and change-point probability was 0.05 (see supplementary Fig.1B for task illustration).

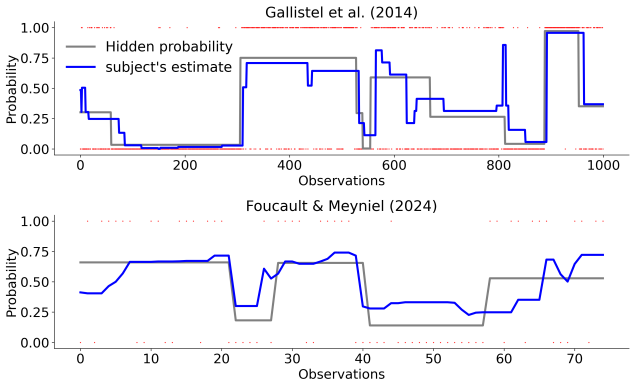

Figure 1: **Probability learning tasks.** Example sequences of observations (red dots) and probability estimates from Gallistel et al. (2014) and Foucault and Meyniel (2024).

Formal description of the models We selected 10 models to approximate subjects' probability estimates across three datasets. These models can be broadly categorized into two groups. The first group consists of different variants of Bayesian models, ranging from the optimal model that makes the most accurate inference of the latent probability (computed numerically here with a hidden Markov model, **HMM** (Behrens et al., 2007); its estimates are very close of the exact Bayesian online change point detection model (Adams and MacKay, 2007) — to models that provide increasingly simplified approximations of Bayesian inference. Such approximations can be variational, replacing some functions in the optimal model with other functions that simplify its computation, like the Hierarchical Gaussian Filter (**HGF**) (Mathys et al., 2011,0) and Volatile Kalman filter (**VKF**) (Piray and Daw, 2020,0). Other approximations simplify the optimal model itself, by reducing the number of variables to estimate. For instance, in theory the optimal problem can be solved by taking into account, on a given trial, all possible positions of the previous change points (parameterized as "run length" i.e. length of a subsequence devoid of change points). The possible number of run lengths being potentially very large, one can decide to approximate the problem by considering only a subset of possible run lengths, as in the **Mixture of Delta Rules** (Wilson et al., 2013), or only the mean run length as in the **Reduced Bayesian Model** (Nassar et al., 2010). Alternatively, instead of updating the current probability estimate on every trial, the optimal model can also be simplified by updating the model only occasionally, when a change point is detected as in the **Change Point Model** (Gallistel et al., 2014; Ricci and Gallistel, 2017).

Each of these approximations remains, to different degrees, computationally intensive. As a result, they are sometimes perceived as offering limited insight into the biological algorithms that the brain actually employs (Bowers and Davis, 2012; Findling et al., 2021; Knill and Pouget, 2004). The second group of models we consider does not attempt to approximate the optimal model, instead they directly aim for com-

putational simplicity. We included models that update their estimates simply based on the proportion of prediction error, without relying on a probabilistic framework. Examples of such models include the Proportional Integral Derivative controller (**PID**), which is a simple, yet very efficient adaptive filtering model (Ritz et al., 2018), the **Delta Rule** (Wagner and Rescorla, 1972) with a constant learning rate, and the **Pearce-Hall model** that adjusts the learning rate based on the unsigned prediction error which serves as an indicator of surprise (Pearce and Hall, 1980). This second group is computationally simpler, and it makes use of prediction errors, whose representation is well established in the brain (Arias-Carrión et al., 2010; Bayer and Glimcher, 2005; Pessiglione et al., 2006; Schultz, 2007) (Figure 2). For these reasons, this second class of models is often considered as more biologically plausible. The key features and number of free parameters for each model are summarized in Table 1, we refer the reader to the references under each model for a complete description.

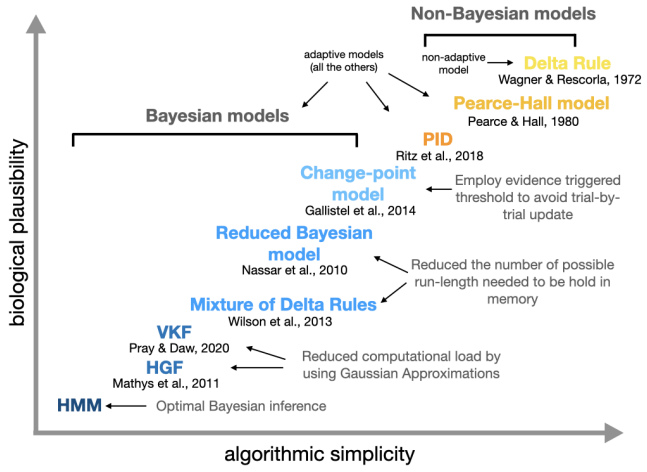

Figure 2: **A gradient of sophistication in probability learning models.** The models included in this paper can be categorized into two main groups. The x-axis represents model complexity, ranging from optimal Bayesian inference to progressively coarser approximations of Bayesian inference, then to non-Bayesian models. The y-axis represents the biological plausibility of these models. Models that require lower memory and computational demands are generally considered more biologically plausible. Note that biological plausibility is a complex and not easily quantifiable concept. This figure is intended to illustrate the general idea that models with high computational cost are often questioned in terms of their biological plausibility. The ordering of models along the sophistication gradient is qualitative and partly reflects our interpretation of the emphasis of the original authors.

Probability weighting function

The linear in log odds (LLO) function is widely used in decision-making research to describe how individuals subjectively distort a probability p (noted $w(p)$ when distorted) when making choices, particularly in risk-based decisions (Gonzalez and Wu, 1999; Bruhin et al., 2010):

$$\ln\left(\frac{w(p)}{1-w(p)}\right) = \ln(\delta) + \gamma \ln\left(\frac{p}{1-p}\right)$$

By fitting this function to our data, where p represents the ideal observer estimates, we quantified distortion of subjective probability. In the LLO framework, δ is the scaling factor that shifts the function up or down, $\delta > 1$ indicates that subjects generally overestimate probabilities; when $\delta < 1$, they generally underestimate probabilities. Meanwhile, γ controls the distortion of probability estimates. When $\gamma < 1$, the function has an inverse S-shape; When $\gamma > 1$,the function shows an S-shape. If δ =1, γ = 1, subjective estimates closely match objective probabilities.

Model recovery analysis

We performed a model recovery analysis to test if different models were behaviorally distinguishable from one another. To do this, we simulated 100 subjects, each containing 10 sessions of 1000 trials, following the same generating procedure as Gallistel et al. (2014), with a constant change point probability of 0.005. Then, we fitted all models to each simulated data set. The resulting model recovery matrix, showing the proportion of best-fit models across the 100 simulations, is presented in Fig. 3 (see Supp. Fig. 2 for parameter recovery). The results suggest that most models were behaviorally identifiable, except for the HGF. This result is likely due to numerical issues the HGF encountered in 6 out of 100 simulations, where the posterior variance over the top (volatility) level became negative. This issue arises from the Taylor approximation used to extrapolate the variational posterior and is a well-known problem of HGF (Mathys et al., 2014; Piray and Daw, 2020,0). For the Reduced Bayesian model, the model is better recovered by the Reduced Bayesian model with underweighted likelihood information. This outcome is unsurprising, as the two models are nested, with the latter including an additional free parameter to adjust the weight of the likelihood.

Model fit and comparison

Model fitting was performed in Python using scipy.optimize with the Powell optimization procedure. As the goal of model fitting was to create models that behave as similarly as possible to humans (rather than to best perform the task), we minimized the mean squared error (MSE) between model predictions and subjects' estimates. For model comparison, we evaluated the performance of each model using a cross-validation (CV) procedure. Specifically, for each subject, we trained the model on N-1 sessions and tested using the left-out session, resulting a 10-fold cross-validation for Gallistel et al.(2014) and Khaw et al.(2017) data (10 sessions of 1000 trials), and 15 folds in Foucault and Meyniel (2024)'s data (15 sessions of 75 trials). Then the final CV-MSE for each model for each subject is the sum of the CV-MSE over all the left-out folds used

Table 1. Model list

Model	Key features	Free parameters (#)
Hidden Markov model (HMM) (Behrens et al., 2007)	- numerical estimation of the optimal solution - use the Markov property for more efficient computation	1
Hierarchical Gaussian Filter (HGF) (Mathys et al., 2011, 2014)	- variational approximation using Gaussian distributions - distinguish multiple level hierarchically organized	3
Volatile Kalman filter (VKF) (Piray & Daw, 2021)	- Uses a Kalman-like filtering algorithm - the Kalman gain is adjusted based on volatility estimated trial-by-trial	3
Mixture of Delta rules (Wilson et al., 2013)	- consider a limited number of run lengths, specified as nodes in the model - the update rule for each node is a Delta rule	4
Reduced Bayesian model (Nassar et al., 2010)	- consider a single run length at a time - update the run length on each trial based on change point probability	1
Reduced Bayesian model with underweighted likelihood (Nassar et al., 2010)	- variant of the Reduced Bayesian model, with an extra parameter to limit the use of likelihood information in computing the probability of a change point	2
Change point model (Gallistel et al., 2014)	- use decision thresholds to update the estimate	2
Proportional integral derivative (PID) controller (Ritz et al., 2018)	- linear combination of 3 terms: Proportional (current prediction error), Integral (leaky integration of past errors) and Derivative (error derivative).	4
Pearce-Hall model (Pearce & Hall, 1980)	- The learning rate is adjusted based on the unsigned prediction error, which serves as an indication of how surprising or unexpected the outcome was.	2
Delta Rule (Wagner & Rescorla, 1972)	- A constant learning rate is used to updates based on the prediction error	1

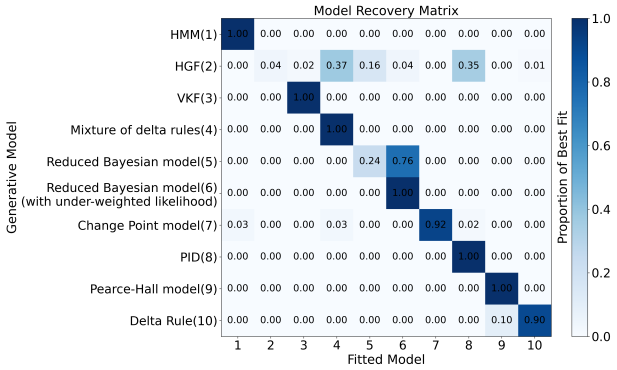

Figure 3: **Model recovery analysis.** Confusion matrix showing the proportion of simulated data sequences (out of 100 iterations) generated by a given model (x-axis) and best fitted by each model (y-axis).

for testing. The cross-validation procedure prevents overfitting and ensures a fair comparison of models with different numbers of free parameters.

Results

Accuracy of subjective probability

We first evaluated the accuracy of subjects' estimates by comparing their reported estimates to the ideal observer estimates derived from the optimal Bayesian model. Subjects' estimates were tightly correlated with the ideal observer estimates in all three datasets (Gallistel et al. (2014): Pearson r = 0.92 ± 0.008, t(9) = 120.14, $p < 10^{-16}$; Khaw et al. (2017): r = 0.88 ± 0.02, t(10) = 44.36, $p < 10^{-13}$; Foucault and Meyniel (2024): r = 0.80 ± 0.01, t(95) = 55.79, $p < 10^{-74}$ (Figure 4, blue dot).

We assessed the linear in log odds (LLO) probability weighting function to quantify distortions in subjects' reported probability estimates relative to the ideal observer estimates. The results across the three datasets confirmed the visual impression from Fig. 4 that distortions are very minor. Both and distortion parameters were very close to 1, with some minor bias toward overweighting small probabilities and underweighting the large probabilities (see Fig. 4, yellow line). This distortion is much less pronounced than the distortion typically found when subjects make a decision/action based on learned probability (for instance, see (Oprea and Vieider, 2024)). Together, these results indicate that subjects produced accurate probability estimates despite the presence of multiple, unpredictable change points. We also computed several linear regressions, one for each trial number i relative to the change point, using subjects' probability estimates as the dependent variable and the generative probabilities before and after the change point as predictors. This allowed us to examine the rate at which subjects adapted to the new generative probabilities after a change point. The results suggested that, in general, subjects tended to adapt quickly—within approximately 10 trials (see Fig. 5, black line).

Model comparison reveals the sophistication of human probability learning

We then compared a wide range of models to characterize the computations subtending human probability learning. Specifically, we fitted 10 models to each subjects' data from each of the three datasets and evaluated their performance using cross-validated mean squared error (MSE). Note that the models were not optimized to perform the task well, but rather to reproduce the observed human responses as accurately as possible. The results, summarized in Fig. 6, show the percentage of subjects in each dataset best fitted (i.e., with the lowest CV-MSE) by each model, along with the average CV-MSE difference for each pair of models.

In Foucault & Meyniel (2024), the HMM was the best-fitting model, accounting for 46.9% (45 out of 96) of subjects, followed by the PID, which was the best-fitting model for 24% (23 out of 96) of subjects. The HMM performed better than

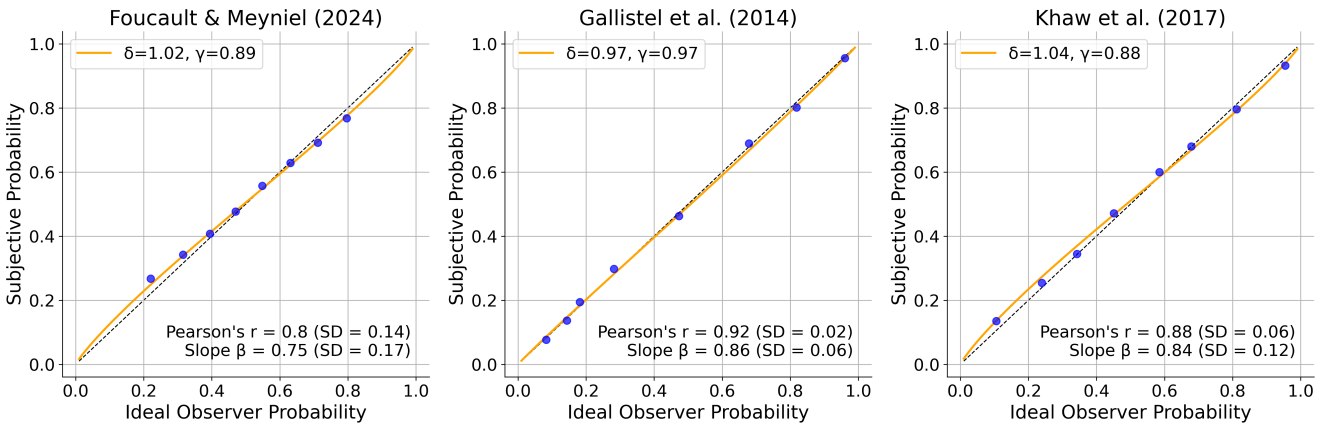

Figure 4: **Subjective probability estimates are remarkably accurate.** (Blue dots). The accuracy of subjects' estimates is assessed in comparison to the ideal observer estimates. The data were binned in 8 quantiles of the ideal observer estimate and averaged within-subjects for visualization purpose. Points and error bars (too small to be seen) show mean ± s.e.m. across subjects. (Yellow line). Probability weighting functions based on the median parameters of a linear in log-odds function estimated from subjects in each dataset. Overall the subjective reported probabilities match closely with the ideal observer probabilities. Dashed line is the identity line.

all other models, with significant CV-MSE differences with all models ($p < .001$ for all comparisons) but the PID model. In Gallistel et al. (2014), the Mixture of delta rules was the best-fitting model, accounting for 40% (4 out of 10) of subjects, followed by the HMM, which best fit 30% (3 out of 10) of subjects. The CV-MSE was significantly smaller in the Mixture of delta rules than in any other model, apart from the HMM and PID. For Khaw et al. (2017), the results are more mixed. The Mixture of delta rules was also the best-fitting model, accounting for 54.5% (6 out of 11) of subjects, followed by the VKF and the Reduced Bayesian model with under-weighted likelihood information, each best fitting 18.2% (2 out of 11) of subjects. However, in this dataset no model had a CV-MSE significantly smaller than most other models.

The results indicate that subjects' estimates are best captured by models that update their estimates on a trial-by-trial basis in a highly adaptive manner. This includes models that optimally compute Bayesian inference, such as the HMM, as well as those that use approximations, like the Mixture of Delta Rules, or the PID. In contrast, models that rely solely on error-driven updating, such as the Delta Rule and the Pearce-Hall model, generally provided a worse fit to subjects' data. Interestingly, the PID model, which is based on proportional error updates, demonstrated better predictive performance. Unlike the Delta Rule, which has a fixed learning rate and is therefore not adaptive, and the Pearce-Hall model, which adjusts its learning rate based only on the most recent unsigned prediction error, the PID model incorporates a leaky sum of all past errors and the rate of change in error. This filtering mechanism allows it to adapt flexibly to sudden changes in the environment.

We performed model predictive checks that capture different features of learning dynamics, and compared them to subjects' data, in order to have a better understanding of why some models perform better than some others beyond the quantitative model comparison based on mean-squared error. We estimated how quickly the estimate of subjects and models (fitted to behavior) become close to the new generative probability that follows a change point, and how quickly the generative probability that prevailed before the change point is forgotten (Figure 5). In the Foucault & Meyniel (2024) data, the HMM model was the only one to capture the very quick adaptation of the subject's estimate to the new generative probability. This very fast updating of the HMM is actually the reason why it performed worse than the Mixture of Delta Rules on the two other data sets, in which subjects adapted their estimates more slowly. Another model predictive check is the trial-by-trial learning rate, computed as the ratio between the update and the prediction (Supp Fig 3). This trial-by-trial learning rate, time-locked to change points, showed a marked transient increase in subjects in the Foucault & Meyniel (2024) dataset that was best captured only by the HMM model. By contrast, this transient increase was very small in the Gallistel et al. dataset, not compatible with HMM but instead with the Mixture of Delta Rules, and no increase in the Khaw et al. dataset.

Overall, these results suggest that subjects' estimates are best captured by models that dynamically adjust the learning rate on each trial.

Discussion

In this study, we trained a wide range of models to approximate subjects' data and compared their performance to gain insight into the computations underlying human probability learning in the face of latent changes. Our results provide evidence that probability learning is adapted dynamically on a

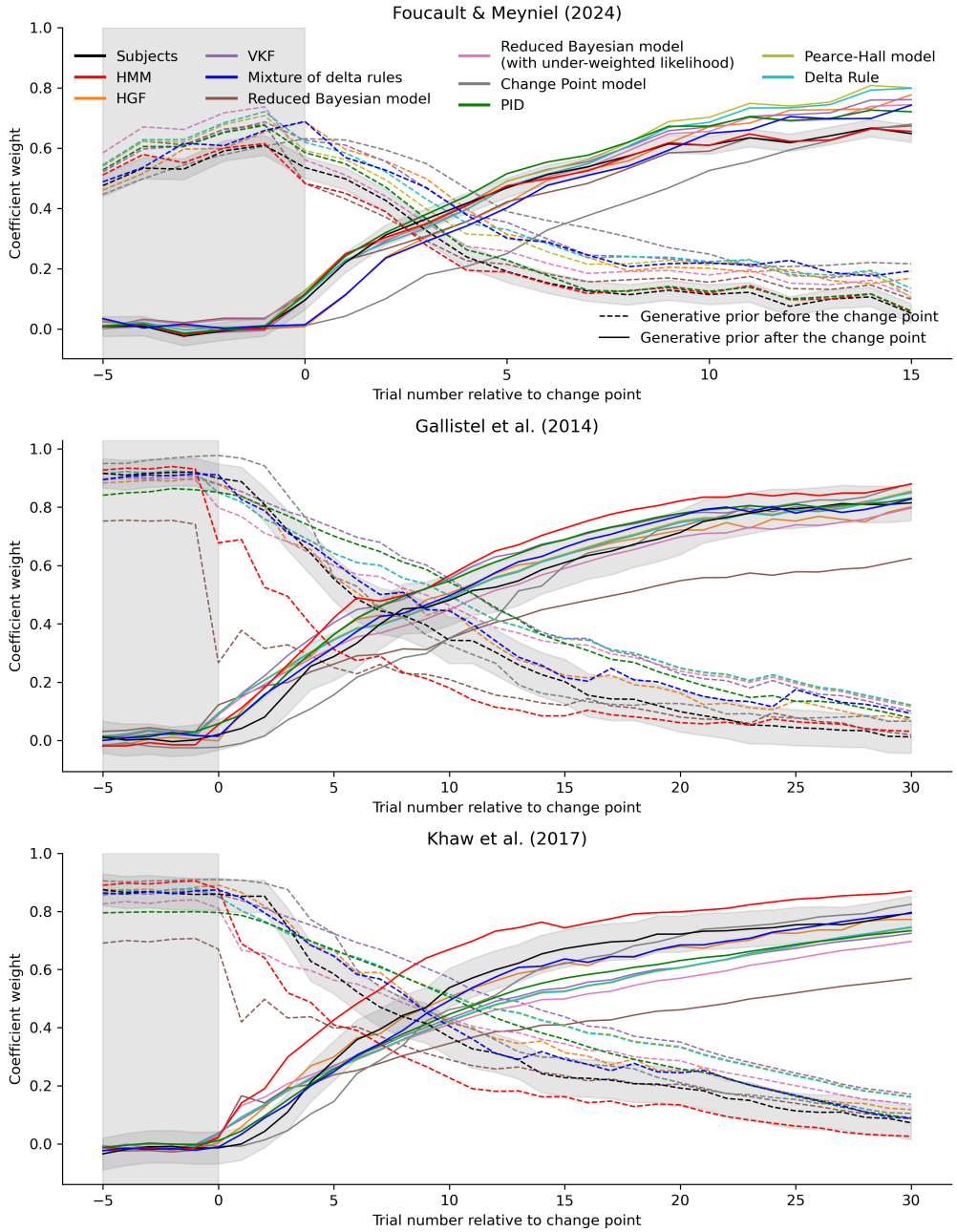

Figure 5: **Model predictive check: the speed of updating after a change point**. The estimates of each subject and each model were linearly regressed, on all trials having the same location relative to change points (e.g. 3 trials before a change point), onto two generative probabilities: the one that prevailed before the change point, and the one that follows the change point. The plot shows the regression weights as a function of trial number relative to change points, to reveal the learning dynamics. As a result of learning, the weights of the generative probability before and after the change point changed across time, as the models' estimates and the participants' estimates progressively align with the new generative probability. Error bars correspond to 95% CI computed across subjects.

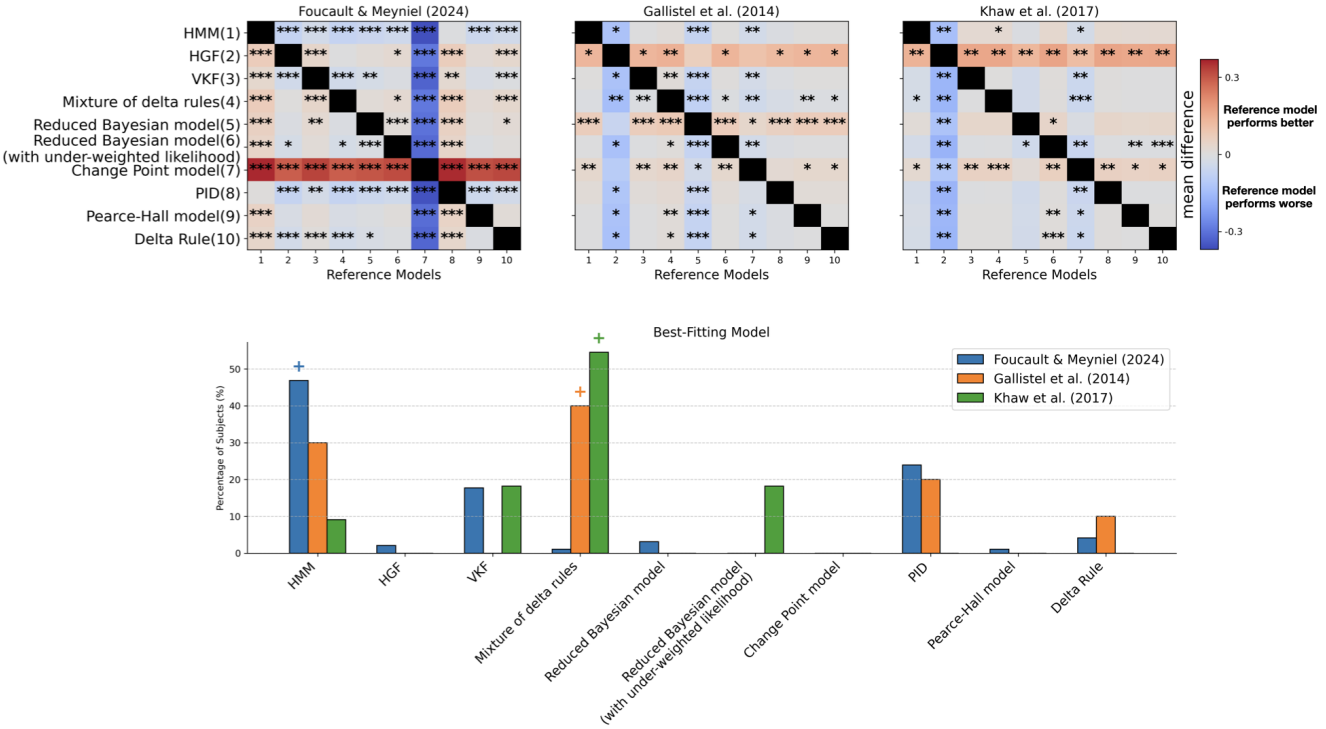

Figure 6: **Model comparison reveals the sophistication of human probability learning**. The upper panel shows the results of the pairwise comparison between each model. Positive values in the mean difference indicate that the reference model (on the x axis) performed better than the model on the y axis, and negative values indicating the reference model performs worse than the other (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ from a paired t-test). The lower panel shows the percentage of subjects best fitted by each model in each of the three datasets. "+" indicated the best-fitting model in each dataset.

trial-by-trial basis, since subjects' probability estimates were best explained by models featuring this property. Furthermore, we show that probability learning tasks in which subjects directly report probability estimates, without a concurrent decision-making task, induce minimal distortions in the reported estimates, in contrast to studies in the field of decision-making. This makes such paradigms particularly valuable for studying human probability learning.

Our results show that the best-fitting models employed trial-by-trial adjustments of the updating process, either by approximating Bayesian inference, or by using an adaptive filtering method (PID model). By contrast, models with fixed or weakly adaptive learning rates, such as the Delta rule (Wagner & Rescorla, 1972) and the Pearce-Hall model (Pearce & Hall, 1980), generally provided poor fits to subjects' data. Similarly, the Change Point model (Gallistel et al., 2014), which does not update estimates on a trial-by-trial basis, also failed to capture subjects' probability estimates effectively. Overall, our model comparison results suggest that the brain adaptively updates probability estimates based on each observation. A key distinction between Bayesian models and those using alternative updating rules is that the Bayesian framework inherently accounts for uncertainty/confidence about the estimate, as reflected by the variance of its prior distribution. When the model is highly uncertain about its estimate (i.e., the prior has a large variance), it updates its estimate more quickly, treating new observations as more informative. Conversely, when the model is more certain about its estimate (low variance), it updates it less, relying more on prior knowledge than new observations. Therefore, the result that Bayesian models provided a better fit to subjects' behavior in most cases suggests that the brain may compute this uncertainty, and that humans may adjust their learning dynamically on a trial-by-trial basis by incorporating uncertainty. The idea is further supported by previous studies showing that humans can report this uncertainty in a way that correlates with the uncertainty of the optimal Bayesian model (Meyniel et al., 2015). In addition, previous neuroimaging studies have identified brain correlates of this uncertainty (Iglesias et al., 2013; McGuire et al., 2014; Tomov et al., 2020).

Interestingly, the PID model performed similarly as the optimal model (HMM) in all datasets. This is a notable finding since this filtering model relies on updating the current estimate using the prediction error, which is computationally simple and biologically plausible (Arias-Carrión et al., 2010; Bayer & Glimcher, 2005). The update is done adaptively on a trial-by-trial basis by filtering previous prediction errors, obviating the need to perform costly computations as in Bayesian inference (even approximate ones). Such filtering algorithms are popular in the engineering field due to their simplicity and ac-

curacy (Ritz et al., 2018). One drawback is that the PID model is oblivious of the structure of the task, whereas Bayesian inference provides, in theory, a general solution to invert any generative model. In more complex tasks, with a richer latent structure, it is likely that the PID model would be surpassed by more complex models in capturing the subject's inference, since simple models may not generalize well (Benjamin et al., 2023; Foucault and Meyniel, 2021). The benchmarking we introduce should include more complex tasks in the future to better compare models.

In our results, we observed that different datasets are best-fitted by different models. Specifically, in Foucault & Meyniel (2024), an optimal Bayesian model (the HMM), provided the best fit to subjects' responses. While in Gallistel et al. (2014) and Khaw et al. (2017), the best-fitting model was the Mixture of Delta Rules. One reason for this difference could be that they these datasets have different volatility levels. In Foucault & Meyniel (2024), the probability of a change point in each session was set at 0.05 (one change point every 20 trials on average). By contrast, in Gallistel et al. (2014) and Khaw et al. (2017), the change point probability was much lower, at 0.005 (one change point every 200 trials on average). The higher volatility in Foucault & Meyniel (2024) may have necessitated the use of a close-to-optimal model, such as the HMM, which computes the full posterior distribution for the estimate at each time step to achieve better adaptability to frequent and abrupt changes. For the Mixture of Delta Rules which is the best-fitting model for Gallistel et al. (2014) and Khaw et al. (2017), it approximates Bayesian inference by retaining only a limited number of possible run lengths and updating their weights onto the final estimate, trial-by-trial, based on the probability that a change point has occurred. This approach reduces computational cost and memory load while sacrificing some adaptability compared to the optimal Bayesian model. In a more stable environment, a simplified approximation of Bayesian inference may be sufficient to provide accurate estimates, which would provide a better balance between computational cost and accuracy (Tavoni et al., 2022). If true, it means that humans flexibly adjust their inference strategies in response to environmental demands. In a recent paper, Verbeke and Verguts (2024) showed that human data are best fitted by different learning strategies depending on the complexity of the environment (e.g. whether optimal action and the associated reward change over time). Supporting this idea, previous neuroscience studies supported the existence of multiple computational systems that can be used depending on the complexity of the task, with an array of memory timescales allowing different subsets of neural populations to be selectively deployed according to the task demands (Bernacchia et al., 2011; Maheu et al., 2019; Meder et al., 2017).

The difference in results between datasets may also be partly explained by procedural differences in the tasks. For example, subjects in Foucault & Meyniel (2024) exhibited more frequent updates, whereas in Gallistel et al. (2014) and Khaw et al. (2017), subjects updated their reported estimate only occasionally. This may be due to specific aspects of the task designs: in Gallistel et al. (2014) and Khaw et al. (2017), updating the estimate incurred a time cost. In contrast, in Foucault & Meyniel (2024), observations occurred at regular intervals, and updating the reported estimates incurred no time cost, which may explain the step-like, occasional updates in the studies of Gallistel and Khaw (Forsgren et al., 2020).

We now discuss some limitations of this study. First, it is based on only three datasets. Ideally, we would like to incorporate more datasets using a similar experimental paradigm, including datasets with a richer task structure, in future studies. Second, the model comparison does not nail down the algorithm used by humans for probability learning. Doing so will require finer analyses of the data beyond the MSE, looking for specific properties of different possible algorithms (Palminteri et al., 2017). Neural recordings should also prove useful to check the existence of variables postulated by algorithmic models.

This study is intended as a preliminary step. All models and datasets are openly available, and we encourage others to contribute by sharing their datasets and model implementations. Our goal is to initiate an open benchmark project, inspired by other community-based projects such as the Algonauts (Cichy et al., 2019), to facilitate collaborative advances in understanding human probability learning.

Acknowledgments

This work was supported by funding from the European Research Council (ERC grant 947105) to F.M.

References

Adams, R. P. and MacKay, D. J. C. (2007). Bayesian Online Changepoint Detection. *arXiv preprint arXiv:0710.3742*.

Arias-Carrión, O., Stamelou, M., Murillo-Rodríguez, E., Menéndez-González, M., and Pöppel, E. (2010). Dopaminergic reward system: a short integrative review. *International Archives of Medicine*, 3(1):24.

Bayer, H. M. and Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, 47(1):129–141.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9):1214–1221.

Benjamin, L., Fló, A., Al Roumi, F., and Dehaene-Lambertz, G. (2023). Humans parsimoniously represent auditory sequences by pruning and completing the underlying network structure. *eLife*, 12:e86430.

Bernacchia, A., Seo, H., Lee, D., and Wang, X.-J. (2011). A reservoir of time constants for memory traces in cortical neurons. *Nature Neuroscience*, 14(3):366–372.

Bowers, J. S. and Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138(3):389–414.

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., and Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, 18(4):590–596.

Bruckner, R., Heekeren, H. R., and Nassar, M. R. (2025). Understanding learning through uncertainty and bias. *Communications Psychology*, 3(1):24.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412.

Cichy, R. M., Roig, G., Andonian, A., Dwivedi, K., Lahner, B., Lascelles, A., Mohsenzadeh, Y., Ramakrishnan, K., and Oliva, A. (2019). The algonauts project: A platform for communication between the sciences of biological and artificial intelligence. *arXiv preprint arXiv:1905.05675*.

Cooper, G. F. (1990). The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42(2-3):393–405.

Findling, C., Chopin, N., and Koechlin, E. (2021). Imprecise neural computations as a source of adaptive behaviour in volatile environments. *Nature Human Behaviour*, 5(1):99–112.

Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, 14(3):119–130.

Forsgren, M., Juslin, P., and Van Den Berg, R. (2020). Further Perceptions of Probability: In Defence of Associative Models – A Commentary on Gallistel et al. 2014.

Foucault, C. and Meyniel, F. (2021). Gated recurrence enables simple and accurate sequence prediction in stochastic, changing, and structured environments. *eLife*, 10:e71801.

Foucault, C. and Meyniel, F. (2024). Two Determinants of Dynamic Adaptive Learning for Magnitudes and Probabilities. *OPEN MIND*, 8:615–638.

Gallistel, C. R., Krishan, M., Liu, Y., Miller, R., and Latham, P. E. (2014). The Perception of Probability. *Psychological Review*, 121:99 – 123.

Gonzalez, R. and Wu, G. (1999). On the shape of the probability weighting function. *Cognitive psychology*, 38(1):129–166.

Iglesias, S., Mathys, C., Brodersen, K., Kasper, L., Piccirelli, M., den Ouden, H., and Stephan, K. (2013). Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. *Neuron*, 80(2):519–530.

Khaw, M. W., Stevens, L., and Woodford, M. (2017). Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*, 91:88–103.

Knill, D. C. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719.

Maheu, M., Dehaene, S., and Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *eLife*, 8:e41541.

Mathys, C., Daunizeau, J., Friston, K. J., and Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, 5.

Mathys, C., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., and Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience*, 8.

McGuire, J., Nassar, M., Gold, J., and Kable, J. (2014). Functionally Dissociable Influences on Learning Rate in a Dynamic Environment. *Neuron*, 84(4):870–881.

Meder, D., Kolling, N., Verhagen, L., Wittmann, M. K., Scholl, J., Madsen, K. H., Hulme, O. J., Behrens, T. E., and Rushworth, M. F. (2017). Simultaneous representation of a spectrum of dynamically changing value estimates during decision making. *Nature Communications*, 8(1):1942.

Meyniel, F., Sigman, M., and Mainen, Z. (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, 88(1):78–92.

Nassar, M. R., Bruckner, R., Gold, J. I., Li, S.-C., Heekeren, H. R., and Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, 7(1):11609.

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

Nassar, M. R., Wilson, R. C., Heasly, B., and Gold, J. I. (2010). An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. *The Journal of Neuroscience*, 30(37):12366–12378.

Oprea, R. and Vieider, F. M. (2024). Minding the Gap: On the Origins of Probability Weighting and the Description-Experience Gap. *working paper, UC Santa Barbara*.

Palminteri, S., Wyart, V., and Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. *Trends in Cognitive Sciences*, 21(6):425–433.

Pearce, J. M. and Hall, G. (1980). A Model for Pavlovian Learning: Variations in the Effectiveness of Conditioned But Not of Unconditioned Stimuli. *Psychological review*, 87(6):532–552.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106):1042–1045.

Piray, P. and Daw, N. D. (2020). A simple model for learning in volatile environments. *PLOS Computational Biology*, 16(7):e1007963.

Piray, P. and Daw, N. D. (2021). A model for learning based on the joint estimation of stochasticity and volatility. *Nature Communications*, 12(1):6587.

Ricci, M. and Gallistel, R. (2017). Accurate step-hold tracking of smoothly varying periodic and aperiodic probability. *Attention, Perception, & Psychophysics*, 79(5):1480–1494.

Ritz, H., Nassar, M. R., Frank, M. J., and Shenhav, A. (2018). A Control Theoretic Model of Adaptive Learning in Dynamic Environments. *Journal of Cognitive Neuroscience*, 30(10):1405–1421.

Schultz, W. (2007). Behavioral dopamine signals. *Trends in Neurosciences*, 30(5):203–210.

Tavoni, G., Doi, T., Pizzica, C., Balasubramanian, V., and Gold, J. I. (2022). Human inference reflects a normative balance of complexity and accuracy. *Nature Human Behaviour*, 6(8):1153–1168.

Tomov, M. S., Truong, V. Q., Hundia, R. A., and Gershman, S. J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nature Communications*, 11(1):2371.

Vaghi, M. M., Luyckx, F., Sule, A., Fineberg, N. A., Robbins, T. W., and De Martino, B. (2017). Compulsivity reveals a novel dissociation between action and confidence. *Neuron*, 96(2):348–354.

Verbeke, P. and Verguts, T. (2024). Humans adaptively select different computational strategies in different learning environments. *Psychological Review*.

Wagner, A. R. and Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. *Inhibition and learning*, pages 301–336.

Wilson, R. C., Nassar, M. R., and Gold, J. I. (2013). A Mixture of Delta-Rules Approximation to Bayesian Inference in Change-Point Problems. *PLoS Computational Biology*, 9(7):e1003150.

Wulff, D. U., Mergenthaler-Canseco, M., and Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, 144(2):140–176.

Supplementary materials

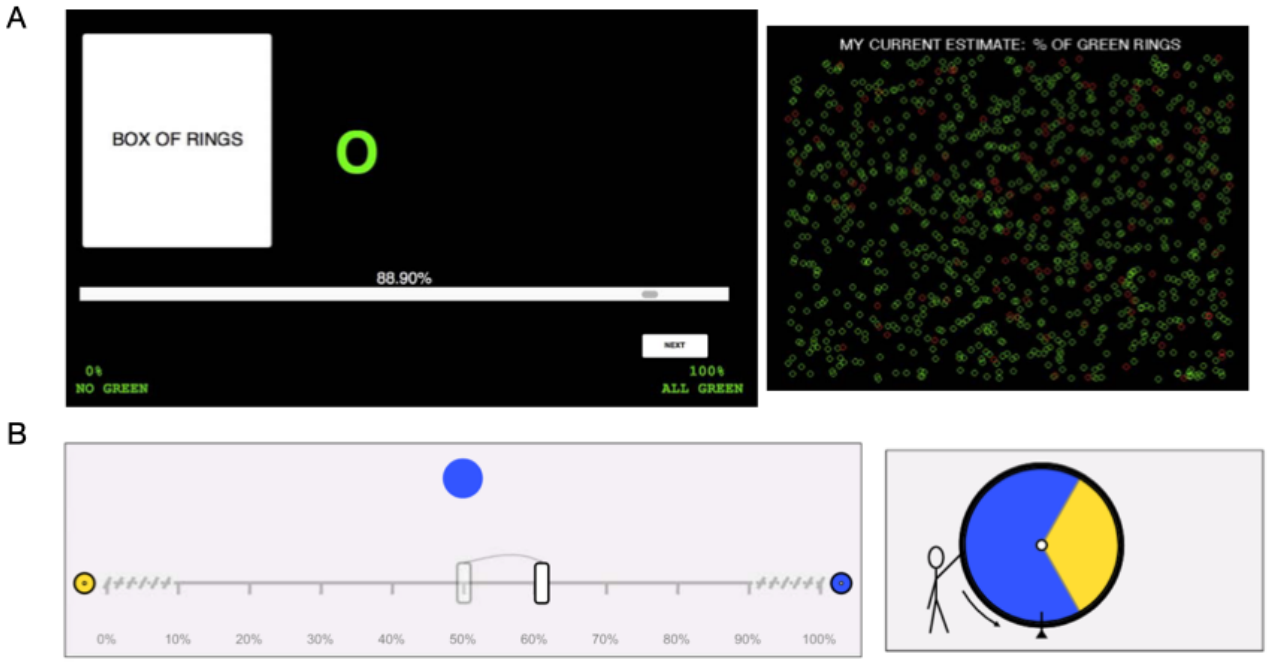

Figure S1: **Illustration of task procedure, adapted from Khaw et al. (2017) and Foucault & Meyniel (2024).** A. Subjects are asked to estimate the proportion of green rings in a box containing 1000 rings (either green or red). They indicate their estimate by adjusting a slider, and a new ring is drawn after they click the 'Next' button. Subjects are informed that the box of rings may occasionally be silently replaced with another box containing a different proportion of green rings. B. Subjects are asked to estimate the probability of seeing a blue vs. yellow dot on the next trial. At any time, they can adjust their estimate by moving a cursor on a slider, using continuous tracking (mouse or touchpad). They are informed that the estimate corresponds to the proportion of blue and yellow on a hidden wheel, which is used to determine the observed colors. Furthermore, the wheel may change at random intervals without warning. The illustration of the wheel is shown during the instruction phase to facilitate task comprehension, but no longer during the task.

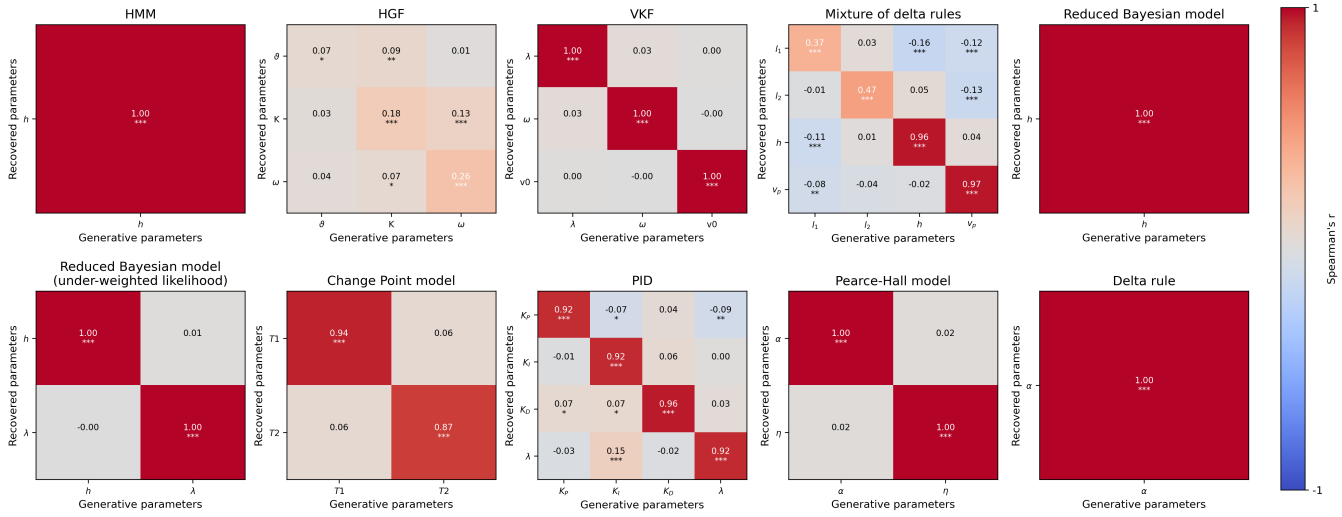

Figure S2: **Parameter recovery analysis.** Spearman's correlations between generative and recovered parameters across 1000 simulations.

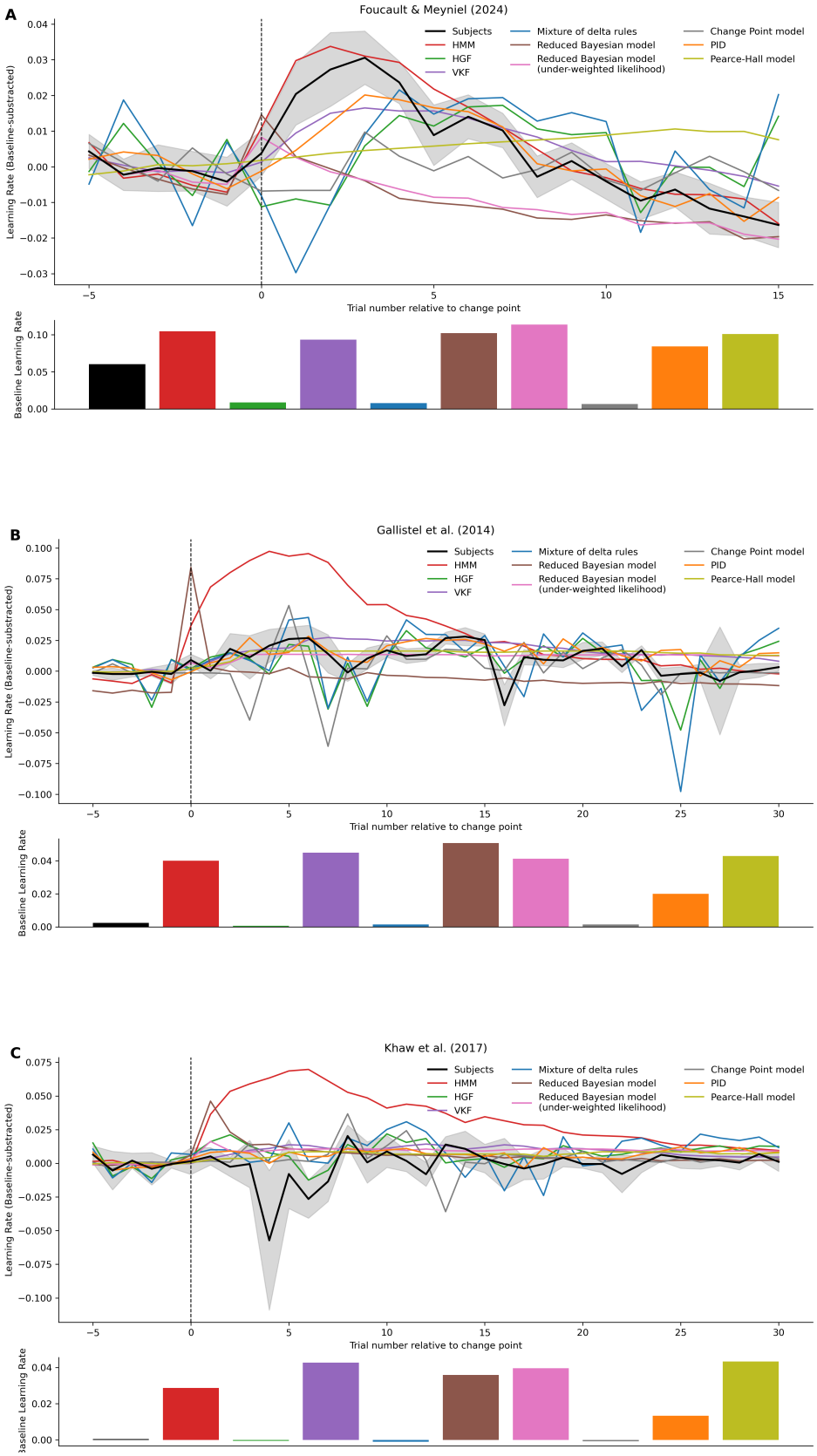

Figure S3: **Model predictive check: Trial-by-trial learning rate in response to a change point.** The baseline of the learning rates (i.e. their value before the change point) is subtracted from the line plots to isolate the transient increase in learning rate that follows a change point in most models. The value of the baseline learning rate is displayed in the bar plots.

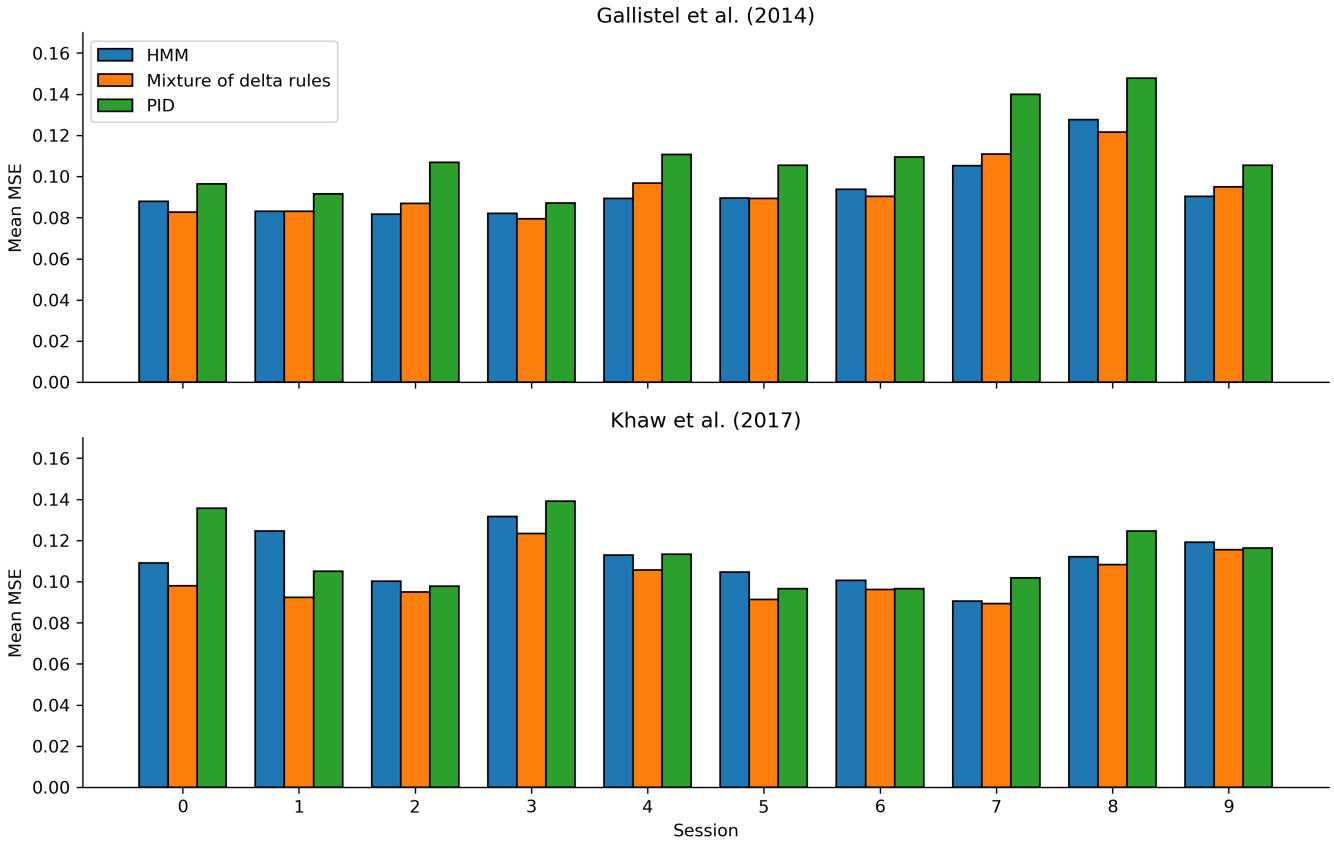

Figure S4: **Mean cv-MSE across subjects of the HMM (optimal Bayesian model), Mixture of delta rules (sub-optimal Bayesian model) and PID (error-updated model) in each session.** This analysis shows that the relative performance among models is generally preserved across sessions. In particular, the Mixture of Delta Rules is the best model already in the first session. This result rules out the possibility that this model would better account for the datasets of Gallistel et al. and Khaw et al than the dataset of Foucault & Meyniel, due to a much lower number of trials in the latter. The number of trials is approximately matched, per subject, in Foucault & Meyniel, and in the first session of Gallistel et al., or Khaw et al.