Further perceptions of probability: in defence of trial-by-trial estimation models

Mattias Forsgren[1,3], Peter Juslin[1] and Ronald van den Berg[1,2,3]

[1]Department of Psychology, Uppsala University, Uppsala, Sweden
[2]Department of Psychology, Stockholm University, Stockholm, Sweden
[3]Equal contributions

Running head:

Further perceptions of probability

Author note

Mattias Forsgren 0000-0003-0394-1626
Peter Juslin 0000-0001-9594-2153
Ronald van den Berg 0000-0001-7353-5960

We have no conflicts of interest to disclose. The research was funded by the Swedish Research Council (Grant 2018-01947) and the Marcus and Amalia Wallenberg Foundation (MAW 2016.0132).

We thank Randy Gallistel and Matthew Ricci for sharing their data with us and for several helpful discussions about their experiments. We also thank Luigi Acerbi for a helpful discussion about the likelihood function.

Correspondence concerning this article should be addressed to Ronald van den Berg Department of Psychology, Albanovägen 12, 11419, Stockholm, Sweden. Email: ronald.van-den-berg@psychology.su.se.

All data analysed in this study can be found at https://osf.io/zhv2r/. This study was not preregistered.

ABSTRACT

Many events we experience are binary and probabilistic, such as the weather (rain or no rain) and the outcome of medical tests (negative or positive). Extensive research in the behavioural sciences has addressed people's ability to learn stationary probabilities (i.e., probabilities that stay constant over time) of such events, but only recently have there been attempts to model the cognitive processes whereby people learn – and track – *non-stationary* probabilities. The old debate on whether learning occurs trial-by-trial or by occasional shifts between discrete hypotheses has been revived in this context. Trial-by-trial estimation models – such as the delta-rule model – have been successful in describing human learning in various contexts. It has been argued, however, that behaviour on non-stationary probability learning tasks is incompatible with trial-by-trial learning and can only be explained by models in which learning proceeds through hypothesis testing. Here, we show that this conclusion was premature. By combining two well-supported concepts from cognitive modelling – delta-rule learning and drift diffusion evidence accumulation – we reproduce all behavioural phenomena that were previously used to reject trial-by-trial learning models. Moreover, a quantitative model comparison shows that this model accounts for the data better than a model based on hypothesis testing. In the spirit of cumulative science, our results demonstrate that a combination of two well-established theories of trial-by-trial learning and evidence accumulation is sufficient to explain human learning of non-stationary probabilities.

KEYWORDS

Probability learning; change-point model; delta-rule; belief updating; hypothesis testing; drift diffusion model

INTRODUCTION

The issue of how people learn and assess probabilities has been pivotal to the behavioural sciences at least since the Enlightenment and studied extensively, especially in psychology and behavioural economics. Typically, this has occurred in the context of assuming *stationary probabilities* in the environment (i.e., probabilities that stay constant over time). This research shows that people are good at learning probabilities from experience with relative frequencies (Edwards, 1961; Estes, 1976; Fiedler, 2000; Peterson & Beach, 1967). Yet, research on heuristics and biases shows that probability assessments are sometimes swayed by subjective ("intentional") aspects, like prototype-similarity (representativeness) or ease of retrieval, leading to biased judgements (Kahneman & Frederick, 2005). People also appear to over-weight extreme probabilities in their decisions when encountering them in numeric form (Tversky & Kahneman, 1992), but under-weight them when they are learned inductively from trial-by-trial experience (Hertwig & Erev, 2009). People frequently have problems with reasoning according to probability theory, leading to phenomena like base-rate neglect and conjunction fallacies (Kahneman & Frederick, 2005; Tversky & Kahneman, 1983), at least if they cannot benefit from natural frequency formats (Gigerenzer & Hoffrage, 1995) that highlight the set-relations between the events (Barbey & Sloman, 2007).

Not all probabilities are stationary, as when, for example, the risks of default in a mortgage market fluctuate over time or the risk of hurricanes changes with a changing global climate. Since modelling how humans learn – and track – *non-stationary probabilities* involves changes in people's beliefs about probability, it has (once again) highlighted the classical issue of whether people learn by trial-by-trial estimation or occasional shifts between discrete hypotheses (Bruner et al., 1956). A neuropsychological and psychophysical literature has suggested a cohort of models that estimate in a trial-by-trial manner (Nassar et al., 2012, 2010; Norton et al., 2019; Wilson et al., 2013, 2018) which is supported by the notion that learning rates are modulated by trial-level prediction errors registered in the anterior cingulate cortex (Behrens et al., 2007; Rushworth & Behrens, 2008; Silvetti et al., 2013). These ideas have now been challenged by a small, mostly recent literature (Gallistel et al., 2014; Khaw et al., 2017; Ricci & Gallistel, 2017; Robinson, 1964). Observations of stepwise, "staircase shaped" response patterns, explicit reports of perceived changes in the underlying probability and other phenomena have been claimed (Gallistel et al., 2014; Ricci & Gallistel, 2017) to be incompatible with trial-by-trial estimation and to require a model built on hypothesis testing. In this Theoretical Note, we scrutinise this notion through simulation and model comparison and find that it was premature: a trial-by-trial model based on two established mechanisms accounts accurately for human data on probability estimation tasks and even outperforms the earlier proposed hypothesis-testing model.

Tracking Probabilities in Non-Stationary Environments

While there is a large literature on how people learn stationary probabilities, there are only a few studies that have addressed learning of non-stationary probabilities. In the studies claimed to support hypothesis testing, participants were asked to estimate the hidden Bernoulli parameter by adjusting a physical lever (Robinson, 1964) or a slider on a computer screen (Gallistel et al., 2014; Khaw et al., 2017; Ricci & Gallistel, 2017). In the latter three of those studies, this was framed as the proportion of green rings in a hypothetical box visualised on a computer screen (Figure 1A). On each trial, the participant could adjust a slider in a range between 0 and 100 percent as their current estimate, before locking in their guess and initiating the next draw from the box (i.e., the next trial). Participants performed 10,000 trials and, importantly, were free to choose to revise their estimate or to leave it unchanged on any trial. Data of interest are the realised outcomes from the Bernoulli process, the underlying true probabilities of the outcomes, and the participant's estimates of these probabilities (Figure 1B). Most participants exhibited stepwise updating behaviour: for long periods they did not adjust their estimates, at other times more often, but never on every trial. One of the studies (Gallistel et al., 2014) included a button labelled "I think the box has changed" that allowed participants to indicate that they believed that there had been a change in the parameter of the Bernoulli process. Half of those participants were also provided with the option to retract their decisions by pressing a button labelled "I take that back" (so called "second thoughts", see Figure 1A).

Two classes of cognitive models: trial-by-trial estimation *vs.* hypothesis testing

As in many areas of the psychology of learning, there are two different ways of explaining how people infer probabilities from experience: models with their origin in the associationist traditions of behaviourism, reinforcement learning, and connectionist models emphasise the continuous updating of beliefs "trial-by-trial", while models with their origin in cognitive psychology emphasise the testing of and discrete shifting between hypotheses.

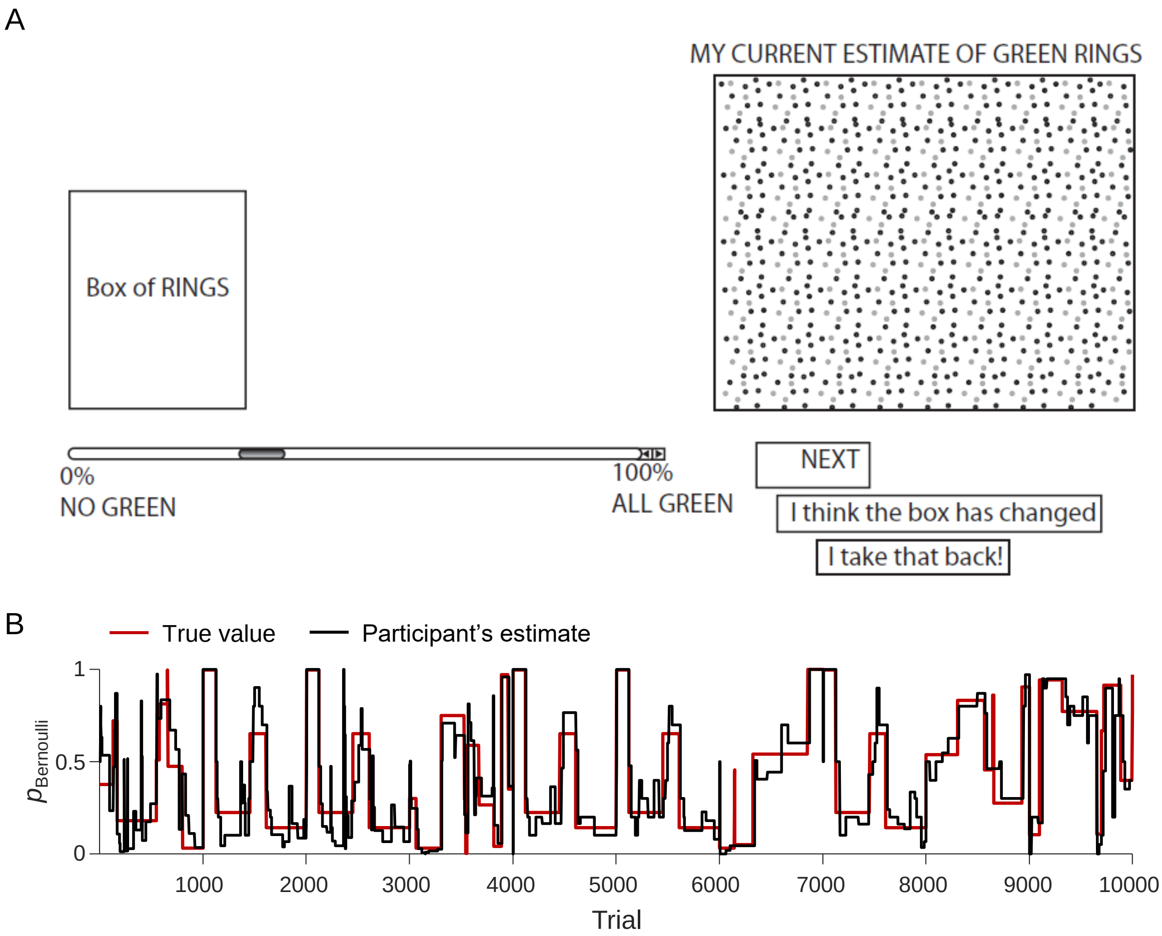

Figure 1 | Experimental paradigm. (A) Screenshot of the task in Gallistel et al (2014). Khaw, Stevens and Woodford (2017) and Ricci and Gallistel (2017) used a similar design but without buttons for reporting that the box has changed or second thoughts. From "The perception of probability," by C. R. Gallistel, M. Krishan, Y. Liu, R. Miller and P. E. Latham, 2014, *Psychological review, Vol. 121,* p. 96-123. Copyright 2014 by American Psychological Association. Reprinted with permission. (B) Example of response data (black) in an experiment where the hidden Bernoulli probability (red) was non-stationary and stepwise (Participant 1 in Gallistel et al, 2014).

A defining feature of trial-by-trial learning mechanisms is that the internal beliefs are updated each time a new data point is observed. One famous example is the delta learning rule which was introduced by Widrow and Hoff (1960) as an algorithm for updating the weights of nodes in a connectionist network (see Widrow & Lehr, 1993, for a review). In psychology, the most famous model based on this rule is the Rescorla-Wagner model of classical conditioning (Rescorla & Wagner, 1972), but it has also been adopted in many other domains (Behrens et al., 2007; Busemeyer & Myung, 1988; Neal & Dayan, 1997; Verguts & Van Opstal, 2014).

In the context of probability estimation, delta-rule learning can be implemented as $\hat{p}_t = (1-\gamma)\,\hat{p}_{t-1} + \gamma\delta_{t-1}$, where $\hat{p}_t$ is the probability estimate at time t, $\hat{p}_{t-1}$ the previous estimate, $\delta_{t-1} = 1 - X_t$ the prediction error at time point $t-1$, and γ the learning rate. The delta-rule accordingly abstracts an online running estimate of the underlying probability and it has the advantage of being recursive: it operates without requiring access to memories going back further than the latest observation.

By contrast, hypothesis-testing models assume that people learn about the world by testing between explicit hypotheses about the state of the world based on confirming or disconfirming feedback (Brehmer, 1974; Bruner et al., 1956). A defining feature of these models is that beliefs are updated in a discrete rather than gradual fashion, because observers hold on to a belief until sufficiently strong evidence has accumulated against it. Hypothesis testing models have been applied to, for example, research on reasoning (e.g. Klayman & Ha, 1987; Oaksford & Chater, 1994; Wason & Johnson-Laird, 1970), categorisation (Ashby & Valentin, 2017; Bruner et al., 1956), and function learning (Brehmer, 1974, 1980). Because a single data point typically provides little evidence about a hypothesis, these models predict that beliefs may sometimes stay unchanged over many outcome observations. Gallistel et al. (2014) formalised a hypothesis-testing model for the learning of non-stationary probabilities, which they called the "If it ain't broke, don't fix it" (IIAB) model. According to this model, participants assess after each new observation whether their current belief is "broke" and only update it if the answer is in the affirmative. The suggestion is that humans do not learn probabilities directly: they learn "change points" in the hidden Bernoulli parameter and use this information and memories of previous outcomes to infer probabilities.

Empirical Phenomena Related to Human Estimation of Non-Stationary Probabilities

To evaluate the plausibility of these two classes of models under non-stationary probabilities, Gallistel et al. (2014) identified a number of important empirical phenomena that any serious model should be able to reproduce. Table 1 provides an overview of these phenomena, which we divide into two categories: those related to slider updates and those related to participants' conceptions of the generative function ("higher-order" beliefs).

We identified an additional phenomenon that has not been reported before but may be informative about the underlying mechanisms: participants regularly make changes to the slider in the *opposite* direction of the colour of the last observation (e.g., decrease their estimate of the probability of a red outcome after observing a red outcome). In the three datasets considered in the present study, 23.6±1.6% (M±SE across participants) of the updates were of this nature. This phenomenon is unexpected under both the IIAB model and standard trial-by-trial models.

Two of the phenomena (10 & 11) cannot be explained by either the proposed hypothesis testing model or a regular trial-by-trial model as specified here. We believe they could do so with certain extensions, which we will come back to in the Discussion. Since these phenomena are a shared issue, and thus not diagnostic of the learning mechanisms that we are contrasting here, we will not consider them in our evaluations of the models.

Table 1.

Empirical phenomena observed in the probability tracking task with the mechanisms of the IIAB and the proposed mechanisms of our extended delta-rule that explain them. All phenomena reported by Gallistel et al. (2014) unless indicated otherwise.

Empirical phenomenon	**Mechanism to explain the phenomenon**	
Related to slider updates	IIAB model	Delta-rule model
1. Stepwise updating of reported estimates of the tracked probability	Slider updates follow belief updates, which only happen when there is sufficient evidence against the present hypothesis.	Beliefs are updated on each trial, but they are accompanied by a slider update only when the discrepancy between the current belief and the slider value exceeds the response threshold.
2. Unimodal step height distributions	Small adjustments happen when the current hypothesis is refined in the troubleshooting stage or when a new hypothesis is close to the present one. Large updates happen when a new hypothesis is very different from the present one.	The response threshold varies across trials. Small slider updates happen when the response threshold is low; large ones may happen when the response threshold is high.
3. Rapid adjustment to changes(*)	A sufficiently low threshold on hypothesis updating.	A sufficiently low response threshold on initiating a slider update.
4. Median response close to true p	Maximum-likelihood updating of the internal estimate of the tracked probability.	Gradient descent updating of the internal estimate of the tracked probability.
5. Inconsistent updating (present paper)	Unexplained by the original model but can be accounted for by adding a variable response threshold.	Accounted for by the variable response threshold.
Related to conception of generative function	IIAB model	Delta-rule model combined with a drift diffusion mechanism
6. Rapid detection of change points	A sufficiently low decision threshold during hypothesis testing (see Phenomenon 3).	A sufficiently low bound in the drift-to-bound change detection mechanism.
7. High hit rates on change-point reports	A consequence of "rapid adjustment to changes".	A consequence of "rapid adjustment to changes".

8. High false discovery rates on change-point reports (**)	A consequence of "rapid adjustment to changes".	A consequence of "rapid adjustment to changes".
9. Occasional changes of mind about the last reported change point ("I take that back")	Incongruent observations are better described by expunging the last change point.	Large prediction errors in the direction opposite to the most recently reported change point are interpreted as evidence that there was no change point after all.
10. The average rate at which changes are reported decreases over time	Unexplained by the original model. Can be explained by modifying priors but at the expense of explaining other phenomena.	Unexplained by our version but explained by previous literature. Participants update their decision bound separation through learning.
11. Declarative perception of periodicity (Ricci and Gallistel, 2017)	Unexplained by the original IIAB model, but can be accounted for by adding a separate function learning process.	Unexplained by the original delta-rule model, but can be accounted for by adding a separate function learning process.

(*) Gallistel et al. (2014) refer to observations of adjustments of the response shortly after a change point as "rapid *detection* of changes" (emphasis added). We use the phrase "rapid *adjustment"* to avoid conflation of this phenomenon and high hit rates, which refers to the observation of participants clicking "I think the box has changed" after a change point.

(**) Gallistel et al. (2014) reported "high hit rates and low false-alarm rates". However, while they calculated the hit rate as reports of a change point in the interval between two change points divided by the number of change points (event level definition), the false alarm rate was calculated as the number of change point reports on trials without a change point divided by the total number of trials without a change point (trial level definition). When using an event-level definition for both metrics, only the hit rate is high and the false-alarm rate is (trivially) close to zero. In this task, we believe it is more informative to look at the false-discovery rate: the number of false alarms divided by the total number of change reports.

The Main Arguments Against Trial-by-Trial Estimation Models

Gallistel et al. (2014) argue that trial-by-trial models are unable to account for several of the phenomena listed above. The first one is the stepwise manner in which participants tend to adjust their estimates of tracked probabilities: they often leave the slider unchanged for long periods of time (Figure 1B), which seems in direct contradiction with any model that updates on a trial-by-trial basis. An additional and closely related argument against those models is based on the distribution of adjustment sizes. Besides making many small adjustments, participants also regularly make large adjustments. Large adjustments are hard to reconcile with the idea of gradual, trial-by-trial updating, because a single observation rarely causes a large change in the estimated probability. Gallistel et al. (2014) argue that large adjustments and periods of constancy instead reflect discrete belief changes.

One potential way to make a trial-by-trial model account both for large slider adjustments and periods with no adjustment is to assume that participants have a "response threshold" that prevents them from making slider updates when the difference between the current slider value and their internal belief is not sufficiently large to justify the effort. Such as threshold could reflect simple "laziness" (recall that they typically performed thousands of trials) or have a more sophisticated basis. While this kind of model produces stepwise response behaviour, it has another problem: it is unable to explain adjustments smaller than the response threshold. As noted by Gallistel et al. (2014), a potential remedy is to make one further assumption, namely that the response threshold can vary across trials. This variability could reflect, for example, fluctuations in attention and motivation, or noise in neural and cognitive processing (Drugowitsch et al., 2016; Faisal et al., 2008). A more sophisticated proposal is that the variable threshold reflects a rational process in which participants trade off costs related to moving the mouse and cognitive processing against accuracy in task performance (Khaw et al., 2017). Gallistel et al. (2014) inspected the behaviour of a trial-by-trial model with a variable response threshold through simulations but were unable to find parameter settings that produced step height distributions resembling the empirically observed distributions. Importantly, however, they seem to have done this by manually trying out a number of different parameter settings, rather than by exploring the space exhaustively. In the present paper, we perform a more systematic search and show that a delta-rule model with a variable response threshold does, in fact, accurately reproduce the empirical distributions.

Another major argument that Gallistel et al. (2014) make against trial-by-trial updating models is based on their observation that participants are able to detect changes in the Bernoulli parameter (Phenomena 6-9 in Table 1). They demonstrated this using a version of the task where participants were asked to press an "I think the box has changed" button whenever they thought that there had been a change in the generative process (see Figure 1A). Some of these participants were also given the option to report "seconds thoughts" about those reports by pushing a button labelled "I take that back." (Figure 1A). Gallistel et al. (2014) interpret the ability to detect and reconsider changes as evidence that participants store a record of the previous change points in memory, over and above a summary representation of the outcomes thus far observed. They argue that such a record is incompatible with trial-by-trial estimation models, which have a much more condensed knowledge state. Here, we show that a delta-rule model extended with a standard drift-to-bound mechanism tracks changes in the underlying

Bernoulli parameter and can account for human reports of and second thoughts about such changes.

Based mainly on the above arguments, Gallistel et al. (2014) ruled out the entire class of trial-by-trial estimation models as a possible explanation for human behaviour on probability estimation tasks. They argued that one instead needs a model with the conceptual richness of a hypothesis testing, "troubleshooting" process that identifies the most likely state of the world to have produced the data. They proposed such a model under the name "If It Ain't Broke don't fix it" (IIAB, described in more detail later) and used simulations to show that there are parameter settings that produce qualitatively similar data patterns as the phenomena observed in human data. Importantly, however, they did not fit the model to any data and they did not perform any quantitative model comparison against alternative models.

Outline of this paper

Gallistel et al. (2014) argued that trial-by-trial estimation models are qualitatively incompatible with human estimation of non-stationary probabilities. This paper presents a re-evaluation of that claim and an extension of their analyses. In the next section, we present the two main contending models: the IIAB hypothesis-testing model proposed by Gallistel et al. (2014) and a trial-by-trial estimation model based on delta-rule learning. Thereafter, we use simulations to examine whether the delta-rule model can reproduce the most important qualitative aspects of human data. Unlike Gallistel et al. (2014), we find that it accurately reproduces those patterns. Having established that there are no qualitative reasons to rule out the delta-rule model, we next examine how well both models account for actual data by fitting them to data from the three previous studies. We find that both models account well for most of the data, even though formal model comparison clearly favours the trial-by-trial model over the IIAB model for almost every participant. To paraphrase Mark Twain (White, 1897), our results indicate that the report of the death of trial-by-trial estimation models was an exaggeration.

MODELS

The IIAB model

We provide a brief description of the "If It Ain't Broke, don't fix it" (IIAB) model here and refer the reader to Gallistel et al. (2014) for a more complete exposition and mathematical details. A key characteristic of this model is that it has a relatively stable internal belief about the tracked probability: it only updates this belief when there is sufficient evidence against the current value. It proceeds in two stages. In the first stage, it tests whether the currently held belief about the tracked probability is “broke”. This test is performed by computing the discrepancy between the belief and the outcomes observed since the last registered change point. If the discrepancy – measured as Kullback-Leibler divergence – exceeds a decision threshold T_1, it is concluded that something is “broke”. Each time this happens, the model enters a “troubleshooting” stage, in which it considers three hypotheses on why the current estimate may be “broke”: (i) there was a change in the generative process (“I think the box has changed”), in which case the model will register a new change point and update its estimate of p_{true} accordingly; (ii) the previously registered change point was a mistake (“I take that back”), in which case the model will expunge the last recorded change point and update its estimate accordingly; (iii) the previous estimate of p_{true} was wrong but the change point record is correct, in which case the model will update its estimate of p_{true} but not register or expunge any change point. Hypothesis (iii) corresponds to concluding that the estimate was “broke” due to sampling error, but it is not assumed that such beliefs are recorded in memory. Gallistel et al. (2014) argue that these “troubleshooting” steps allow the IIAB model to explain behavioural phenomena related to the participant’s knowledge about the generative function (phenomena 6-11 in Table 1). Since the updated estimate is always the average of all observations since the last believed change point, the model must retain the full sequence of observations since the second to last change point.

The original version of the IIAB model has just two parameters: threshold T_1 mentioned above and an additional threshold T_2 that is used in the troubleshooting stage. While both thresholds are fixed, the evidence in the first stage is scaled by the number of trials since the last change (the sample size). The IIAB model will therefore become increasingly sensitive to small discrepancies between the current belief and the most recently observed evidence when no change point has been detected for a while.

Predictions related to slider updates can be derived directly from the model’s internal belief state about the tracked probability. Predictions related to a participant’s reports of suspected changes in the generative process and changes of mind about those reports can be derived directly from the model’s “troubleshooting” stage. Hence, this is a rich model that makes predictions about all of the empirical phenomena listed in Table 1 except 11.

The delta-rule model

In contrast to the IIAB model delta-rule models update the estimate after *every* new observation. The most basic version of the delta-rule model does this using a recursive function of the form

$$\hat{p}_{\mathrm{B},t} = (1-\lambda)\,\hat{p}_{\mathrm{B},t-1} + \lambda E_t, \tag{1}$$

where $\hat{p}_{\mathrm{B},t}$ is the estimated probability of the tracked event (in this case: of observing a green ring) on trial *t*, $E_t = \hat{p}_{\mathrm{B},t-1} - X_t$ is the *prediction error*, and $\lambda \in [0,1]$ is the *learning rate*. This model thus proceeds by constantly adjusting its estimate of the tracked probability in the direction of the latest observed outcome: seeing a green ring slightly increases the observer's estimate of the proportion of green rings in the box and seeing a red one decreases it. The higher the value of the learning rate, λ, the larger the trial-by-trial adjustments. In environments with frequent, abrupt changes in the generative process, it is beneficial to have a high learning rate because that will allow the model to catch up quickly to those changes. By contrast, in stable or very slowly changing environments it is better to have a slow learning rate, to avoid the estimates being overly sensitive to occasional unexpected outcomes. The environments used in previous studies on non-stationary probability tracking (Gallistel et al., 2014; Khaw et al., 2017; Nassar et al., 2010; Norton et al., 2019) are often a mixture of those two situations: long periods of stability with occasional, abrupt changes (Figure 1B). In such environments, it can be a disadvantage to have a single, fixed learning rate. Several modifications to the standard delta-rule model have been proposed that might work better in mixed environments, for example the addition of a second kernel (Gallistel et al., 2014) and the use of a dynamic learning rate (Nassar et al., 2010). However, it has been shown in a similar task that the basic model typically performs as well as or even better than more complex alternatives (Norton et al., 2019). While we will consider two variants later (see Results), our main focus will be on the most basic, single-parameter version of the delta-rule, as specified by Equation (1).

The cumulative prediction error as a predictor of changes in the generative process

Gallistel et al. (2014) rightly point out that the delta-rule by itself cannot account for participant data related to explicit change point reports (phenomena 6-9 in Table 1). This is not surprising since the delta rule is a learning mechanism. To explain change point reports, it needs to be combined with a decision-making mechanism. One of the most established decision-making mechanisms to date is the drift-diffusion mechanism (Bogacz et al., 2006; Ditterich, 2006; Ratcliff, 1978), which finds broad support in behavioural, neurophysiological, and computational studies (Ratcliff, 1978; Ratcliff et al., 2016; Wagenmakers, 2009). Here, we will explore if it can also explain change point reports in probability estimation tasks.

A central quantity in delta-rule models is the trial-by-trial prediction error, that is, the difference between the predicted and observed outcome. When the generative process is stable and the observer's estimate has homed in on a value close to the true value of the tracked variable, prediction errors tend to cancel each other out over trials (Figure 2, first 100 trials). After an abrupt change in the generative process (Figure 2, trial 100), however, there will typically be a burst of relatively large prediction errors with a sign that indicates the direction of the change. Hence, the cumulative prediction error is indicative of changes in the generative process: a value close to zero suggests a stable process; a large negative value suggests that there was a recent increase in the Bernoulli parameter; a large positive value suggests that there was a recent decrease in the Bernoulli parameter. Because of its diagnostic value, observers could use the cumulative prediction error to detect changes in the generative process when tasked to do so. This can be modelled by adding a standard drift-to-bound accumulator to the model and let it trigger an "I think the box has changed" response whenever the cumulative prediction error exceeds a decision bound (Figure 2). Fully in line with the philosophy of delta-rule models, this cumulative error can be updated recursively and imposes negligible memory requirements.

Importantly, drift-diffusion mechanisms can also explain "second thoughts", which are known as "changes of mind" in the decision-making literature. This is done by introducing a temporary second bound at the moment that an initial decision has been made (e.g., Resulaj, Kiani, Wolpert, & Shadlen, 2009; Van den Berg et al., 2016). This bound will be crossed if the immediate post-decision information is sufficiently inconsistent with the original decision, triggering a change-of-mind response. A typical way to implement this bound is to use two parameters, specifying its height and lifetime. Because we have very little data on changes of mind (115 reports by 5 participants in a total of 50,000 trials), we take a simpler approach by setting the change-of-mind bound equal to the original bound but in the opposite direction of the detected change point, such that the lifetime of the bound is the only additional parameter required to model these rare responses.

Response threshold

Previous studies (Gallistel et al., 2014; Khaw et al., 2017; Robinson, 1964) have considered the possibility that participants do not adjust the slider when the difference to their internal belief is too small. This could arise from participants economising their time costs[1]. Additionally, cognitive processes are noisy (Drugowitsch et al., 2016; Faisal et al., 2008) and participants' levels of motivation and attention might fluctuate over time, why the discrepancy required for an update may vary. We will first model this as in Gallistel et al. (2014): as a threshold value for the required discrepancy drawn from a constrained Gaussian distribution. We will then test models where the threshold is drawn from a beta distribution. We parameterise both thresholds by their mean and variance.

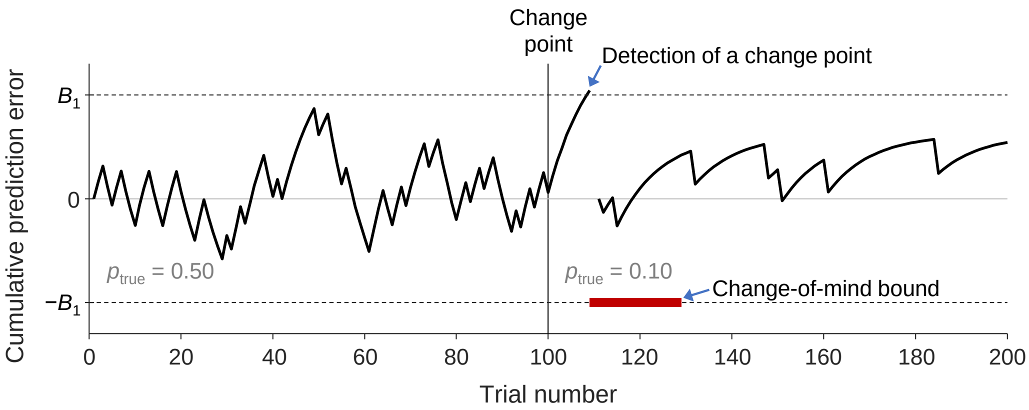

Figure 2 | A proposed mechanism to detect changes in a Bernoulli process based on accumulation of prediction errors. Simulation of the cumulative prediction error in a delta-rule model with a learning rate of 0.10. The true value of the Bernoulli parameter is 0.50 for the first 99 trials and then abruptly changes to 0.10. Before the change, the cumulative prediction error hovers around 0, because positive and negative errors cancel each other out. At around trial 50 there is almost a false alarm. Immediately after the change, the cumulative prediction error quickly increases, because more positive estimation errors are experienced than negative ones. The cumulative prediction error hits decision bound B_1=3.0 at trial 109 which triggers an "I think the box has changed" response, resets the cumulative prediction error to 0, and instates a temporary change-of-mind bound (which is not being crossed in this example). The shape of the cumulative prediction error looks different after the change, because after the model has learned the new value of p_{true}, the trial-by-trial prediction errors are 0.10 (on 90% of the trials) and −0.90 (on 10% of the trials) while they were −0.50 and 0.50 (in 50% of the trials each) before the change.

Response noise and lapse rate

To account for inaccuracies in predicted slider settings – due to factors such as motor noise and model mismatch – we included response noise in all models. This noise was implemented as a beta distribution centred on the model's predicted response, m, and was

[1] For example, in the experiment by Gallistel et al. (2014), trials with an update took on average three times longer (4.22 ± 0.18 seconds) than trials without an update (1.39 ± 0.01 seconds). Responding on each trial would almost have tripled the median session time – from around 25 minutes to around 70 minutes.

applied to trials on which a slider update was predicted. Since the variance of the beta distribution has an upper bound (equal to $m - m^2$), we parameterised it as a relative value between 0 (no variance) and 1 (maximum variance). The (relative) variance was fitted as a free parameter. Moreover, we included a small lapse rate (1/1000) to account for lapses in attention and to avoid numerical instabilities in model variants without any other sources of stochasticity (such as the original IIAB model).

RESULTS

This section consists of two parts. First, following the approach by Gallistel et al. (2014), we perform simulations to re-assess the conclusion that a delta-rule model cannot reproduce the main qualitative phenomena observed in human data (Table 1). Next, we perform a likelihood-based model comparison in which we quantitatively compare this model to the main contender, the IIAB model. Thereafter, we inspect the likelihood-based model fits in greater detail and test two alternative models from the literature.

Reassessment of the conclusion that delta-rule model predictions are qualitatively inconsistent with data

The simulation results by Gallistel et al. (2014) suggested that delta-rule estimation models are unable to produce slider updates that are qualitatively similar to human behaviour. In particular, they were unable to find parameter settings that reproduced the distributions of step widths and step heights observed in human data (phenomena 1-3 in Table 1) and concluded that trial-by-trial models are, therefore, fundamentally unfit to account for human estimation of non-stationary probabilities. Here, we reconsider this finding by using an approach that differs from theirs in an important way: instead of manually trying out parameter settings, we systematically explore parameter space using an optimisation method. Specifically, we let the algorithm search for the setting that minimises the root mean squared deviation (RMSD) between the data and the model prediction for the summary statistic of interest (histograms of step width and height, cumulative number of updates, etc).

In this analysis, we use the exact same delta-rule model as tested by Gallistel et al. (2014), which has three parameters: the learning rate (λ), the mean of the (Gaussian) response threshold distribution (μ_τ), and the coefficient of variation of this distribution (cv_τ); no response noise or lapse rate was included in the model at this stage. Just like Gallistel et al. (2014), we constrain cv_τ to have a maximum value of 0.33. In contrast to their findings, we find that this model reproduces the step width and step height distributions very well (Figure 3). It also does an excellent job in reproducing the other phenomena related to slider updates: the cumulative number of updates, the median response values, and the cumulative distribution of the latency between changes in the Bernoulli parameter and the next slider update.

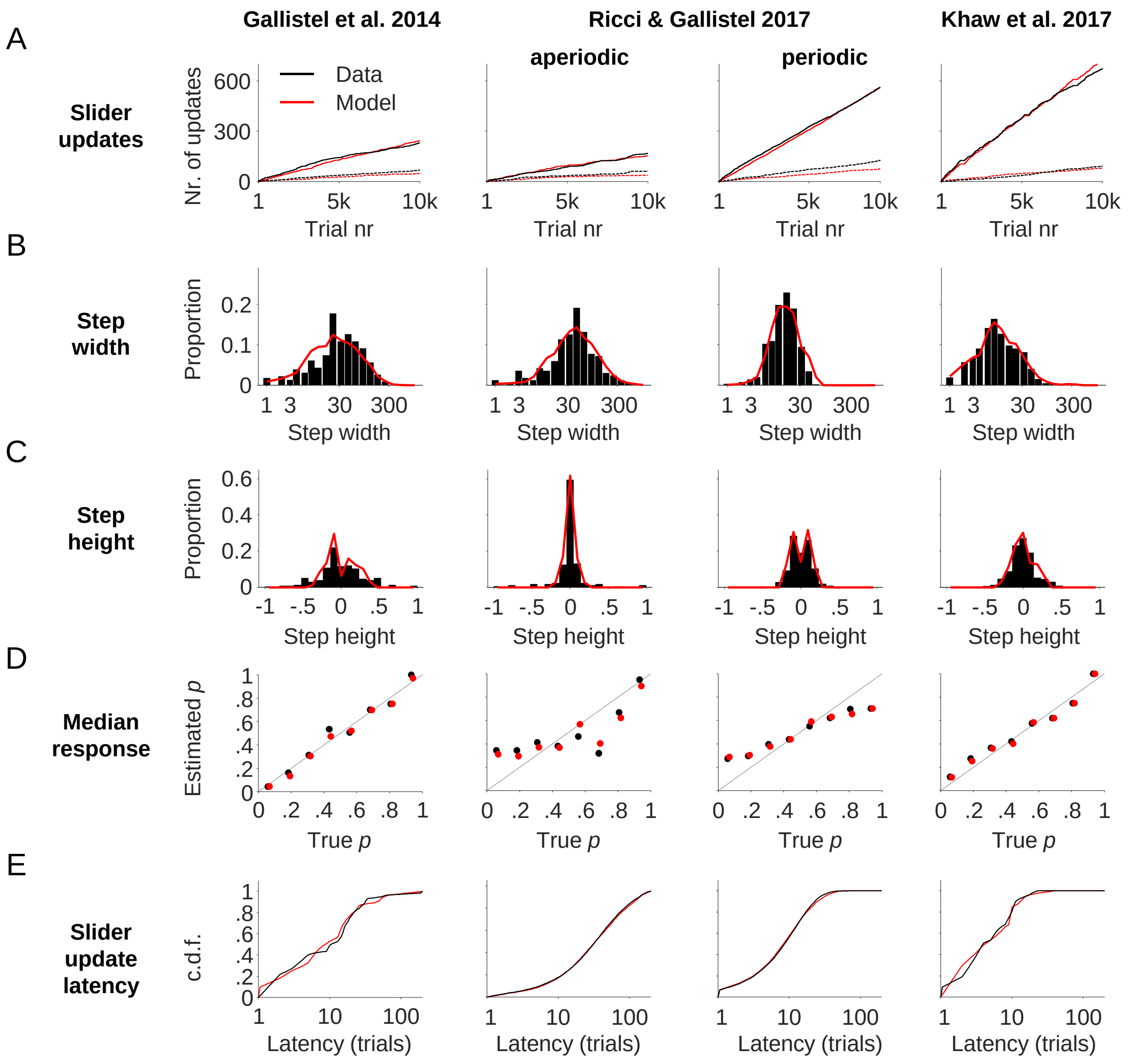

Figure 3 | Evaluation of qualitative predictions by the delta-rule model related to slider settings. Results are shown for Participant 1 in each of the 4 analyzed datasets. The model simulations results (red) were obtained by minimizing the root mean squared deviation (RMSD) with the data (black). (A) Total number of slider updates (solid) and number of inconsistent slider updates (dashed) as a function of trial number. (B) Distribution of the number of trials between consecutive slider updates. (C) Distribution of the magnitude of slider updates on trials with an update. (D) Median estimate of the tracked probability versus the median true value. (E) Cumulative distribution of the number of trials between a change in p_{true} and the next slider update.

Next, we extend the model with a drift diffusion mechanism on the prediction error and test if the resulting model can account for phenomena related to the conception of the generative function (phenomena 6-9 in Table 1). We find that the model accurately reproduces these phenomena too (Figure 4): the cumulative number of "I think the box has changed" responses; the cumulative number of "I take that back" responses; the cumulative distribution of the latency between a change in the generative function and the observer's detection of the change; the hit rates, false discovery rates, and false alarm rates of box-change detections.

In conclusion, the predictions of a delta-rule model combined with a standard evidence accumulation mechanism are qualitatively consistent with human tracking and detection of changes in the parameter underlying a Bernoulli process. This means that the main argument that Gallistel et al. (2014) presented against trial-by-trial models does not hold and may stem from an inexhaustive exploration of parameter space.

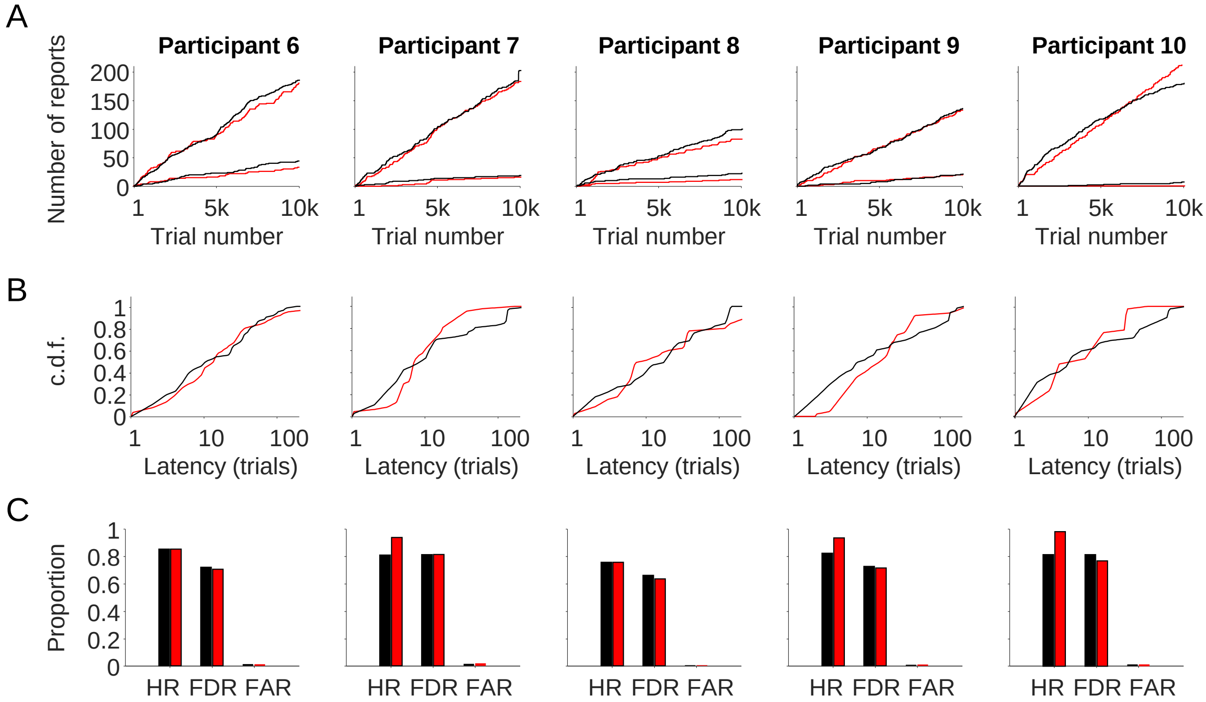

Figure 4 | Evaluation of qualitative predictions by the delta-rule model related to detection of changes in the generative function (delta-rule model). The model simulations results (red) were obtained by minimizing the root mean squared deviation (RMSD) with the data (black). (A) Total number of "I think the box has changed" reports (solid) and "I take that back reports (dashed). (B) Cumulative distribution of the number of trials between a change in p_{true} and the next "I think the box has changed" report. (C) Hit rates, false discovery rates, and false alarm rates on change point detections.

Likelihood-based model comparisons

The results so far show that just like the IIAB model, the delta-rule model is capable of explaining previously established facts about human performance on probability tracking tasks. But which of the two models explains them *better*? Although the above approach of inspecting summary statistics is useful for checking if a model's predictions are qualitatively consistent with well-established facts, it cannot be used for quantitative model comparison. The main problem – as also noted by Gallistel et al. (2014) – is that there is no obvious way to weight misestimates in one summary statistic against misestimates in another, which makes it impossible to formulate a single measure to base judgements on.

To compare the models in a quantitative and more principled manner, we will next evaluate them based on likelihoods computed from raw data (see Supplemental Materials for details). This method has two major advantages over evaluating models based on their predicted summary statistics. First, it is a much more stringent evaluation because it takes *all* aspects of the data into account and describes them using a *single* set of parameters. Second, it allows one to evaluate model performance using a single, formal measure, such as the Akaike Information Criterion (Akaike, 1974) or cross-validated log likelihoods.

We fit the models to the raw data from four experiments (Table 4) reported in the three previous studies[2]. In each experiment, the number of trials per participant varied from 9,000 to 10,000 and were divided over 9 or 10 sessions. In total, the data consists of 286,890 trials performed by 29 participants over 287 sessions. All data can be found at https://osf.io/zhv2r/. We limit these analyses to the slider update data, because "I think the box has changed" and "I take that back" responses were collected for only 10 and 5 of the participants, respectively.

We first compare the two models contrasted in Gallistel et al. (2014): a single-kernel delta-rule model with a variable response threshold and the IIAB model. We fit the models to all sessions jointly, that is, with a single set of parameters per participant. The delta-rule model accounts for the data overwhelmingly better than the IIAB model (Figure 5A): for each of the 29 participants, the delta-rule model is favoured over the IIAB model by a difference of at least 18020 log likelihood points (M±SE: 28654 ± 904)[3]. Hence, not only is the delta-rule model viable from a qualitative perspective, its quantitative account of the raw data is much better than that of the alternative model proposed by Gallistel et al. (2014).

[2] There is one other study using the same paradigm (Robinson, 1964), but it has no preserved record of the data known to us.

[3] When fitting the models separately to each session, the average difference is 285±40 in favour of the delta model. Considering that log likelihoods scale linearly with the number of trials, this difference is comparable to that obtained by fitting the full datasets.

Table 4. *Overview of Datasets Used to Evaluate the Models.*

Exp. ID	Study	Underlying function	Number of participants	Number of trials per participant	Number of trials per session	Total number of sessions
E1	Gallistel et al. (2014)	Stepwise	10	10,000	1,000	100
E2	Ricci & Gallistel (2017)	Continuous (aperiodic)	5	10,000	1,000	50
E3	Ricci & Gallistel (2017)	Continuous (periodic)	3[4]	9,000	1,000	27
E4	Khaw et al. (2017)	Stepwise	11	9,990	999	110

There are two major differences between the models that could explain the enormous difference in goodness of fit. First, they have different belief updating mechanisms: hypothesis testing in the IIAB model and trial-by-trial updating in the delta-rule model. Second, the delta-rule model includes a threshold on the slider updates. Hence, it could be that the IIAB model performs poorly not because of its assumptions about how people update their internal beliefs, but rather due to lacking a response threshold. To examine the evidence for the belief updating mechanisms specifically, one must equalise the models in terms of the assumption about the response threshold. Therefore, we next fit a variant of the IIAB model with the exact same response threshold mechanism as in the delta-rule model. This version has a much better goodness of fit, but it is still outperformed by the delta-rule model for 25 out of 29 participants, with an average log likelihood difference of 271 ± 44 across all participants (Figure 5B). This dramatic change in the log likelihood difference suggests that a response threshold is of primary importance to quantitatively account for the data.

A response threshold can be implemented in many ways and which version is chosen can strongly affect the model fit (see Khaw et al., 2017). So far, we have followed Gallistel et al. (2014) by assuming a variable threshold in the shape of a Gaussian distribution with a constraint on the magnitude of the noise. We will now test an alternative version by making two changes. First, we remove the constraint on the amount of variance ($cv_\tau \leq 0.33$) because its justification is unclear to us and it may have limited both models' ability to account for participants'

[4] This experiment had 4 subjects, but we suspect that for one of them the responses were flipped between two sessions. We excluded this subject from our analyses.

response behaviours. Indeed, for all but one of the participants we find that the fitted coefficient of variation of the response threshold was at the maximum of 0.33. Second, we switch to a beta distribution because, unlike the Gaussian distribution, it produces responses that are properly bounded between 0 and 1. The goodness of fit increases substantially for both the IIAB and delta-rule model, by 650 ± 130 and 495 ± 121 log likelihood points, respectively. The delta-rule model still outperforms the IIAB model for 27 out of 29 participants, with an average difference of 125 ± 20 (Figure 5C).

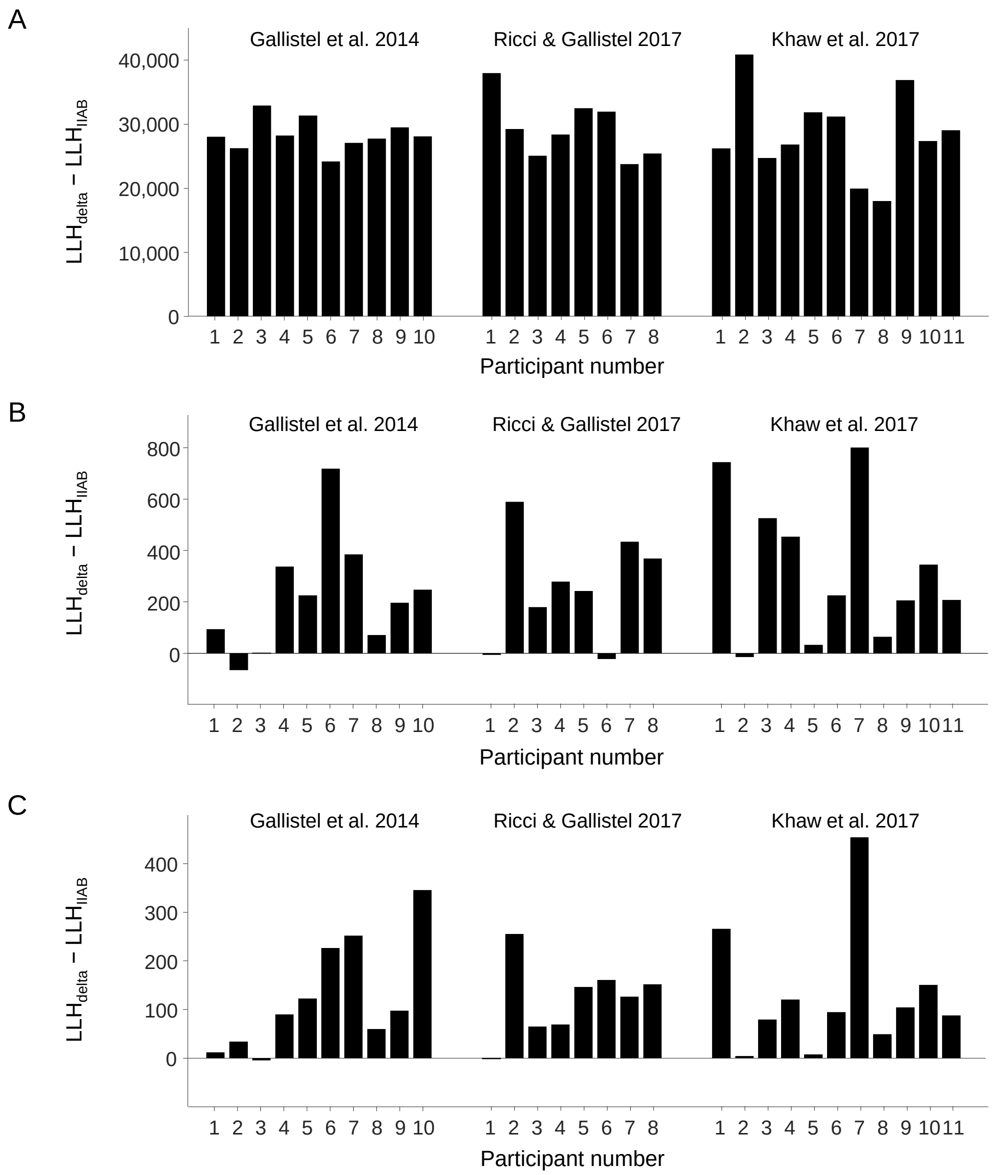

Figure 5 | Model comparison results. Model performance is expressed as the log likelihood of the delta-rule model (LLH_{delta}) relative to that of the IIAB model (LLH_{IIAB}). Positive numbers indicate a better fit for the delta-rule model. (A) A delta-rule model with a constrained Gaussian response threshold versus the original IIAB model (without a response threshold). (B) A delta-rule model with a constrained Gaussian response threshold versus an IIAB model with the same response threshold. (C) A delta-rule model with a beta-distributed response threshold versus an IIAB model with the same response threshold.

Altogether, these results show that from a quantitative model comparison perspective the delta-rule accounts better for the data than the IIAB model. We checked that this conclusion is robust to changes in the assumptions about the lapse rate and the response noise (see Supplemental Materials). Because the beta distribution provides a much better fit, we will employ it in the remaining analyses.

Evaluation of qualitative phenomena under maximum-likelihood parameters

Likelihood-based model comparison is a powerful tool to evaluate models against each other in a quantitative and principled way. However, results of such relative comparisons are of little value if none of the models provides a decent account of the data. To verify that this is not the case, we next examine the models' qualitative predictions under maximum-likelihood parameters. Using these parameter settings, the delta-rule model reproduces the qualitative phenomena related to slider settings almost as well as in the earlier RMSD-based fits (Figure 6). Moreover, it also accounts well for the raw, trial-by-trial slider settings (Figure 7). The maximum-likelihood fits of the original IIAB model (i.e., without response threshold) are very poor (Figures S2 and S3 in Supplemental Materials). After adding a response threshold, the fits become visually of similar quality to those of the delta-rule model (Figures S4 and S5 in Supplemental Materials), which once again highlights that the assumption of a response threshold seems important to account for the data.

Parameter estimates

Response threshold distributions in the delta-rule model. Inspection of the maximum-likelihood estimates of the response thresholds suggests that there is large variation in the trial-to-trial thresholds (Figure 8). As a result, the choice of whether or not to update the slider on any given trial is only partially determined by the discrepancy between the internal belief and the current slider value. Previous literature (Biele et al., 2009; Gonzalez & Dutt, 2011) has suggested a completely discrepancy-independent mechanism called "inertia" where the decision to update is determined by the flip of a weighted coin. We tested this mechanism by replacing the response threshold with a constant probability of updating on each trial, implemented as a free parameter. This mechanism makes the fits substantially worse for 27 of the 29 participants, with an average of 69 ± 14 log likelihood points over all participants. This suggests that the update decision at least in part depends on the discrepancy between the internal belief and the current slider value.

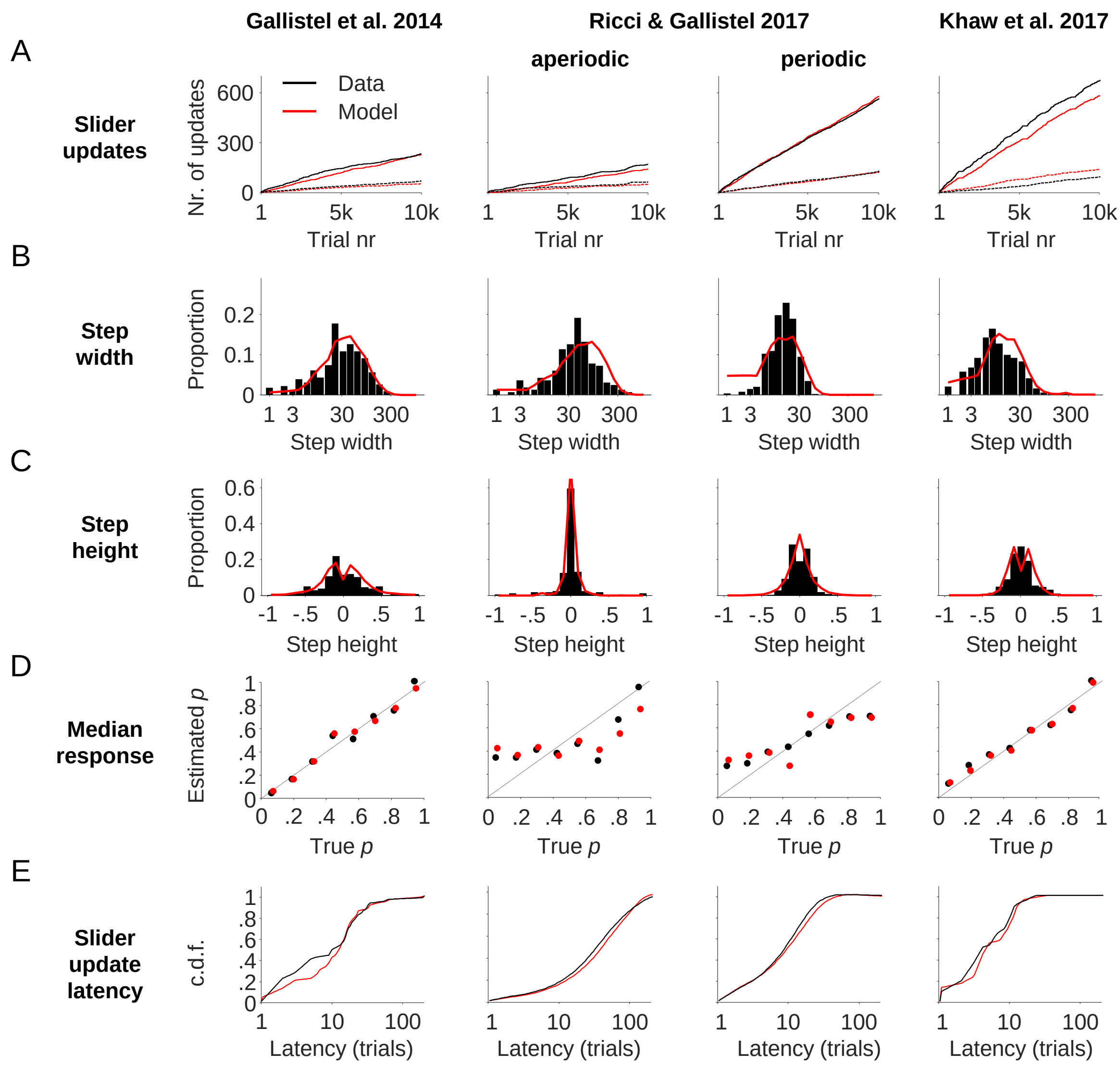

Figure 6 | Delta-rule model behavior under maximum-likelihood parameter estimates. Data (black) are shown for Participant 1 in each of the 4 analyzed datasets. The model predictions (red) were obtained by simulating responses using the maximum-likelihood estimates of the parmater values. (A) Total number of slider updates (solid) and number of inconsistent slider updates (dashed) as a function of trial number. (B) Distribution of the number of trials between consecutive slider updates. (C) Distribution of the magnitude of slider updates on trials with an update. (D) Median estimate of the tracked probability versus the median true value. (E) Cumulative distribution of the number of trials between a change in p_{true} and the next slider update.

Response noise in the delta-rule model. The median estimate of the (relative) variance of the beta response noise distribution is 0.058 (IQR: 0.041). To get an intuition of the magnitude of this noise, we performed a model simulation. Using each participants' maximum-likelihood parameter estimates, we computed the RMSD between predicted slider updates before and after adding response noise, in a fictitious experiment in which the tracked probability was uniformly distributed between 0 and 1. We find that the RMSD equals 0.118 ± 0.007. This seems reasonable, because it is in the same order of magnitude but smaller than the (model-free) RMSD between the tracked probability and the actual participant responses (0.189 ± 0.009). Hence, the model assigns approximately half of the slider error magnitude to response noise.

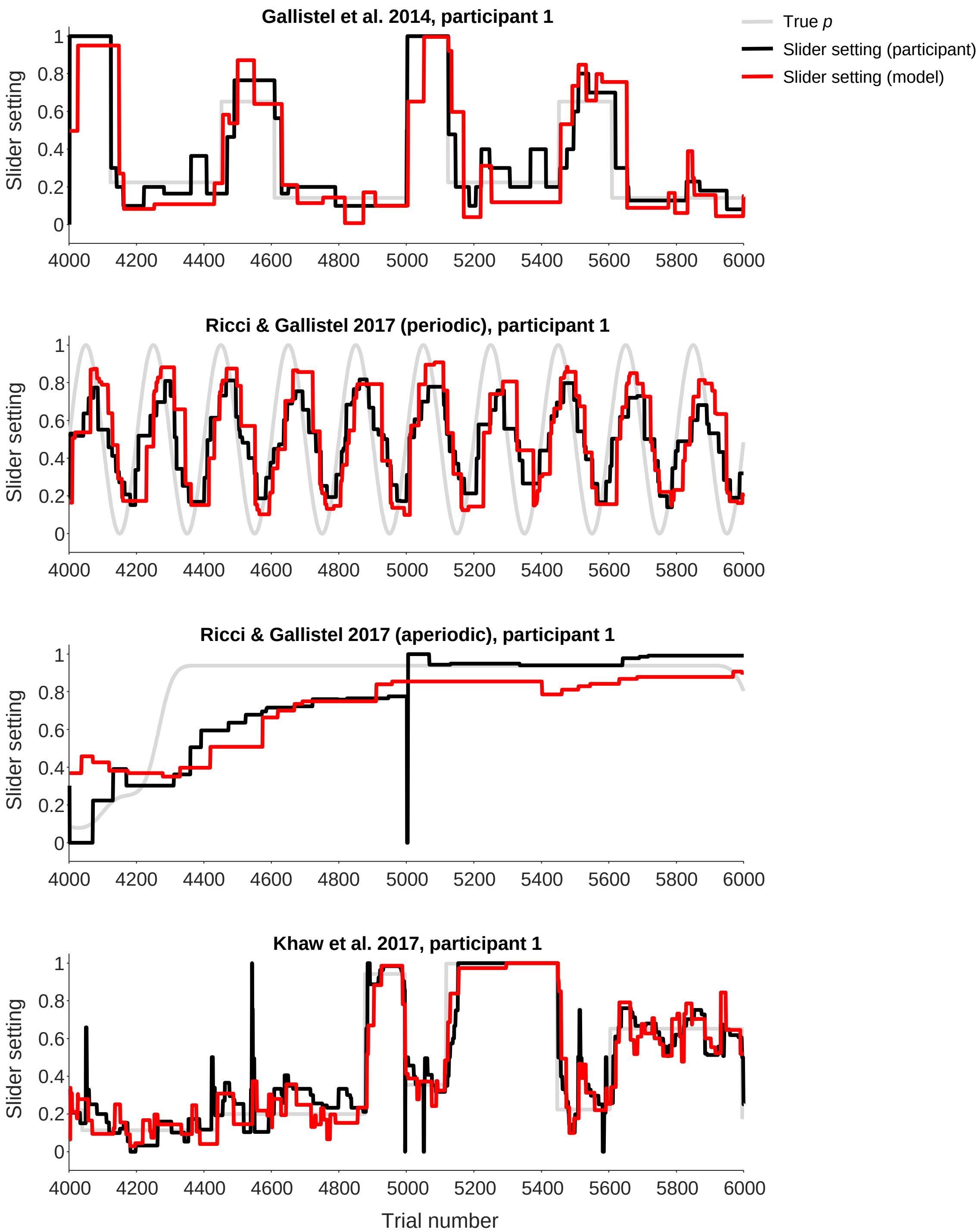

Figure 7 | Examples of trial-by-trial slider settings of delta-rule model under maximum-likelihood parameter estimates. For visualisation purposes, only the central 2,000 trials are shown for each dataset.

Decision threshold in the IIAB model. The decision threshold parameter in the IIAB model – which controls when the model considers the current belief to be "broke" and in need of an update – is estimated to be close to 0 for every participant (M=0.032, SE=0.018). This means that the IIAB model captures the data best when setting its parameters in such a way that it essentially becomes a trial-by-trial estimation model and accounts for stepwise behaviour through the response threshold.

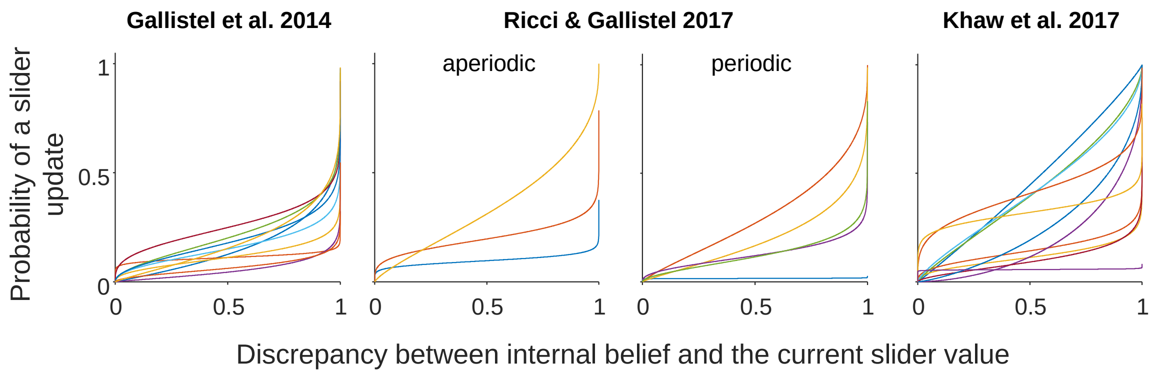

Figure 8 | Maximum-likelihood estimates of the variable response thresholds in the delta-rule model (different colors indicate different participants). The threshold is visualised as the cumulative probability distribution of making a slider update as a function of the size of the discrepancy between the internally held belief about the tracked probability and the current slider value. For most participants, the probability of performing a slider update increases with this discrepancy.

Two-kernel delta-rule model

Under conditions where there are large and infrequent changes, as in much of the experiment data considered in this study, the standard version of the delta-rule faces a problem. If a lot of weight is put on the most recent history (by having a high learning rate), the model will quickly catch on to changes but exhibit excessive volatility during the long periods where the true probability is unchanged. If, on the other hand, recency is given only a little weight, the model will avoid excessive volatility but be slow to catch on to sudden changes. As a potential solution, Gallistel et al. (2014) considered a two-kernel variant that keeps track of two running averages with different learning rates. The model switches between these two running averages, allowing it to keep up with sudden changes while avoiding excessive volatility. Gallistel et al. (2014) rejected this model because it was allegedly unable to produce unimodal step height and step width distributions, which is not the case when we model it with a beta response threshold. We find that this version outperforms the regular delta-rule by 38.8 ± 6.1 log-likelihood points (see Supplemental Materials for implementation details). Having the flexibility to weight evidence differently at different times thus seems important.

An approximately Bayesian delta-rule model

Nassar et al. (2010) suggested a delta-rule variant inspired by the same Bayesian change point detection model (Adams & MacKay, 2007) as the IIAB model. Their “approximately Bayesian delta-rule model” explicitly considers two hypotheses after each new observation: either there has been a change in the true, covert probability or there has not. Unlike the IIAB model, it performs no discrete hypothesis testing but instead balances the relative evidence of these two possibilities trial-by-trial. This balancing can be rewritten (see Nassar et al., 2010, and Supplemental Materials for details) as a delta-rule with an adaptive learning rate.

We find that this model outperforms the IIAB model by 70 ± 25 log likelihood points, but performs worse than the regular delta-rule model by 55 ± 20 log-likelihood points. Nassar et al. (2010) also suggested a non-normative variant that allows underweighting of likelihoods by raising them to a power. When the power is equal to 0, this model reduces to the regular delta-rule model for all but the first few trials (and can thus not perform much worse than that model). This non-normative variant performs better than the regular delta-rule model, by 31 ± 11 log-likelihood points, but often underweights likelihoods heavily (Figure S6 in Supplemental Materials). In sum, this version of an adaptive learning rate does seem to improve on the regular delta-rule model if it is allowed to deviate from normativity.

Slider updating consistency

Why do people regularly make a slider update that is inconsistent with their last observation, such as decreasing their estimate of the probability of red outcomes after observing a red outcome? In a basic delta-rule model, response updates are always consistent with the most recent observation: observing a red ring increases the estimate of the probability of observing a red ring and observing a ring of the other colour decreases it. In the IIAB model, the “second thoughts” mechanism might on rare occasions cause inconsistent updating (4.50 ± 0.65% of all slider updates under the maximum likelihood parameter values).

One potentially important source of inconsistent updating is the response threshold. For example, a momentarily high threshold might suppress a downwards adjustment of the slider but it will never suppress a downwards adjustment of the internal belief. If the threshold on the next trial happens to be lower, and the new observation increases the internal belief by less than it was decreased on the previous trial, the reported estimate will be adjusted downwards – which would be inconsistent with the last observation. Indeed, the maximum-likelihood fits of the IIAB model and delta-rule model with a response threshold predict that 31.8±0.8% and 22.1±0.9% of the slider updates, respectively, are inconsistent. Hence, the IIAB model slightly overestimates the empirical proportion of 23.6±1.6% (BF_{10} = 305; two-tailed paired-samples *t*-test), while the predictions of the delta-rule model are consistent with the data (BF_{10} = 0.36).

DISCUSSION

Previous studies where participants track a non-stationary Bernoulli distribution (Gallistel et al., 2014; Khaw et al., 2017; Ricci & Gallistel, 2017; Robinson, 1964) have consistently observed stepwise, "staircase-like" response patterns. It has been claimed that this pattern and related phenomena are inconsistent with trial-by-trial learning models and are instead indicative of discrete, stepwise learning through hypothesis testing (Gallistel et al., 2014; Ricci & Gallistel, 2017). This claim constitutes a serious challenge to the neuropsychological literature which connects trial-by-trial learning of probabilities (Nassar et al., 2012, 2010; Norton et al., 2019; Wilson et al., 2013, 2018), encoding of prediction errors in the anterior cingulate cortex (Behrens et al., 2007; Rushworth & Behrens, 2008; Silvetti et al., 2013) and the experience of surprise (Lavín et al., 2014; Preuschoff et al., 2011).

In the present paper, we argue that the rejection of trial-by-trial learning in human probability estimation was premature because it was based on an incomplete investigation of the predictions made by delta-rule models: parameter space was explored manually and no model fitting was performed. To reassess the earlier drawn conclusions, we reanalysed data from three previous experiments (Gallistel et al., 2014; Khaw et al., 2017; Ricci & Gallistel, 2017) using rigorous model fitting and model comparison methods. Our findings demonstrate that a dual process of two broadly supported computational theories – the delta-rule for online learning of a latent variable and the drift-diffusion model for making categorical decisions – makes predictions that are qualitatively highly consistent with the observed phenomena. We thereby account for them by reference to the assumptions of two of the most well-established theories of learning and evidence accumulation rather than by introducing new assumptions that are specifically tailored to account for said phenomena. Moreover, quantitative model comparison showed that the delta-rule model actually accounts *better* for the data than the proposed IIAB model in which learning proceeds through hypothesis testing. These conclusions hold across all tested data sets and are robust to changes in the modelling assumptions about the shape of the response threshold distribution, the assumed lapse rate, and the presence of response noise. In the (paraphrased) words of Mark Twain (White, 1897), we conclude that the report of the death of trial-by-trial estimation models was an exaggeration. We will immediately add, however, that we do not take this to imply the death of hypothesis testing models. Ultimately, we would expect – as is true in most areas of cognitive science – the mind to be able to draw on several different cognitive processes to estimate a property so fundamental to adaptation as probability. Our central claim here is that two of those might be delta-rule learning and drift diffusion decision making.

Theoretical importance and implications of a variable response threshold

Adding a variable response threshold greatly improves model fits. One reason is that participants make inconsistent updates which are incompatible with the original models, why their likelihoods are punished each time an inconsistent update occurs. The variable threshold allows the models to account for this. The response threshold thus does not merely "soak up noise" but is required by both the IIAB and delta-rule model to explain inconsistent updating and other empirical phenomena (Table 1). We therefore emphasise that a variable response threshold does not represent a "nuisance term", akin to adding an error term to a regression, but constitutes a theoretical proposition which is tentatively supported by our results.

Evaluation of the fitted response thresholds revealed that many distributions were so broad that the choice of whether or not to update on any given trial becomes partly stimulus-independent. Completely stimulus-independent thresholds have elsewhere (Biele et al., 2009; Gonzalez & Dutt, 2011) been termed "inertia". For two of the 29 participants, a coin-flip mechanism did indeed provide a better quantitative fit than the response threshold mechanism. However, for the vast majority of participants it did not, which suggests that updating is at least in part driven by stimulus-dependent factors (as also concluded by Khaw et al., 2017). For other participants, we obtained threshold distributions such that the probability of updating the response increased with the discrepancy between the current response and the internal estimate. Updates were disproportionately unlikely under very small discrepancies and disproportionately likely under very large discrepancies. We interpret this as a *resistance* to updating, as opposed to a suppressive *threshold* – the term we have hitherto used. Participants are reluctant to update (perhaps due to the motor cost) but balance this against their wish to respond correctly. They care about not being *very* wrong, but not so much about being *exactly* right. In economics, the idea that learning can be influenced by a trade-off between the costs of updating and the gains from a more accurate belief has been formalised in the "rational inattention" literature (Sims, 2003). The stepwise response pattern in the present Bernoulli distribution task has been taken to support this idea (Khaw et al., 2017). The reluctance interpretation of our threshold distributions stated above is different to rational inattention in that it supposes that the overt *response,* and not the covert *belief,* is affected by the trade-off. Our modelling here does not answer which version is correct and we do not hold our findings against rational inattention as a framework. We merely raise this point to caution against too high "blanket" confidence in belief level interpretations, which might be appropriate for some tasks but not for others.

Inertia and resistance (or rational inattention) are, seemingly, two distinct theoretical propositions as to how the mind times response updates. It may be that there is true heterogeneity in what mechanisms are used or that there exists a single mechanism which can express itself in two (ostensibly) different ways. Regardless of how internal estimates are updated, the process which mediates their expression as overt behaviour is scientifically interesting in itself and deserves further attention.

Observation weighting is intrinsic to the theories

We equalised the delta-rule model and the IIAB model on the assumption of a variable response threshold to show that this, although important, is not what drives the conclusions. Another difference is that the delta-rule model effectively performs *unequal weighting* of *all* observations while the IIAB model performs *equal weighting* of a *substring* of observations (those that occurred since the last or second to last change point). The weighting schemes are defining features of the theories the models embody. The IIAB model implements the theory that "the perception of Bernoulli probability is a by-product of the real-time construction of a compact encoding of the evolving sequence by means of change points" (Gallistel et al., 2014). Under unequal weighting of observations, the model contradicts this theory – the percept is no longer deduced from the change points. Associative theories instead suppose that the percept is no by-product but *learned in itself* by gradual adaptation. The delta-rule model has no conception of change points and can therefore not use them to define the relevant observations. The way that observations are weighted thus cannot be held constant across models; instead, it is an integral and defining feature of the mechanisms that we have sought to contrast.

Alternative models

We also found that a delta-rule which simultaneously estimates two kernels (Supplemental Materials) performs better than the regular, one-kernel delta-rule. Taken literally, this model continuously entertains two beliefs and selects one to report on each trial.

However, studies have indicated that learning rates in similar tasks are not fixed but adapted as a function of the prediction error modulated by the estimated volatility (Behrens et al., 2007; McGuire et al., 2014) or possibly other aspects of the choice environment (Lee et al., 2020). With this in mind, one could view the two-kernel model as an analogue for a single kernel model with an adaptive learning rate. We are therefore reluctant to interpret our results as evidence that people actually simultaneously hold dual beliefs about a single probability. Future studies might want to pool a larger number of datasets (from non-Bernoulli distribution tasks too) and compare various adaptive learning rate models to multi-kernel models. A multi-kernel interpretation also suggests that people should be able to report several earnest estimates at any one point, which should be possible to observe in an experiment.

Despite the supposed importance of an adaptive learning rate, an approximately Bayesian delta-rule model from the neuropsychological literature (Nassar et al., 2010) performs better than the hypothesis testing model but worse than the regular delta-rule. Allowing it to underweight likelihoods helps, in line with a previous observation (Nassar et al., 2010). However, with this change the model’s original theoretical claim (that people are approximate Bayesians who balance two hypotheses trial-by-trial) becomes less distinct from the more general notion of the learning rate being inconstant. Our tentative interpretation is that the common problem of the (normative) Bayesian delta-rule and the IIAB is not that they adapt observation weights (which is supported by other evidence, see Behrens et al., 2007; Krugel et al., 2009) but could be that they do this by considering a limited number of discrete hypotheses.

Costello and Watts (2014, 2016, 2018) have suggested that a range of results from various probability judgement and decision tasks, including the present paradigm, could arise from normatively correct judgements being perturbed by constant memory noise. They simulated a hypothesis testing model (Costello & Watts, 2018) with the same two stages/three hypotheses structure as the IIAB. They argue that, if there is constant memory noise, updates from re-estimation will be biased towards 0.5 and updates from acceptance of a new hypothesis will be biased towards the extremes. These effects should cancel out, making the estimates accurate on average (phenomenon 4, Table 1). If estimates are actually made trial-by-trial, and hypothesis testing is a separate drift-diffusion process, Costello and Watts’s (2018) framework predicts a constant bias towards 0.5, which seems inconsistent with the available data.

Unexplained phenomena

Phenomena 10 and 11 (Table 1) cannot be explained by neither the IIAB nor the delta-rule model as implemented here. However, both models could in principle be extended to do so.

Gallistel et al. (2014) noted that participants' frequency of change point reports on average decreased per session (phenomenon 10). They concluded that the IIAB cannot explain this under the regular priors used to explain the other qualitative phenomena, and that they had to substitute special priors tailored to this summary statistic. It is, however, easy to imagine a process where the threshold in the troubleshooting stage, T_2, is not fixed but adapted over time. This would result in a changing change point detection frequency. Analogously, the drift-diffusion literature explains this kind of effect as decision bound separation being adapted through learning (Liu & Watanabe, 2012; Zhang & Rowe, 2014).

In Ricci and Gallistel (2017), some participants were able to correctly report having drawn from a sinusoidal during the debriefing (phenomenon 11). A central theoretical proposition of the IIAB (see pp. 106, Gallistel et al., 2014) is that people do not perceive probabilities per se but "deduce" them from a (sparse) memory of change points. To generate a declarative belief of a continuous functional form from a discrete set of memories, it would require some function learning mechanism (e.g., Brehmer, 1974) which interpolates between the "datapoints". For the delta-rule model, we need the mechanism to be recursive. There exist several such function learning models, some of which are specifically adapted to non-stationary environments (Speekenbrink & Shanks, 2010) and some of which use a version of delta-rule learning (DeLosh et al., 1997). The perhaps most famous of the latter is the EXAM model (Mcdaniel & Busemeyer, 2005).

In sum, we do not view phenomena 10 and 11 as evidence against either model but rather as avenues of future research. Investigating phenomenon 10 involves opening a black box by trying to establish a structured explanation of aspects which we here model as free parameters. Investigating phenomenon 11 would involve attaching a third process of function learning to what we suggest could be a dual process of delta-rule online learning and drift diffusion decision making (in line with the "Linnaean" approach to cognition; Millroth et al., 2021).

Limitations of modelling

The trial-by-trial learning models tested here are recursive: they update a compact knowledge state and do not require any sequence memory. However, any recursive function can be reformulated as an iterative function (Church, 1936b, 1936a; Turing, 1937) which repeatedly generates a new knowledge state from a sequence memory. Hence, to what extent people retain the sequences they have observed is ultimately not a question that can be answered by model comparison alone. We have demonstrated that a recursive, compact knowledge state model is *possible,* which it previously was thought not to be, but future studies should perform falsification tests (Popper, 1968) through experimental manipulation.

Concluding remarks

We have demonstrated that it was premature for the previous literature to rule out trial-by-trial learning models of probability perception. In the spirit of cumulative science (Walter Mischel, 2009), the raw data and observed phenomena can be better explained by a dual process of delta-rule online learning and drift-diffusion evidence accumulation. That being said, this previous research has highlighted that a complete theory of probability perception must account for hypotheses about the generative process and how these affect our online estimates. Outside the laboratory, probabilities are learnt from experiences in their context. It seems likely that external, higher-level beliefs about this context – about volatility, sequentiality and trends in the generative process – can influence our online beliefs.

CREDIT AUTHOR STATEMENT

Mattias Forsgren: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Reviewing & Editing. **Peter Juslin:** Conceptualization, Methodology, Formal Analysis, Writing – Original Draft, Writing – Reviewing & Editing, Supervision, Project Administration, Funding Acquisition. **Ronald van den Berg:** Conceptualization, Methodology, Software, Validation, Formal Analysis, Writing – Original Draft, Writing – Reviewing & Editing, Visualization, Supervision, Project Administration, Funding Acquisition.

REFERENCES

Adams, R. P., & MacKay, D. J. C. (2007). *Bayesian Online Changepoint Detection*.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In *Handbook of categorization in cognitive science, 2nd ed.* (pp. 157–188). Elsevier Academic Press. https://doi.org/10.1016/B978-0-08-101107-2.00007-5

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X07001653

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*. https://doi.org/10.1038/nn1954

Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology*, *53*(3), 155–167. https://doi.org/10.1016/j.jmp.2008.05.006

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*. https://doi.org/10.1016/0030-5073(74)90002-6

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica*. https://doi.org/10.1016/0001-6918(80)90034-7

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). A study of thinking. In *A study of thinking*. John Wiley and Sons.

Busemeyer, J. R., & Myung, I. J. (1988). A New Method for Investigating Prototype Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/0278-7393.14.1.3

Church, A. (1936a). A note on the Entscheidungsproblem. *Journal of Symbolic Logic*. https://doi.org/10.2307/2269326

Church, A. (1936b). An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics*. https://doi.org/10.2307/2371045

Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*. https://doi.org/10.1037/a0037010

Costello, F., & Watts, P. (2016). People’s conditional probability judgments follow probability theory (plus noise). *Cognitive Psychology*, *89*, 106–133. https://doi.org/10.1016/j.cogpsych.2016.06.006

Costello, F., & Watts, P. (2018). Probability Theory Plus Noise: Descriptive Estimation and Inferential Judgment. *Topics in Cognitive Science*, *10*(1), 192–208. https://doi.org/10.1111/tops.12319

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 968–986. https://doi.org/10.1037/0278-7393.23.4.968

Ditterich, J. (2006). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks*, *19*(8), 981–1012. https://doi.org/10.1016/j.neunet.2006.05.042

Drugowitsch, J., Wyart, V., Devauchelle, A.-D. D., & Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*, *92*(6), 1398–1411. https://doi.org/10.1016/j.neuron.2016.11.005

Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology*. https://doi.org/10.1037/h0041970

Estes, W. K. (1976). The cognitive side of probability learning. *Psychological Review*. https://doi.org/10.1037/0033-295X.83.1.37

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews. Neuroscience*, *9*, 292–303. https://doi.org/10.1038/nrn2258

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*. https://doi.org/10.1037/0033-295X.107.4.659

Gallistel, C. R., Krishan, M., Liu, Y., Miller, R., & Latham, P. E. (2014). The perception of probability. *Psychological Review*. https://doi.org/10.1037/a0035232

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*. https://doi.org/10.1037/0033-295X.102.4.684

Gonzalez, C., & Dutt, V. (2011). Instance-Based Learning: Integrating Sampling and Repeated Decisions From Experience. *Psychological Review*. https://doi.org/10.1037/a0024558

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. In *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2009.09.004

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. *The Cambridge Handbook of Thinking and Reasoning*.

Khaw, M. W., Stevens, L., & Woodford, M. (2017). Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*. https://doi.org/10.1016/j.jmoneco.2017.09.001

Klayman, J., & Ha, Y. W. (1987). Confirmation, Disconfirmation, and Information in Hypothesis Testing. *Psychological Review*. https://doi.org/10.1037/0033-295X.94.2.211

Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S. C., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of*

America. https://doi.org/10.1073/pnas.0905191106

Lavín, C., Martín, R. S., & Jubal, E. R. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*. https://doi.org/10.3389/fnbeh.2013.00218

Lee, S., Gold, J. I., & Kable, J. W. (2020). The human as delta-rule learner. *Decision*, *7*(1), 55–66. https://doi.org/10.1037/dec0000112

Liu, C. C., & Watanabe, T. (2012). Accounting for speed–accuracy tradeoff in perceptual learning. *Vision Research*, *61*, 107–114. https://doi.org/10.1016/j.visres.2011.09.007

Mcdaniel, M. A., & Busemeyer, J. R. (2005). The conceptual basis of function learning and extrapolation: Comparison of rule-based and associative-based models. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/BF03196347

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally Dissociable Influences on Learning Rate in a Dynamic Environment. *Neuron*. https://doi.org/10.1016/j.neuron.2014.10.013

Millroth, P., Collsiöö, A., & Juslin, P. (2021). Cognitiva Speciebus: Towards a Linnaean Approach to Cognition. *Trends in Cognitive Sciences*, *25*(3), 173–176. https://doi.org/10.1016/j.tics.2020.12.005

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*. https://doi.org/10.1038/nn.3130

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0822-10.2010

Neal, R. M., & Dayan, P. (1997). Factor Analysis Using Delta-Rule Wake-Sleep Learning. *Neural Computation*. https://doi.org/10.1162/neco.1997.9.8.1781

Norton, E. H., Acerbi, L., Ma, W. J., & Landy, M. S. (2019). Human online adaptation to changes in prior probability. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1006681

Oaksford, M., & Chater, N. (1994). A Rational Analysis of the Selection Task as Optimal Data Selection. *Psychological Review*. https://doi.org/10.1037/0033-295X.101.4.608

PETERSON, C. R., & BEACH, L. R. (1967). MAN AS AN INTUITIVE STATISTICIAN. *Psychological Bulletin*. https://doi.org/10.1037/h0024722

Popper, K. (1968). *The Logic of Scientific Discovery* (2 rev. ed.). Hutchinson.

Preuschoff, K., ’t Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise:

Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2011.00115

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Rescorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement BT - Clasical conditioning II: current research and theory. In *Clasical conditioning II: current research and theory*.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*. https://doi.org/10.1038/nature08275

Ricci, M., & Gallistel, R. (2017). Accurate step-hold tracking of smoothly varying periodic and aperiodic probability. *Attention, Perception, and Psychophysics*. https://doi.org/10.3758/s13414-017-1310-0

Robinson, G. H. (1964). Continuous Estimation Of A Time-Varying Probability. *Ergonomics*, *7*(1), 7–21. https://doi.org/10.1080/00140136408930721

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. In *Nature Neuroscience*. https://doi.org/10.1038/nn2066

Silvetti, M., Seurinck, R., & Verguts, T. (2013). Value and prediction error estimation account for volatility effects in ACC: A model-based fMRI study. *Cortex*. https://doi.org/10.1016/j.cortex.2012.05.008

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690. https://doi.org/10.1016/S0304-3932(03)00029-1

Speekenbrink, M., & Shanks, D. R. (2010). Learning in a changing environment. *Journal of Experimental Psychology: General*, *139*(2), 266–298. https://doi.org/10.1037/a0018620

Turing, A. M. (1937). On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*. https://doi.org/10.1112/plms/s2-42.1.230

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*. https://doi.org/10.1037/0033-295X.90.4.293

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*.

https://doi.org/10.1007/BF00122574

van den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N. M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. *ELife*, *5*(FEBRUARY2016). https://doi.org/10.7554/eLife.12192

Verguts, T., & Van Opstal, F. (2014). A delta-rule model of numerical and non-numerical order processing. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/a0035114

Wagenmakers, E. J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*(5), 641–671. https://doi.org/10.1080/09541440802205067

Walter Mischel. (2009). Becoming a Cumulative Science. *APS Observer*, *22*(1).

WASON, P. C., & JOHNSON-LAIRD, P. N. (1970). A CONFLICT BETWEEN SELECTING AND EVALUATING INFORMATION IN AN INFERENTIAL TASK. *British Journal of Psychology*. https://doi.org/10.1111/j.2044-8295.1970.tb01270.x

White, F. M. (1897, June). Mark Twain amused, humorist says he even heard on good authority that he was dead. *New York Journal and Advertiser*.

Widrow, B., & Hoff, M. E. (1960). Adaptive Switching Circuits. In *Technical report no. 1553-1*.

Widrow, B., & Lehr, M. A. (1993). ARTIFICIAL NEURAL NETWORKS OF THE PERCEPTRON, MADALINE, AND BACKPROPAGATION FAMILY. In *Neurobionics*. https://doi.org/10.1016/b978-0-444-89958-3.50013-9

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2013). A Mixture of Delta-Rules Approximation to Bayesian Inference in Change-Point Problems. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1003150

Wilson, R. C., Nassar, M. R., Tavoni, G., & Gold, J. I. (2018). Correction: A Mixture of Delta-Rules Approximation to Bayesian Inference in Change-Point Problems. *PLOS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1006210

Zhang, J., & Rowe, J. B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in Neuroscience*, *8*(7), 651–687. https://doi.org/10.3389/fnins.2014.00069