Imprecise Probabilistic Inference from Sequential Data

Arthur Prat-Carrabin
Columbia University

Michael Woodford
Columbia University

Abstract

Although the Bayesian paradigm is an important benchmark in studies of human inference, the extent to which it provides a useful framework to account for human behavior remains debated. We document systematic departures from the predictions of Bayesian inference, even on average, in the estimates by experimental subjects of the probability of a binary event following observations of successive realizations of the event. In particular we find under-reaction of subjects' probability estimates to the evidence ('conservatism') after only a few observations, and at the same time over-reaction after a longer sequence of observations. This is not explained by an incorrect prior, nor by many common models of Bayesian inference. We uncover the autocorrelation in estimates, which suggests that subjects carry imprecise representations of the decision situations, with noise in beliefs propagating over successive trials. But even taking into account these internal imprecisions, we find that the subjects' updates are inconsistent with the rules of Bayesian inference. We show how subjects instead considerably economize on the attention that they pay to the information relevant to decision, and on the degree of control that they exert over their precise response, while giving responses fairly adapted to the experimental task. A "noisy counting" model of probability estimation reproduces the several patterns we exhibit in subjects' behavior. In sum, human subjects in our task perform reasonably well while greatly minimizing the amount of information that they pay attention to. Our results emphasize that investigating this economy of attention is crucial in understanding human decisions.

Author Note

Arthur Prat-Carrabin https://orcid.org/0000-0001-6710-1488
Michael Woodford https://orcid.org/0000-0001-5485-5280

We thank Carlos Alos-Ferrer, Duarte Gonçalves, Frank Heinemann and Simas Kucinskas for helpful comments, Justinas Grigaitis and Haoran Wang for their help as research assistants, and the National Science Foundation for research support (grant SES DRMS 1949418). Parts of this study have been presented at the 2020 Marlene Porsche Graduate School of Neuroeconomics Symposium, the 2021 Society for NeuroEconomics Annual Meeting, the 2022 Annual Meeting of the Cognitive Science Society, and the 2022 Computational Psychiatry Course. Requests for the data and the code can be sent via email to the corresponding author. This study was not preregistered. Correspondence concerning this article should be addressed to Arthur Prat-Carrabin. Email: arthur.p@columbia.edu.

Introduction: Are People "Intuitive Statisticians"?

A fundamental question about human decision making under uncertainty concerns the accuracy with which people are able to assess the probabilities of different possible outcomes. A considerable body of evidence suggests that people are not good at estimating the Bayesian posterior probabilities of the few hypotheses that could underlie the evidence presented to them (typically, just two hypotheses, in 'bookbag-and-poker-chips' tasks; Benjamin, 2019). At the same time, an important strand of the literature has argued that people are surprisingly good at estimating probabilities under certain circumstances — that man is "an intuitive statistician," in the phrase of Peterson and Beach (1967) — namely, ones in which a subject must estimate a probability, on the basis of a sequence of random samples, each obtained with this probability (e.g., in Gallistel *et al.,* 2014, subjects are asked to estimate the proportion of red vs. green rings in a box, on the basis of random samples from the box.) Since the latter kind of task is arguably of greater ecological relevance, one might wonder if it is not reasonable to assume that people base their decisions on essentially correct probability beliefs, despite their poor grasp of abstract principles of statistical inference.

More precisely, one might wonder if these studies do not support the validity of the hypothesis of "rational expectations" as formulated by Muth (1961), which proposed that while individual forecasters do not make correct statistical forecasts, their forecasts are *on average* correct.[1] For example, Figure 1A plots (on the vertical axis) the probability estimates of subjects in the study of Khaw *et al.* (2017a) as a function of the correct Bayesian posterior mean probability, given the evidence shown to that point. There is evidently a great deal of error in individual estimates; but as shown in panel B, the average response, conditional on the correct Bayesian estimate, is close to the diagonal. This result has a special importance in the field of economics, if one supposes that to the extent that market outcomes depend only on the aggregate behavior of a large number of people who form independent beliefs, these market outcomes will (as in Muth's analysis) be the same as in a world where market participants were all perfect Bayesians.

Here we reconsider this question, by focusing, like such studies as Peterson and Beach (1967), Gallistel *et al.* (2014), and Khaw *et al.* (2017a), on a task in which subjects must estimate the probability of a binary event on the basis of a sequence of successive realizations, with a new probability estimate elicited after each new realization. But we simplify the environment, relative to many classic studies, in a crucial respect. While the probability of the event shifts over time in studies like Gallistel *et al.* (2014), and at times unknown to the subject ('change points',) in our experiment the unknown probability remains the same for a fixed number of trials, and the subject is always told when a new probability has been drawn. This means that in our experiment a subject will repeatedly be asked to estimate the probability under precisely the same sequence of evidence that they have previously experienced. The fact that a small number of possible evidentiary states are repeatedly experienced allows us to document in detail each subject's distribution of responses to each possible evidentiary state.

[1] In more recent economic literature, the "rational expectations hypothesis" is usually taken to assert that all agents are perfect Bayesian statisticians. However, Muth (1961) asserts only the accuracy of "averages of expectations in an industry" (p. 316), and denies that his hypothesis asserts "that predictions of entrepreneurs are perfect or that their expectations are all the same" (p. 317).

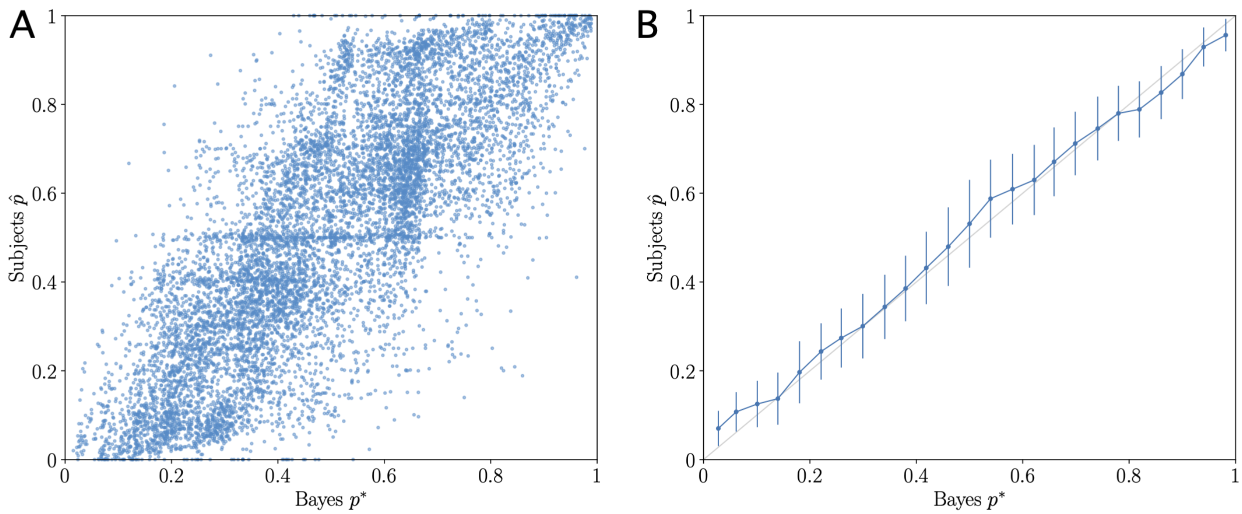

Figure 1: Subjective probability estimates plotted against the correct Bayesian posterior mean, in an experiment with occasional shifts in the true probability. (A) All trials on which the subject changes their reported estimate. (B) Mean estimates, and standard error of the mean (whiskers), for each range of values of the correct Bayesian estimate; the horizontal location of each of these ranges corresponds to the mean Bayesian estimate for that bin. (Data: Khaw *et al.*, 2017b.) See Khaw *et al.* (2017a) for details of the experiment and the method of calculation of the correct Bayesian posterior probability beliefs.

Just as with the data shown in Figure 1, we find in the case of our simpler task that a linear regression of the subjective estimate on the correct Bayesian posterior mean given the evidence available on that trial would yield a line close to the diagonal (see Figure 3 below). Nonetheless, we find systematic departures from correct Bayesian estimates on the part of our subjects, even on average. However, the biases in our subjects' estimates exhibit neither systematic "conservatism" of the kind found, for example, by Phillips and Edwards (1966), nor the systematic over-reaction to evidence of the kind reported by Brown and Bane (1975). Instead, we find a conservative bias in subjects' estimates after only a small amount of evidence about a new regime has been presented; but after further evidence is observed the bias switches to over-reaction (to the cumulative evidence). Both of these biases are pervasive in our data, but would be obscured if we were to look only for a static relationship between the optimal Bayesian estimate and subjects' responses, as in Figure 1. We conjecture that they are also obscured in the case of an experimental design, like the ones just described with random change points, in which there is less experimental control of the degree of evidence that the subjects believe themselves to have at each point in time.

We further investigate the source of these patterns of bias, and conclude that they are consistent with a relatively simple view of our subjects' behavior. Rather than basing their responses on a precise awareness of the body of evidence presented to them, but using an incorrect formula to produce a probability estimate from that information, it seems that (i) subjects base their responses on very little information, and (ii) they do not precisely control their exact response, allowing an element of arbitrary randomness in it. Subject to these constraints, their responses are about as accurate as they can be.

Moreover, it appears that rather than considering on each trial what the probability estimate should be, given the body of evidence provided to that point, subjects instead consider only how much they should *adjust* their existing estimate in the light of the new evidence just received. Of course, if such an adjustment decision were made on the basis of information that included a precise awareness of one's current estimate, making a decision about how much to *change* one's estimate and making a decision about *what estimate* to announce should be equivalent problems. But our hypothesis is that the decision about how much to adjust the estimate takes little account of what the existing estimate is; it pays attention to the last binary outcome that has been observed — indeed, we find that the sign of the adjustment nearly always agrees with this latest piece of evidence — but to very little else. Deciding on the change to make on the basis of this limited information set has very different implications than a decision about the estimate to announce on the basis of the same information set, and the model of limited-information *changes* is much more consistent with our data.

The idea that people may have a better idea of how accurately the change in their estimate matches the latest increment to the evidence than of how the absolute size of their estimate matches the cumulative import of the evidence that they have seen is consistent with a common property of perception in sensory domains, where it is often observed that people can more accurately judge changes or differences in some sensory magnitude (say, height) than they can judge the absolute magnitude of an individual stimulus (Laming, 2011). Another common feature of judgments about sensory magnitudes is that they are random. Similarly, we find that a given increment to the observed evidence results in an adjustment of subjects' estimates by a random amount, rather than a deterministic adjustment as would be implied by correct Bayesian inference.

We thus find that many features of our subjects' estimates are consistent with a "noisy-counting" model of probability estimation. In this model, subjects adjust their estimate up or down (depending on the sign of the most recent observation) by a random amount, and keep track only of the estimate that they have reached through this process, which is essentially a noisy count of the net number of pieces of evidence in one direction or the other.

Because the distribution of possible adjustments on any given trial is assumed not to be conditioned on information other than the most recent observation, it is not possible for adjustment to be larger in the case of early pieces of evidence and smaller in the case of later observations, as would be true under correct Bayesian updating (with correct beliefs about the process generating the observations). Even under the hypothesis that the average size of adjustments is optimized (taking as given the constraint on the information upon which the adjustment can be conditioned, and the amount of random variation in the adjustment on each trial), it will inevitably be the case that adjustments will be too small (relative to the Bayesian benchmark) on the early trials after a change in the true probability, and too large on the later trials. Thus the pattern of early under-reaction and later over-reaction that we observe is precisely what such a model predicts.

In section 1, we explain our experimental design, and discuss the ways in which average responses under each informational condition are systematically biased relative to the Bayesian benchmark. We further show that the biases cannot be summarized by a simple monotonic transformation of the probability indicated by the evidence into a distorted average response, as in the model of Zhang *et al.* (2020); instead, the biases are different depending on the amount of evidence that has been observed. In section 2, we present evidence indicating that subjects' responses are not simply biased, but that they appear to make use of only a limited-precision record of the evidence presented to that point, and indeed that the biases can largely be attributed to the limited informational basis for subjects' responses.

In section 3, we then ask whether subjects' estimates are optimal conditional on the limited information used in producing them, as in the Bayesian models of perceptual bias proposed by authors such as Wei and Stocker (2015) or Petzschner *et al.* (2015), and find that they are not. We also show that the subjects' estimates are not consistent with Bayesian estimation conditional on both noisy evidence and an incorrect prior, as in the model of bias in probability estimates proposed by Zhu *et al.* (2020). Instead, we conclude that in addition to being based on limited information, subjects' estimates incorporate a certain degree of random response error, in addition to the noise in their internal representation of the evidence that they have observed. In section 4 we examine a family of 'quasi-Bayesian' models inspired by the literature, in which the prior and the likelihood are under- or over-weighted in the application of Bayes' rule; but we do not find that it provides the most satisfying account of subjects' behavior. Finally, in section 5 we present our "noisy-counting" model, and show that simulations of this model have many of the properties of our subjects' estimates documented in the previous sections. Section 6 offers a concluding discussion.

1 Biases Relative to the Bayesian Benchmark

In this section, we explain our experiment, and document that subjects' probability estimates differ from correct Bayesian estimates even on average, contrary to the results shown in Figure 1 in the case of a more complex task.

1.1 Our behavioral task and the Bayesian benchmark

In a computer-based task, subjects are repeatedly asked to infer the proportion of red rings in a box containing red and green rings, based on the presentation of rings randomly drawn from the box. Specifically, an experimental session is divided in 200 blocks of five trials. At the beginning of each block, a new (virtual) box is prepared, with a proportion of red rings, p, randomly sampled from the uniform distribution on $[0, 1]$. In an effort to convey this uniform prior to the subjects, the task instructions indicate twice that "all proportions of red rings, from 0% to 100%, are equally likely". At each trial a ring is drawn from the box; its color, x, is a Bernoulli random variable that takes the value R (denoting a red ring) with probability p, and G (denoting a green ring) with probability $1 - p$. It is presented to the subject, who is then asked to provide an estimate, $\hat{p}$, of the proportion of red rings, using a slider (Fig. 2). The ring is then replaced in the box, and a new trial begins. After five trials, the block ends, and the subject is notified that a new block — with a new proportion of red rings — begins. For each estimate provided, the subject receives a number of points which is a decreasing affine function of the squared error, $(\hat{p} - p)^2$. (No feedback was provided, except in a short practice phase.) At the end of the experiment, the points are converted into a financial reward. The average reward received by the subjects was \$22. Additional

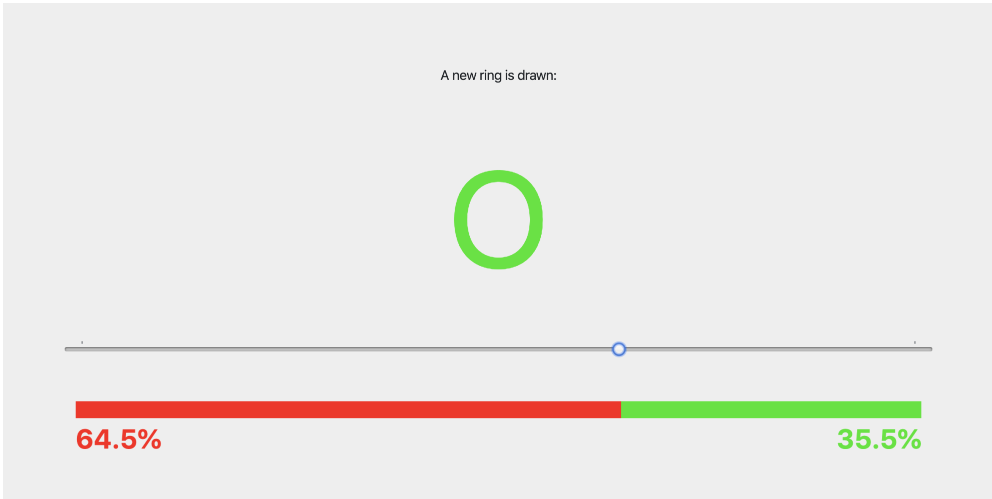

Figure 2: Subjects were asked to estimate the proportion of red rings in a box upon the presentation of random ring draws. At each trial, a ring is drawn (here, a green ring) and it is presented to the subject, who then uses a slider to provide her estimate of the proportion of red rings.

details on the task can be found in Methods. Finally, we also run a variant of the experiment in which blocks of trials do not end after a fixed length of five trials, but instead end with probability 0.2 after each trial (resulting in geometrically-distributed sequence lengths). The results pertaining to this variant strengthen our conclusions. We present them in Methods.

The posterior of a Bayesian observer inferring the proportion of red rings, in the task just presented, is a Beta distribution with parameters $n_R + 1$ and $t - n_R + 1$, where t is the number of rings observed, and n_R the number of red rings among them. The optimal estimate, obtained by maximizing the expected reward, is the posterior mean,

$$p^* = \frac{n_R + 1}{t + 2}. \tag{1.1}$$

We denote by s a given sequence of ring draws (with length anywhere between one to five), and by $p^*(s)$ the optimal estimate implied by the sequence s. To respond optimally in the task (assuming one has the correct belief about the process underlying the observations), it is thus sufficient to keep track of two quantities — the number of red rings observed and the total number of rings — and to compute the ratio given by Eq. 1.1. We now compare the responses provided by the subjects to the optimal estimates.

1.2 Bias in the average estimates

Over the course of the 200 blocks of trials, many sequences of ring draws are sampled more than once. Although the optimal estimate is a deterministic function of the sequence, the responses of subjects to identical sequences are not identical. In other words, there is (intra-subject) variability in the responses. A possibility is that this variability is explained by the presence of motor noise, which may prevent the subjects from selecting the optimal slider position, but that on average the responses of subjects are equal to the optimal estimates. Looking at the average responses in comparison to the optimal estimates, we note that, in most cases, these two quantities are different (Fig. 3, blue solid line).

We point to the symmetry, around 1/2, of the subjects' average responses plotted against the optimal estimates (Fig. 3, blue solid line), which suggests that the deviations from optimality do not simply result from sampling noise. Nevertheless, we test the hypothesis that the average responses equal the optimal estimates. For this test, as for many of the tests we run below, we compute the p-value of the F statistics implied by the squared errors in the predictions of a 'restricted' hypothesis, which assumes that the data obey some constraints characterized by a given set of parameters, and the predictions of an alternative, less constraining, 'unrestricted' hypothesis, which comes with a greater number of parameters. We carry the F-tests both with the responses of all subjects, pooled together, and with the responses of each subject, taken separately. In the latter case, we report the proportion of subjects for whom the p-value of the test is below .01, and the median p-value across subjects.

We run the F-test of the restricted hypothesis that for any sequence of ring draws, s, the average of subjects' responses is equal to the optimal estimate ($\mathrm{E}[\hat{p}|s] = p^*(s)$), against the unrestricted hypothesis that the average response may be any function of the optimal estimate ($\mathrm{E}[\hat{p}|s] = m(p^*(s))$). The restricted hypothesis is rejected for all the subjects (Table 1, line 1). (We also conduct a series of t-tests, which support this result; see Table 6 in Methods). We thus turn to alternative hypotheses, that predict a bias in subjects' responses.

1.2.1 The average response not simply a monotonic transformation of the optimal Bayesian estimate

A simple hypothesis that results in a bias in responses is one in which the average response is a linear transformation of the optimal estimate, i.e., $\mathrm{E}[\hat{p}|s] = ap^*(s) + b$. A different approach, however, has been reported to account successfully for the biases of human subjects in a variety of studies involving probability and frequency estimations. In this approach, the distortion in a subject's estimate of a probability results from a linear transformation of the log-odds of the two outcomes (Zhang and Maloney, 2012; Zhang *et al.*, 2020). The log-odds (or logit) transformation of a probability is the function

$$\mathrm{Lo}(p) \;\equiv\; \ln\frac{p}{1-p}, \tag{1.2}$$

and the hypothesis we now investigate posits that the log-odds of the subject's response $\hat{p}$ is an affine function of the log-odds of the probability to be estimated, p, i.e., $\mathrm{Lo}(\hat{p}) = a \cdot \mathrm{Lo}(p) + b$, or equivalently

$$\hat{p} = \mathrm{Lo}^{-1}\big(a \cdot \mathrm{Lo}(p) + b\big), \tag{1.3}$$

where $\mathrm{Lo}^{-1}(\cdot)$ is the logistic function (the inverse of $\mathrm{Lo}(\cdot)$).

We test both the restricted hypothesis that the average of subjects' responses is a linear transformation of the optimal estimate ($\mathrm{E}[\hat{p}|s] = ap^*(s) + b$), and the restricted hypothesis that it results from a linear transformation of the log-odds of the optimal estimate (i.e., $\mathrm{E}[\hat{p}|s] = \mathrm{Lo}^{-1}\big((a \cdot \mathrm{Lo}(p^*(s)) + b)\big)$), against the unrestricted hypothesis, as in the previous section, that it may be any function of the optimal estimate ($\mathrm{E}[\hat{p}|s] = m(p^*(s))$). Both

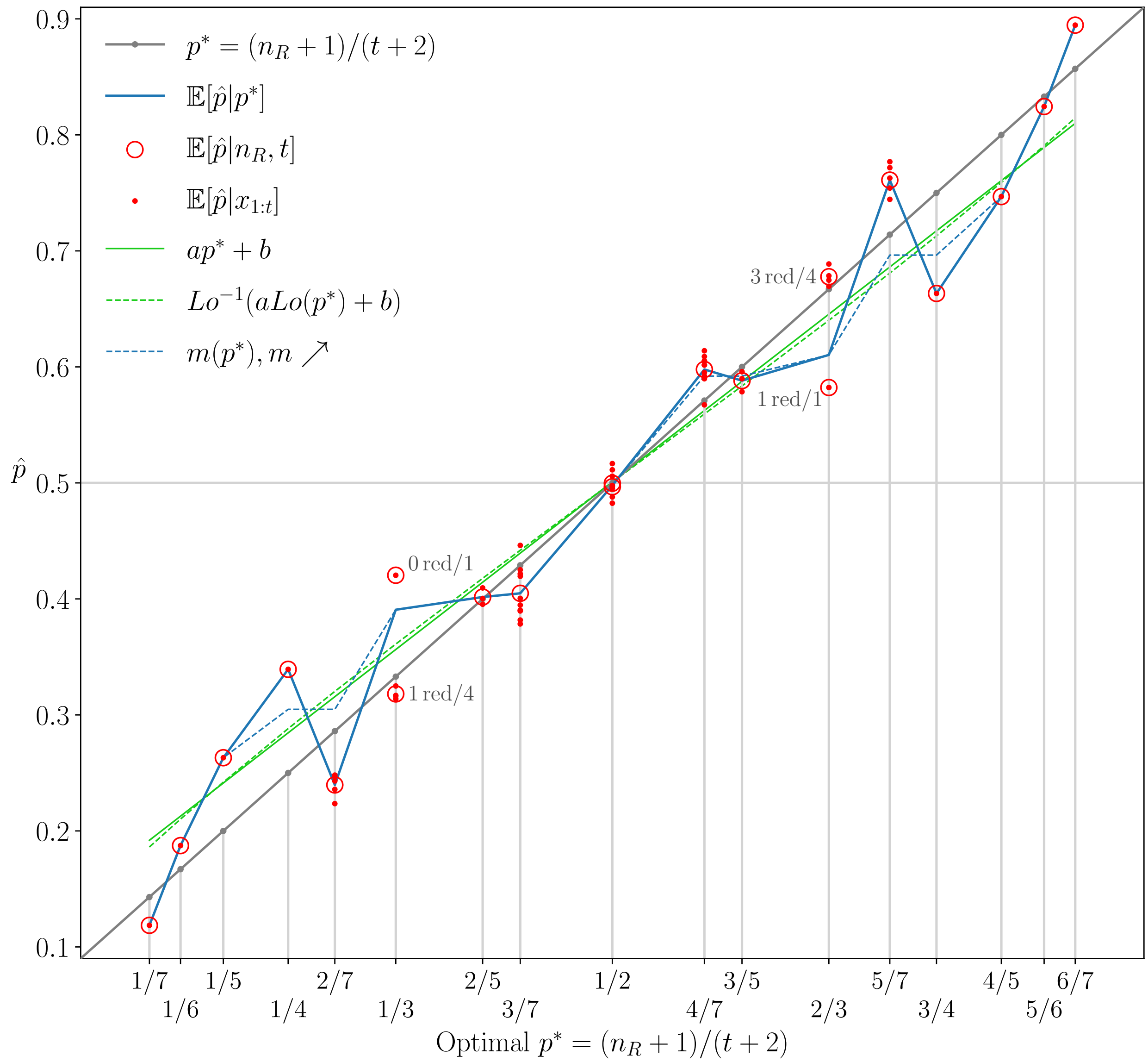

Figure 3: Subjects' responses are poorly captured by functions of the optimal Bayesian estimate alone. Subjects' average responses, and responses derived from various hypotheses, as a function of the optimal estimate p^*. *Grey line and dots:* optimal estimates. *Blue solid line*: average responses as a function of the optimal estimates. The responses to all the sequences that yield the same optimal estimate are pooled together. Largest standard error of the mean (sem): .004. *Red circles:* Average responses, when pooling together the responses to the sequences that have the same length, t, and contain the same number of red rings, n_R (thus the order of the rings may differ). Largest sem: .004. For $p^* = 1/3$ and $2/3$, two pairs (n_R, t) yield the same optimal estimate p^*; in these cases n_R and t are indicated. *Red dots:* Average responses to each of the 62 possible sequences in the task, as a function of the optimal estimate. Largest sem: .014 (median sem: .007). In all cases, the responses of all the subjects are pooled together. *Green solid line*: affine function of the optimal estimate that best fits subjects' responses (least squares). *Green dashed line*: best-fitting function of the optimal estimate such that the log-odds of an average response is an affine function of the log-odds of the optimal estimate. *Blue dashed line*: non-decreasing function of the optimal estimate that best fits subjects' responses.

restricted hypotheses are strongly rejected (Table 1, lines 2-3). We note that the two hypotheses result in an average response that is an increasing function of the optimal estimate, and which underestimates the probabilities greater than .5 and overestimates the probabilities smaller than .5 (Fig. 3, green lines). Subjects' responses, however, exhibit more intricate patterns, which both approaches fail to capture. We now examine these patterns more closely.

Over the range of the different optimal estimates implied by the various sequences of rings that appear in the experiment, the responses of subjects seem alternatively below and above the optimal estimate, and the curve of the average response as a function of the optimal estimate crosses nine times the first bisector. Besides, although at first glance the subjects' responses seem to increase as a function of the optimal estimate, a close inspection reveals that between the two increasing optimal estimates 5/7 and 3/4 (and between the symmetrical estimates 1/4 and 2/7,) the subjects' average responses decrease sharply (Fig. 3, blue solid line). In other words, when presented with two sets of evidence, the Bayesian observer concludes that the first one suggests a larger proportion of red rings than the second one, but the subjects draw the *opposite* conclusion: that the first set of evidence suggests a smaller proportion of red rings than the second one.

To test whether subjects' responses do decrease where optimal estimates increase, we consider the restricted hypothesis that the average of subjects' responses to a sequence of ring draws, s, is a non-decreasing function, m, of the optimal estimates, $p^*(s)$. We fit the function m to subjects' data, with the constraint that it not be decreasing, a problem known as isotonic least-square regression (Best, 1990; Pedregosa *et al.,* 2011; Fig. 3, blue dashed line). As for the unrestricted hypothesis, we note that the pair of optimal estimates mentioned above, between which subjects' responses seem to decrease, correspond to sequences of two different lengths: two rings (for $p^* = 1/4$ and $3/4$) and five rings (for $p^* = 2/7$ and $5/7$). Looking at the responses of subjects separately for each sequence length, it appears that the average response of subjects, conditional on a given sequence length, increases as a function of the optimal estimate (in contrast with the responses unconditional on the sequence length; Fig. 4A). Thus we run the F-test of the restricted hypothesis mentioned above ($\mathrm{E}[\hat{p}|s] = m(p^*(s)), m$ non-decreasing) against the weaker hypothesis that the average response is a function of both the optimal estimate and the number of draws in the sequence, $m(p^*(s), n(s))$, such that for each sequence length, n, the function $p^* \mapsto m(p^*, n)$ is a non-decreasing function of the optimal estimate, p^*. (This last constraint, however, proves non-binding, for all subjects.) The restricted hypothesis is strongly rejected (Table 1, line 4). We conclude that the responses of the subjects are not well captured by any non-decreasing function of the optimal estimate.

Furthermore, the analysis just presented suggests that subjects' responses are not satisfactorily predicted by the optimal estimate alone, and that they seem to also vary with the number of rings presented. We point to two other results that support this hypothesis. First, if the average response of subjects were a function of the optimal estimate and did not vary with the number of samples, then two sequences of ring draws that implied the same optimal estimate, but that had different lengths, should yield the same response. For instance, the one-draw sequence containing one red ring and the four-draw sequences containing three red rings and one green ring all result in the same optimal estimate, $p^* = 2/3$. However, the average responses of subjects in these two cases exhibit a sizable and statistically significant difference of .096 (Fig. 3, red circles; t-test p-value, pooling all subjects: 1.5e-272; individual tests: p-value $< .01$ for 95% of subjects; median p-value: 5.6e-23). Second, we run an F-test of the restricted hypothesis that the average response is a function of the optimal estimate ($\mathrm{E}[\hat{p}|s] = m(p^*(s))$) against the unrestricted hypothesis that the average response is a function of both the optimal estimate and the number of rings presented ($\mathrm{E}[\hat{p}|s] = m(p^*(s), n(s))$), and find that the restricted hypothesis is rejected at the .01 level for 95% of the subjects (Table 1, line 5). We conclude that the average response of subjects is not accounted for by a transformation of the sole optimal estimate, but that it also varies with other features of the observed sequence of ring draws, such as the length of the sequence. Hence, we examine more closely, below, the responses of subjects in response to different numbers of rings presented.

	Restricted hypothesis $\mathrm{E}[\hat{p}\|s] = \ldots$	*Unrestricted hyp.* $\mathrm{E}[\hat{p}\|s] = \ldots$	*Subjects pooled* p-value	*Individual tests* $\% < .01$	median $p$val.
(1)	$p^*(s) = (n_R + 1)/(t + 2)$	$m(p^*(s))$	0.0***	1.	1.4e-151
(2)	$ap^*(s) + b$	$m(p^*(s))$	0.0***	.95	6.1e-98
(3)	$\mathrm{Lo}^{-1}\big(a \cdot \mathrm{Lo}(p^*(s)) + b\big)$	$m(p^*(s))$	0.0***	.95	2.1e-95
(4)	$m(p^*(s)) \nearrow$	$m(p^*(s), n(s)) \nearrow$	0.0***	.95	4.7e-79
(5)	$m(p^*(s))$	$m(p^*(s), n(s))$	3.7e-322***	.95	7.3e-42
(6)	$p_\alpha(s) = \frac{n_R+\alpha}{t+2\alpha}$	$m(p^*(s), n(s))$	0.0***	1.	2.9e-111
(7)	$p_{\alpha\rho\lambda}(s)$ (see Eq. 4.7)	$m(s)$	1.46e-54***	.81	2.6e-17

Table 1: F-tests of various hypotheses about subjects' responses. Each line shows the result of a Neyman-Pearson F-test of a restricted hypothesis on the mean estimate of subjects given a sequence of ring draws, $\mathrm{E}[\hat{p}|s]$, against an unrestricted hypothesis with more parameters. A p-value of 0.0 indicates that the value is smaller than computer precision (about 1e-323). The restricted hypotheses at lines 1 to 3 assume that the mean estimate is a function of the optimal estimate, $p^*(s)$; each hypothesis posits a different function. $\mathrm{Lo}(\cdot)$ is the log-odds function (Eq. 1.2). Lines 4 and 5 test the hypotheses that the mean estimate is an arbitrary function of the optimal estimate (a non-decreasing function, in line 4), against the unrestricted hypothesis that it also depends on the sequence length, $n(s)$. (In line 4, the arrow '$\nearrow$' means, in the case of the restricted hypothesis, that $m(\cdot)$ is required to be a non-decreasing function, and in the case of the unrestricted version, that $m(\cdot, n)$ is required to be non-decreasing for each n.) Line 6 tests the hypothesis that subjects perform Bayesian inference but use an incorrect prior (a symmetric Beta distribution with parameter α). Note that the hypothesis that subjects misweight the likelihood in their application of Bayes' rule is equivalent to the incorrect-prior hypothesis tested in line 6 (see Eq. 4.5). Line 7 tests the hypothesis that subjects have an incorrect prior and misweight the prior and the likelihood in their application of Bayes' rule. In all cases the restricted hypothesis is rejected at the .01 level for at least 81% of subjects, and in most cases for 95% or more.

1.2.2 Under-reaction to small samples ("conservatism"), but over-reaction to larger samples

As noted above, for a given number of draws, the average estimate of subjects increases with the optimal estimate (Fig. 4A). The difference between the estimate of subjects and the optimal estimate, i.e., the bias, varies across different numbers of draws (Fig. 4B). Specifically, upon observation of the first ring draw, the optimal estimate is .67 if the ring is red (and .33 if it is green), but the average estimate of subjects is .58 (or .42 for a green ring), a statistically significant difference (see Table 7 in Methods). Consistently, after two and three ring draws, the subjects are biased towards .5, i.e., they underestimate the proportion of red rings when the optimal estimate is greater than .5, and they overestimate it when the optimal estimate is less than .5. In other words, after observing no more than three ring draws, subjects' estimates exhibit "conservatism," as in many previous studies (Peterson *et al.*, 1965; Phillips and Edwards, 1966; Benjamin, 2019).

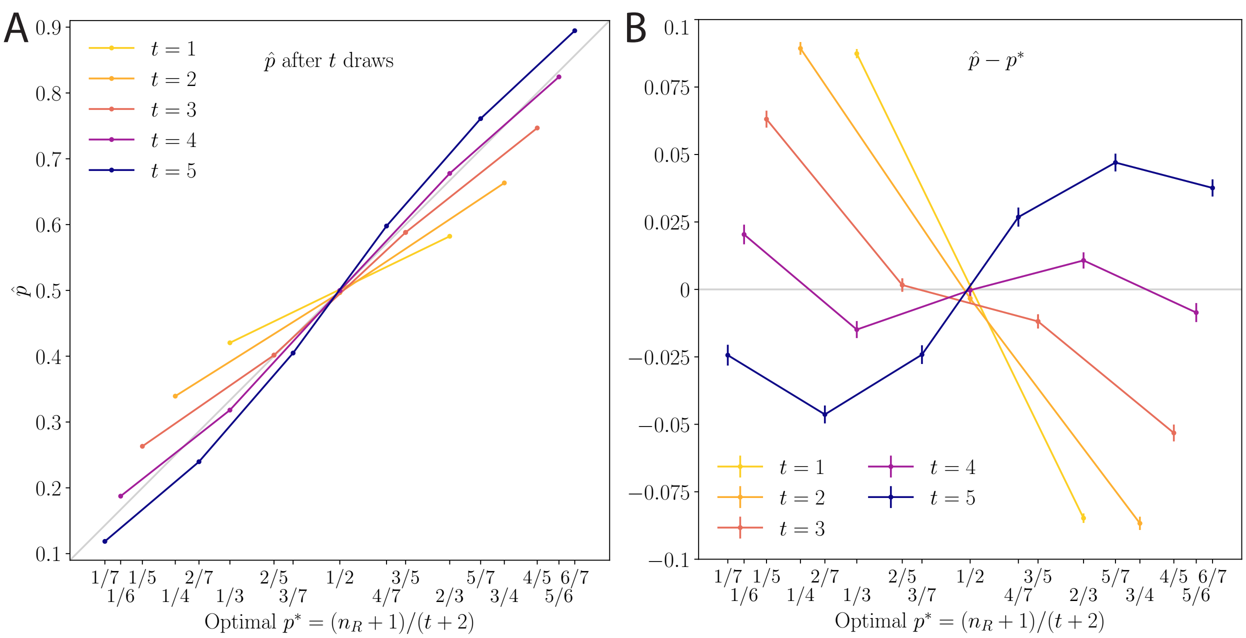

Figure 4: Bias reversal: subjects exhibit conservatism after the first three samples, and the opposite at the fifth sample. (A) Subjects' response, $\hat{p}$, as a function of the optimal estimate, p^*, after the presentation of t rings, with t from 1 (yellow) to 5 (blue). (B) Bias, i.e., difference between the subjects' response, $\hat{p}$, and the optimal estimate, p^*, after the presentation of t rings, as a function of the optimal estimate. The error bars indicate the standard error of the mean.

We note that the negative biases in subjects' responses, for optimal estimates greater than .5 (and the positive biases for optimal estimates less than .5), are smaller after the third draw than after the first draw. From this trend one might assume that the responses of subjects converge towards optimality as additional rings are drawn. However, upon observing the fourth ring, if the four rings observed are red, the subjects, here also, underestimate the proportion of red rings, but if three rings are red and one is green, they overestimate the proportion, a behavioral pattern at odds with the predictions of conservatism (as in both cases the optimal estimate is greater than .5). Moreover, the conservatism pattern is entirely reversed upon observation of the fifth ring: if three or more rings are red (and thus the optimal estimate is greater than .5), the subjects exhibit a sizable, positive bias, i.e., their responses are closer to the extreme value 1 than are the optimal estimates. Symmetrically, if two rings or less are red, the bias is negative, and the responses of subjects are closer to 0. In other words, after five rings the subjects are biased away from .5. In summary, subjects under-react to the evidence after one, two and three ring draws, and over-react to the evidence after five ring draws (we clarify that this is an overreaction to the cumulative evidence from the five observations, and not just to the last observation). Figure 4B illustrates the reversal of the bias as the number of draws increases from one to five.

1.3 Bayesian updating from an incorrect prior

A dynamic pattern of bias of the kind documented here cannot be explained by a model that assumes that the subject's response is based on the correct Bayesian posterior mean, but that this correct response is subjected to a (possibly nonlinear) transformation and that random response error is added. Here we examine the hypothesis that subjects behave like Bayesian observers, but misinterpret the information about the proportion of red rings that is available to them before any ring is drawn (specifically, that "all proportions of red rings, from 0% to 100%, are equally likely"). As seen in Eq. 1.1, the optimal estimate is derived from the number of red rings drawn and the length of the sequence presented. How the Bayesian observer uses these two inputs to determine the best response depends, in addition, on the prior. The optimal estimate is obtained using the correct, uniform prior, but a possibility is that the subjects make Bayesian inferences on the basis of an incorrect prior. Under this assumption, a subject's response is a different function of the number of red rings and of the length of the sequence than the optimal estimate. Here, we consider the case of a symmetric Beta-distributed prior, with parameter $\alpha > 0$. Symmetric Beta distributions spans a diverse range of distribution shapes, from distributions concentrated around the center of the $[0, 1]$ segment, to distributions that allocate a high probability to values close to the extremes, 0 and 1. Moreover, the symmetry of the problem, in our task, warrants the use of a symmetric prior. Finally, this family includes the correct prior, which corresponds to the case $\alpha = 1$.

The optimal estimate, for a Bayesian observer equipped with a symmetric Beta prior with parameter α, is

$$p_\alpha = \frac{n_R + \alpha}{t + 2\alpha}, \tag{1.4}$$

where n_R is the number of red rings drawn, and t the length of the sequence. Using Eq. 1.1, this can be reformulated as

$$p_\alpha = p^* + (2p^* - 1)\frac{1 - \alpha}{t + 2\alpha}, \tag{1.5}$$

where p^* is the truly optimal estimate. Hence, in contrast with the hypotheses examined above, the responses, here, and thus the bias over the course of the experiment, are predicted to vary as a function of both the optimal estimate and the length of the sequence. However, it follows from Eq. 1.5 that a subject, in this erroneous-prior hypothesis, either overestimates large proportions (and underestimates small proportions), if $\alpha < 1$, or the converse, if $\alpha > 1$, but it cannot exhibit both patterns, in contrast with the biases observed in our subjects' responses. Accordingly, the hypothesis that the estimates of a subject are, on average, equal to the estimates of a Bayesian observer equipped with a symmetric Beta prior is strongly rejected for all our subjects (Table 1, line 6), confirming that this hypothesis does not offer a satisfying account of subjects' responses (we present additional evidence against this hypothesis in section 3.2 below.)

1.4 Bayesian inference with incorrect beliefs about the data-generating process

So far we have considered several hypotheses about subjects' average responses: that they are equal to the estimates of the optimal, Bayesian observer; that they are captured by a transformation of these optimal estimates; and that they result from Bayesian inference with an incorrect prior. We have rejected these hypotheses, on the basis of their failure to reproduce the qualitative patterns we find in subjects' responses. But there are other ways in which subjects may have incorrect beliefs about the underlying process generating the observations: for instance, they may believe that the successive samples from a given box are *not* independent and identically distributed. Inspired by the literature, we examine several such hypotheses. First, we follow Yu and Cohen (2008), who show that 'sequential effects' in some behavioral tasks are well accounted for by the assumption that subjects have a prior belief in the non-stationarity of the probability underlying the observations. We thus assume, as in their models, that subjects are Bayesian, but believe that the proportion of red rings is subject to random changes at unannounced times (although the instructions explicitly mention that within each block of five trials the proportion does not change). We also examine a different hypothesis inspired by Meyniel, Maheu, and Dehaene (2016), whereby subjects believe that there is a sequential dependency in the ring draws, i.e., that the probability of a red ring depends on whether the preceding ring was red, or green. Formally, subjects are assumed to be inferring — in a Bayesian fashion — the conditional probabilities (or 'transition probabilities') of the red and green rings.

We consider in addition the possibility that subjects start each block of trial with a different prior on the proportion of red rings. This would occur, for instance, if they were uncertain about the prior, but gradually learned it, from one block to the next. Thus we consider a hierarchical Bayesian model, in which subjects learn the prior throughout the experiment. We also consider the hypothesis that the prior is randomly chosen at the beginning of each block (independently from the preceding block). And finally, we compare the behavior of subjects when the preceding block contains a majority of red rings, and when it contains a majority of green rings. We note that apart from this last analysis, the hypotheses we examine are ones in which the subjects are assumed to be conducting sound Bayesian inference, but on the basis of incorrect beliefs.

Through a careful qualitative and quantitative examination of the behaviors predicted by these hypotheses, the statistical testing of their various implications, the analysis of the performance, in comparison to our other models, of the models derived from these hypotheses, and the inspection of model simulations, we confidently reject all the hypotheses listed in this section: they do not provide a satisfying account of the behavioral data. In the interest of space, here we have only presented these hypotheses; the interested reader will find the details of how we test their implications in Methods. Below, we continue our examination of subjects' responses, and exhibit how they seem to derive from imprecise retrieval of the information presented, in contrast with the models considered so far.

2 Evidence of Judgments Based on an Imprecise Cognitive State

We have seen that subjects' judgments cannot be understood as simply a noisy report of, or some nonlinear transformation of, the optimal response on that trial, given all of the evidence that has been to that point. We next argue that there is reason to conclude that subjects' judgments do not make use of a precise record of that evidence while nonetheless computing a suboptimal response; instead, they appear to make use only of an imprecise record of the evidence that has been presented. This suggests that subjects' errors should be attributed not so much to a mistaken understanding of how to use the information available to them, as to cognitive processing that economizes on the amount of information that must be retained and retrieved while the subjects perform the task.

The information that is used in producing subjects' judgments seems to be less precise than a perfect record of the evidence presented, in at least two respects. First, subjects' judgments seem to be based on a noisy record of past evidence, rather than a perfectly precise one. We conclude this because their judgments involve random noise, that cannot be attributed purely to noise in the way the response is generated on the basis of a deterministically evolving cognitive state; instead, there appears to be random noise in the available record of past observations, evident from the way that the noise seems to propagate from one trial to the next. And second, whereas a perfect record of the evidence presented would require a subject's belief state to vary along two dimensions, subjects' responses appear to reflect the evolution of an internal cognitive state that moves randomly among a set of states that can be ordered on a line. Thus even apart from the random noise in the cognitive state, it appears to involve a compressed representation that is not informative enough to make optimal responses possible.

2.1 Noise in cognitive processing propagates to successive trials

We turn first to the evidence for noise in the record of past observations that is available to the decision process. It is evident that our subjects' responses involve noise; since the same finite number of possible evidentiary states (i.e., finite sequences of ring draws) are visited multiple times by each subject, we can see that a given subject does not always give the same response in response to given evidence. Yet this noise might be viewed as arising in the process by which subjects' responses are generated, on the basis of an internal state that includes a perfect record of the presented evidence. For example, one might assume that instead of maximizing the expected reward from their response, subjects respond probabilistically according to a "softmax" operator, but that the relative probabilities of different responses are determined by the (correctly calculated) expected rewards. (In the experimental game theory literature, the "quantal response equilibrium" solution concept

(Palfrey, 1995) assumes behavior of this kind.) Alternatively, Khaw *et al.* (2017a) propose a model of subject behavior in a more complex version of our task, in which the subject's response on each trial is based on a noisy, imprecise readout of the history of ring draws; but although this readout, when producing a response, is imprecise, the model assumes that a perfect record of the complete the history of ring draws is available to the decision-maker at all times. Such a model predicts that responses are random (conditional on the history of ring draws), but the randomness in the response given on some trial does not in any way corrupt the accuracy of the record that is available to be (imperfectly) accessed on subsequent trials. Here we show that no model of this kind can account for the nature of the randomness in subjects' responses.

2.1.1 Autocorrelation of subjects' responses

There is an important testable prediction of any model in the above class. Under the assumption that the subject's cognitive state evolves as a deterministic function of the history of evidence (whether in fact it represents a complete record of that evidence), with noise arising only in the way that the response is generated on the basis of this cognitive state, one obtains the prediction that, conditional on the sequence of ring draws, responses are independent events. We propose to test this prediction in the case of our subjects' responses.

Let the 'excursion' of a given subject at a given trial be the difference between the subject's response provided upon the observation of this trial's sequence of rings, and the average of the subject's responses provided at all the trials that feature the same sequence. An excursion thus measures how far a subject's response is from her average response when presented with identical sequences. Under the assumptions of deterministic cognitive states and noise in response selection, an excursion at trial $t+1$ is a random variable independent from the excursion at the preceding trial, t. We find, however, that the excursion at a given trial is highly correlated with the excursion at the preceding trial (Pearson correlation coefficient: 0.62; p-value below computer precision. Individual coefficients: mean: 0.62, standard deviation: 0.14; all p-values $< .01$, median p-value: 4.2e-95). In other words, if on a given trial a subject provides a larger response than she does on average in trials that feature the same sequence of rings, then in the subsequent trial she also provides a response larger than average (Fig. 5A). This suggests that independent noise in response selection does not alone account for the variability in subjects' responses.

A possible explanation of the correlation between subjects' successive excursions is that each *block* of five successive trials is characterized by an excursion from the average that uniformly impacts the responses at all the trials in the block. To illustrate this possibility, we consider a subject whose response at a given trial, in a given block, is the sum of the average response to sequences identical to that observed at this trial, and of a block-dependent excursion, which is a centered random variable, sampled once at the beginning of the block (it represents, for instance, the variability of the subject's cognitive state at the beginning of the block). For this subject, two successive responses are correlated — not because of a mechanistic relation between the two, but because both are influenced by the common, initial excursion. Moreover, the correlation coefficient between responses at more distant trials, e.g., at trial t and at a preceding trial $t-l$, is also positive, and it does not depend on the distance between the trials, l, as all responses in a block are equally impacted by

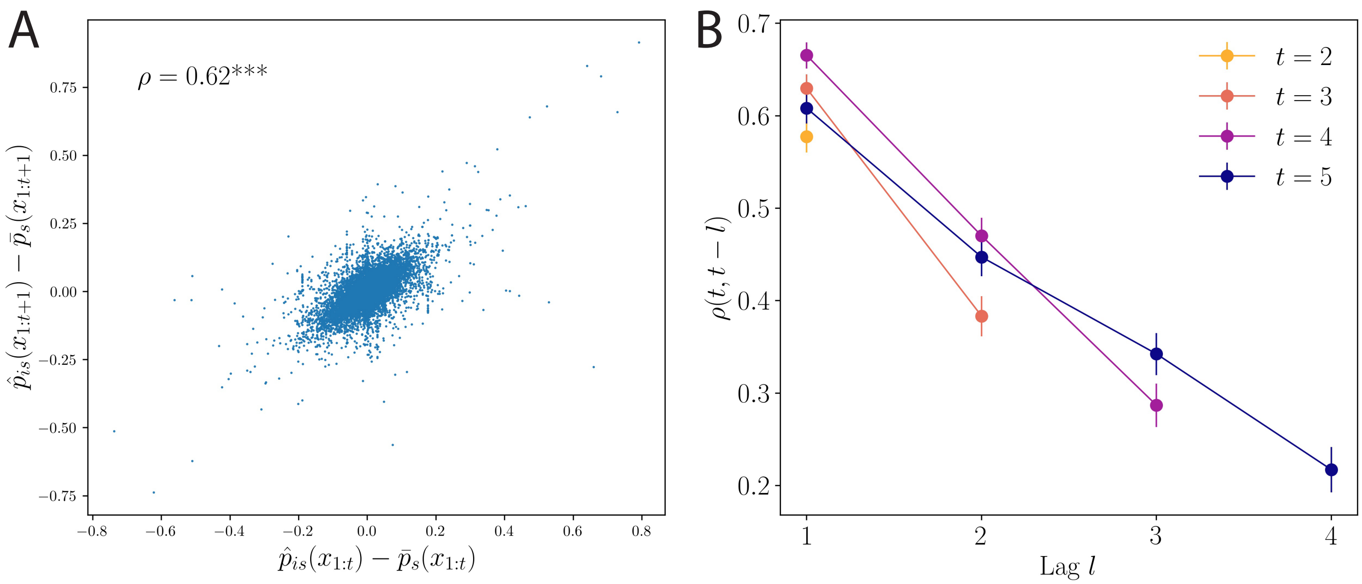

Figure 5: Autocorrelation in subjects' responses. (A) Excursions in responses (i.e., distances from the average response,) at a given trial vs. excursions in responses at the preceding trial. Specifically, let $\hat{p}_{is}(x_{1:t})$ be the estimate given by subject s after a sequence of t draws, $x_{1:t}$, in the ith block of trials, and $\bar{p}_s(x_{1:t})$ the mean estimate of subject s averaged over the blocks of trials featuring the same sequence, $x_{1:t}$. The excursion of subject s in trial t of block i is the difference $\hat{p}_{is}(x_{1:t}) - \bar{p}_s(x_{1:t})$. Panel A shows the positive correlation between the excursions of subjects at trials $t+1$ and the excursions at trial t. In other words, a larger-than-average response at a given trial is predictive of a larger-than-average response at the subsequent trial. (B) Coefficients of correlation between subjects' excursions at trial t and subjects' excursions at trial $t-l$, with lag l ranging from 1 to $t-1$. Error bars indicate the 90% confidence interval.

the same excursion (we also consider a more elaborate model in which the initial prior in each block is subject to a random shock. The behavior of this model is similar to the one just described; see Methods). Turning to the responses of subjects, we find that there is indeed a positive autocorrelation in their responses at distant trials. However, in contrast with the prediction just presented, the autocorrelation decreases, approximately linearly, as a function of the distance between two trials. The Pearson correlation coefficient between the most distant trials, the first and last ones, is 0.22, i.e., roughly three times less than that of successive trials (Fig. 5B). An hypothesis consistent with these observations is that a subject's response is derived from an internal cognitive state which is not a deterministic function of the history of ring draws, but rather represents a noisy record of the presented evidence, and that noise propagates through successive cognitive states. This hypothesis implies that a response at a given trial should provide information about the response at the next trial — although one might expect a priori that the latter response should only depend on the observed sequence of ring, while the former response should be irrelevant. Hence we further investigate, below, the extent to which a response at some trial contains information about the response at the following trial.

2.1.2 The information content of the preceding response

The results just presented indicate that the response of a subject in a given trial is predictive, to some extent, of the responses in the following trials. This suggests a different account of subjects' responses than the ones investigated thus far. Under this new account, the average subjects' response $\hat{p}_{t+1}$ at trial $t+1$ can be derived from the preceding response, $\hat{p}_t$, and from the new drawn ring, x_{t+1}, rather than from the sequence of rings presented, $x_{1:t+1}$. To compare these two accounts, we seek to obtain quantitative measures of the respective shares, in the response of a subject, that can be derived from the preceding response, $\hat{p}_t$, and from the sequence presented before the new ring, $x_{1:t}$. Information theory provides measures suitable for this purpose. Specifically, we decompose, first, the entropy of a response at trial $t+1$ as

$$H(\hat{p}_{t+1}) = I(\hat{p}_{t+1}; x_{t+1}) + I(\hat{p}_{t+1}; \hat{p}_t, x_{1:t}|x_{t+1}) + H(\hat{p}_{t+1}|\hat{p}_t, x_{1:t+1}). \tag{2.1}$$

The first term of the right-hand side of Eq. 2.1 is the mutual information (MI) between the response at trial $t+1$, $\hat{p}_{t+1}$, and the last ring draw, x_{t+1}. It is a measure of the amount of information about the subject's response $\hat{p}_{t+1}$ that is obtained by learning the last ring that the subject observed, x_{t+1}. The second term is the MI between $\hat{p}_{t+1}$ and the pair composed by the two variables we are interested in — the preceding response, $\hat{p}_t$, and the sequence observed, $x_{1:t}$ — conditional on the last ring, x_{t+1}. It measures the amount of information about $\hat{p}_{t+1}$ that is obtained by learning $\hat{p}_t$ and $x_{1:t}$, once the knowledge of x_{t+1} has been taken into account (below, we further decompose this term in order to examine the information content of each variable, $\hat{p}_t$ and $x_{1:t}$). Finally, the third term is the 'residual' entropy, i.e., the information that is left in $\hat{p}_{t+1}$ once $\hat{p}_t$, $x_{1:t}$, and x_{t+1} have been taken into account.

We estimate each of these quantities in behavioral data. The 'naïve' approach to deriving the entropy from a dataset, which consists in directly using the empirical frequencies in lieu of the probabilities in the expression of the entropy, is known to suffer from a strong bias. Hence we use, to estimate these quantities, the 'Best Upper Bounds' estimator, which minimizes a polynomial approximation of the bias of the naïve estimator (Paninski, 2003). We also computed these quantities with a different method, which relies on Bayesian estimates (Wolpert and Wolf, 1995; Nemenman *et al.*, 2002), and obtain similar results (see Table 8 in Methods.)

Panel A

	Value	Share
$H(\hat{p}_{t+1})$	6.39	100%
$= I(\hat{p}_{t+1}; x_{t+1})$	0.25	3.96%
$+I(\hat{p}_{t+1}; \hat{p}_t, x_{1:t}\|x_{t+1})$	4.34	67.9%
$+H(\hat{p}_{t+1}\|\hat{p}_t, x_{1:t+1})$	1.80	28.3%

Panel B

	Value	Share
$I(\hat{p}_{t+1}; \hat{p}_t, x_{1:t}\|x_{t+1})$	4.34	100%
$= I(\hat{p}_{t+1}; x_{1:t}\|x_{t+1})$	0.92	21.1%
$+I(\hat{p}_{t+1}; \hat{p}_t\|x_{1:t+1})$	3.42	78.9%

Table 2: A larger share of a response's information stems from the preceding response than from the evidence presented. Breakdown of the entropy of a response, following Eq. 2.1 (Panel A) and Eq. 2.2 (Panel B).

We find that the last ring presented, x_{t+1}, represents just under 4% of the entropy in subjects' responses, while the residual entropy, not accounted for by the preceding response nor by the sequence of ring draws, represents 28.3% of the entropy in responses (Table 2A). In order to examine the information contents of the sequence of rings up to trial t, $x_{1:t}$, and of the preceding response, $\hat{p}_t$, we now turn to the second term of Eq. 2.1, which we further decompose as the sum of the MI of $\hat{p}_{t+1}$ and $x_{1:t}$, conditional on x_{t+1}, and the MI of $\hat{p}_{t+1}$ and $\hat{p}_t$, conditional on the whole sequence $x_{1:t+1}$, i.e,

$$I(\hat{p}_{t+1}; \hat{p}_t, x_{1:t}|x_{t+1}) = I(\hat{p}_{t+1}; x_{1:t}|x_{t+1}) + I(\hat{p}_{t+1}; \hat{p}_t|x_{1:t+1}). \tag{2.2}$$

For any observer whose response at trial $t+1$ is a deterministic function of the sequence of rings observed, $x_{1:t+1}$, the information in the response, $\hat{p}_{t+1}$, is entirely contained in the sequence up to trial t, $x_{1:t}$, once the last ring, x_{t+1}, is known; and once the whole sequence $x_{1:t+1}$ is known, no additional information resides in the preceding response, $\hat{p}_t$, i.e., $I(\hat{p}_{t+1}; \hat{p}_t|x_{1:t+1}) = 0$. In other words the first term in the right-hand side of Eq. 2.2 represents 100% of the left-hand side, and the second term represents 0%. This is the case, for instance, of the optimal observer. By contrast, for the subjects, these shares are 21.1% and 78.9%, respectively (instead of 100% and 0%; Table 2B). The share of the preceding response, $\hat{p}_t$, once the whole sequence is known, is thus more than 3.7 times the share of the observed sequence, $x_{1:t}$, once x_{t+1} is known, while it would be zero if the response was a function of the sequence. In other words, only a small fraction of the information content of a response is contained in the sequence of ring draws, while a larger fraction is instead contained in the preceding response. Thus subjects' responses seem to derive from representations of the decision situations — cognitive states — that are not perfect records of the sequences presented. Cognitive states, rather, seem to be imprecise records of the history of ring draws, with noise propagating through successive states as new evidence is presented. Below, we examine the kind of information that is carried by cognitive states, and in particular whether they seem to comprise the summary statistics that are used by the optimal observer to modulate its responses as a function of the presented evidence.

2.2 The cognitive state appears to be one-dimensional

We have shown that subjects' responses are based on a cognitive state that provides only a noisy record of previous experience, with noise that propagates from one trial to the next. But in addition to this, the set of possible states between which stochastic transitions occur in response to additional observations appears to be insufficiently differentiated to allow the range of different responses that would be required for optimal Bayesian estimates. We have shown above (Equation 1.1) that optimal responses depend on two pieces of information: both n_R and n_G, or alternatively, both the net evidence in favor of red, $n_R - n_G$, and the total number of observations, t. Indeed, although one might think that the quantity p^* defined in Equation 1.1 suffices as a summary statistic, the current value of p^* suffices to determine the current optimal slider position but it does not provide the information needed in order to know how to adjust the value of p^* after another ring draw is observed; that would depend both on the old value of p^* and the value of t. For instance, the optimal adjustment to bring to the current optimal response after a red ring is drawn is

$$p_R^* - p^* \;=\; \frac{1}{t+3}\,(1-p^*), \tag{2.3}$$

where p_R^* is the new optimal response. A two-dimensional summary of observed evidence (such as any invertible transformation of the pair (n_R, n_G),) is necessary to determine both the optimal response on the current trial and the way in which the summary description should be updated in response to further observations. Thus optimal decision making requires access to a cognitive state that varies over two dimensions as evidence accumulates, reflecting changes both in the degree to which the evidence supports a higher posterior mean estimate of p and in the degree of certainty that the evidence provides about the value of p.

Instead, our evidence suggests that the cognitive state moves along a line, or more generally a one-dimensional manifold, implying that both the degree to which experience favors a higher estimate and the degree of certainty are not independently represented. We maintain the assumption that a subject's response derives from an internal cognitive state, r, which we assume to be a mental representation of the decision situation (characterized by the sequence of rings observed), a point in a space of possibly high dimension (e.g., a vector representing the activity of every neuron in the brain). The observation of a new ring, x, results in a new cognitive state which depends deterministically or stochastically on the preceding cognitive state, and on the color of the new ring. We do not have access to the cognitive states, but we do observe the responses, $\hat{p}$, which we assume to depend, here also deterministically or stochastically, on the cognitive states. Therefore, the mean responses, at a trial t and at subsequent trials, are functions of the cognitive state at trial t. As functions from the (possibly high-dimensional) space of cognitive states to the [0,1] response scale, these mean responses act as "projections" that can indirectly inform us on the topology of the cognitive states. In particular, if the manifold of cognitive states occurring in our experiment has dimension n, its projection on the space of average responses cannot have a dimension higher than n. If instead the projections of the cognitive states have dimension n' smaller than the space into which they are projected, it suggests that the dimension of the cognitive states may be as small as n'.

For instance, for all sequences of t draws, $x_{1:t}$, we compute the average of the obtained responses, $\langle \hat{p}_t | x_{1:t} \rangle$, and for each of these sequences followed by a red ring, $(x_{1:t}, x_{t+1} = R)$, we compute the average adjustment, $\langle \hat{p}_{t+1} - \hat{p}_t | x_{1:t}, x_{t+1} = R \rangle$. We can thus look at the adjustments as a function of the preceding responses, averaged for each sequence. For the optimal observer, as predicted by Eq. 2.3 above, the adjustments depend not only on the preceding response, but also on the number of pieces of evidence observed, t. For each t, the adjustment is a linear function of the preceding response, but the slope and intercept of this linear function depend on the number of rings drawn (Fig. 6, top left panel). By contrast, the adjustments of subjects seem to all fall along a single line, regardless of the amount of evidence accumulated (Fig. 6, top right panel). In other words, in the plane formed by these two quantities (the average responses and the average adjustments), the sets of points obtained for the optimal observer occupy a bidimensional area, while those obtained for the subjects are close to a unidimensional curve.

We can look at different planes, formed by different quantities: for instance, the average adjustments after observing a green ring, $\langle \hat{p}_{t+1} - \hat{p}_t | x_{1:t}, x_{t+1} = G \rangle$, paired with the average responses, $\langle \hat{p}_t | x_{1:t} \rangle$, and the average response after observing a red ring, $\langle \hat{p}_{t+1} | x_{1:t}, x_{t+1} = R \rangle$, paired with the average response after observing a green ring, $\langle \hat{p}_{t+1} | x_{1:t}, x_{t+1} = G \rangle$. In all cases, we find that the the optimal responses describe a bidimensional area, in which responses are differentiated, in particular, by the different numbers of samples presented in the corresponding sequences, t, while the responses of subjects appear to fall along a unidimensional curve, and seem not to depend on t (Fig. 6).

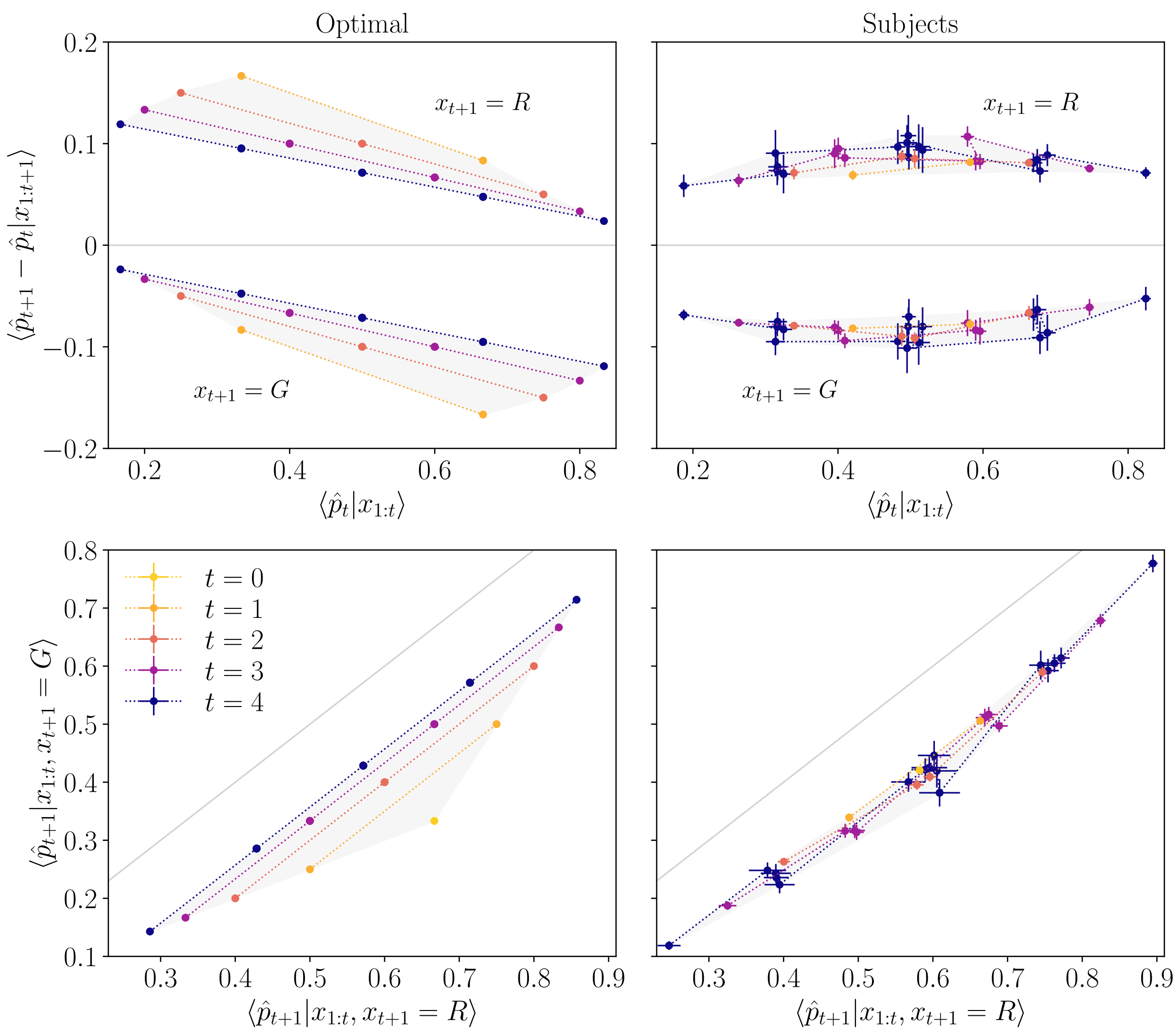

Figure 6: Subjects' responses seem to reflect unidimensional cognitive states, in contrast with those of the Bayesian observer. *Top row:* average updates, $\hat{p}_{t+1} - \hat{p}$, after observing the sequence $x_{1:t}$ and a red or a green ring ($x_{t+1} = R$ or G), vs. average responses after observing the same sequence ($x_{1:t}$). *Bottom row:* average responses after observing the sequence $x_{1:t}$ and a green ring ($x_{t+1} = G$), vs. average responses after observing the same sequence ($x_{1:t}$) and a red ring ($x_{t+1} = R$). The grey line is the identity line. *Left column:* Optimal responses. *Right column:* Subjects' responses.

We propose to measure quantitatively the extent to which the responses of subjects can be characterized by an underlying state that is unidimensional. We consider the responses as points in the three-dimensional space given by the average response at trial t and those at trial $t+1$ following a red or a green ring. We compute the line and the plane that minimize the distance to these points—a *linear* dimensionality-reduction procedure, known as Principal Component Analysis (PCA). We then compute the error of each response, as its distance from the corresponding average response projected on the PCA line or plane, and we examine the Fraction of Variance Unexplained (FVU), defined as the mean squared error divided by the variance of responses. With just one principal component, we find that the FVU is 17.73%. With two components, it is 17.62% (and with three components it is 17.61%). Thus the second PCA component increases the variance explained, but only by a modest amount (0.11 percentage point). The relative share explained by a single component could certainly be improved, at least marginally, if non-linear transformations were allowed; but this linear analysis suggests that the responses of subjects can be accounted for, to a large extent, by a unidimensional underlying state, which imperfectly reflects the sequence of observations.

3 Responses are Not Bayesian, Even Conditional on the Imprecise Cognitive State

We have argued that subjects' responses appear to be a function of a cognitive state which provides only imprecise information about the sequence of rings that has been observed. This raises the question whether the bias and variability of subjects' responses can be fully accounted for by the imprecision of this cognitive state. Are their responses perhaps optimal, conditional on the fact that they must be based on the imprecise cognitive state, as in Bayesian accounts of perceptual biases (e.g., Petzschner *et al.,* 2015; Wei and Stocker, 2015), "rational inattention" models of imperfect economic decisions (Caplin and Dean, 2015; Caplin *et al.,* 2020), or the model of imprecise probability estimation proposed by Khaw *et al.* (2017a)? In fact, there are testable predictions of this hypothesis that can be formulated independently of a specific mode of the evolution of the cognitive state, and that allow us to reject the hypothesis of Bayesian optimality of our subjects' responses.

3.1 Are subjects' probability estimates well-calibrated?

In our task, subjects would maximize the expected value of their reward, given that their response must be some function of the current cognitive state r, by reporting a probability estimate equal to

$$\hat{p}(r) \;=\; \mathrm{E}[p\,|r], \tag{3.1}$$

where the conditional expectation is computed using the joint distribution of p and r implied by the prior from which p is drawn (we assume in this section that subjects hold the correct prior), and by the stochastic dynamics of the cognitive state in response to independent, identically-distributed draws from the Bernoulli distribution with parameter p. But if this

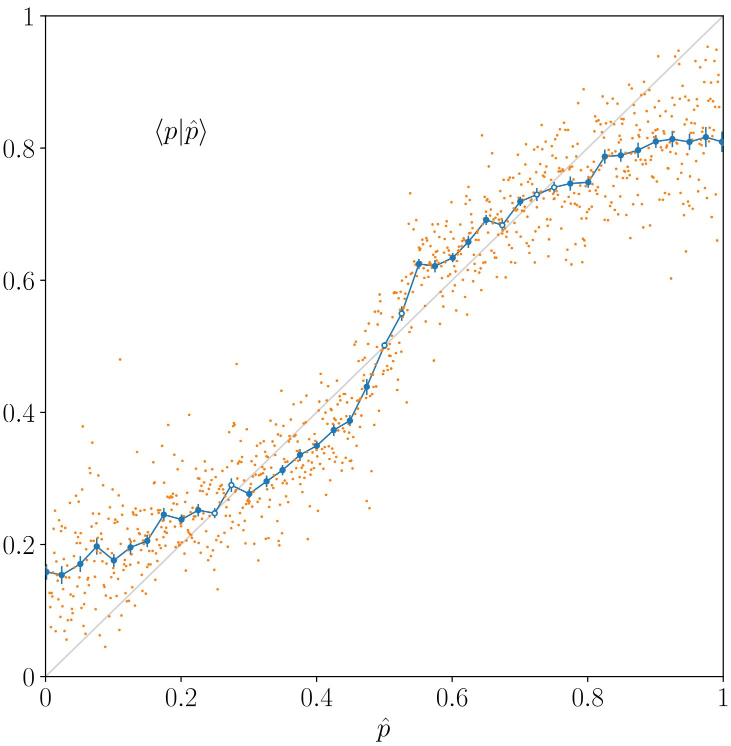

Figure 7: The calibrated-responses Bayesian property is not verified in subjects' responses. Empirical average of the true proportions p conditional on each response $\hat{p}$ (orange dots) and conditional on each .025-long response interval I_i (blue line). The lengths of the error bars equal twice the standard error of the mean. Filled blue points indicate the cases in which the t-tests of equality between the means of the true probabilities and of the responses are rejected at the .01 level (see Methods).

is true, the law of iterated expectations implies that we should observe that

$$\mathrm{E}\left[p|\hat{p}\right] \;=\; \mathrm{E}\left[\mathrm{E}\left[p|r,\hat{p}\right]\,|\hat{p}\right] \;= \mathrm{E}\left[\mathrm{E}\left[p|r\right]\,|\hat{p}\right] \;=\; \hat{p}, \tag{3.2}$$

for each slider setting $\hat{p}$ that is ever observed, where the conditional expectation on the left is computed using the joint distribution of p and $\hat{p}$, and the middle equality follows from $\hat{p}$ being a deterministic function of r (Eq. 3.1). In other words, subjects' reported probability estimates should be "well-calibrated" in the terminology of Dawid (1982).

This prediction is a property of the joint distribution of p and $\hat{p}$ that would have to hold, regardless of what we assume about the structure of the cognitive state r and its stochastic dynamics. Hence it is straightforward for us to test it, subject of course to the caveat that we have only a finite sample from the joint distribution of p and $\hat{p}$ implied by our subjects' decision process.

An examination, in subjects' data, of the empirical average of the true proportion p for each response $\hat{p}$ suggests that Eq. 3.2 is not everywhere verified. Instead, when the subject's response is above .75 (respectively, below .25), the average true proportion is below the subject's response (respectively, above), a behavior that has been called 'over-confidence' in the literature (Erev *et al.,* 1994). Closer to the middle value, .5, the subjects exhibit the opposite behavior, i.e., the average true proportions are further from .5 than are the responses of the subjects (Fig. 7).

Before proceeding to a statistical test of the hypothesis defined by Eq. 3.2, we note that many of the possible responses $\hat{p}$ are chosen once, or a small number of times (on average, 51% of the responses of a subject are chosen only once by this subject, over the course of the experiment). Thus an empirical estimation of the average of the true proportions conditional on a response (the left-hand side of Eq. 3.2) would be established on the basis of a very small sample. Hence instead of considering separately each possible response $\hat{p}$ we divide the set of possible responses into disjoints intervals I_i of length .025 (except at each end of the response scale, where the interval is half this length in order to obtain a middle interval that contains the response .5 at its center; see Methods). We test the prediction, implied by Eq. 3.2, that for each interval I_i the mean true proportion conditional on I_i is equal to the mean response conditional on I_i, i.e.,

$$\mathrm{E}[p|\hat{p} \in I_i] = \mathrm{E}[\hat{p}|\hat{p} \in I_i]. \tag{3.3}$$

We test the restricted hypothesis defined by Eq. 3.3 against the unrestricted hypothesis that allows the mean true probability to be any function, m, of the interval I_i, i.e., $\mathrm{E}[p|\hat{p} \in I_i] = m(I_i)$. The restricted hypothesis is strongly rejected when pooling together the responses of the subjects (p-value: 8.4e-219). As for the individual tests, the restricted hypothesis is rejected at the .05 level for 96% of subjects, and at the .01 level for 76% of subjects (median p-value: 3.4e-10). We conclude that the responses of most subjects do not satisfy the calibrated-responses Bayesian property. (We conducted additional tests of Eq. 3.3, including ANOVA F-tests and a series of t-tests, and reached the same conclusion; see Methods.)

3.2 Are the changes in subjects' estimates consistent with Bayesian updating?

A limitation of the test of the Bayesian optimality of subjects' responses considered in the previous section is that it tests (and rejects) a property of estimates that would be optimal under the correct prior distribution for the true value of p, but that would not have to hold, even if subjects' estimates are consistent with Eq. 3.1, if the prior distribution used to compute the conditional expectation in Eq. 3.1 is some other distribution. One might suppose instead that subjects' responses are optimal Bayesian estimates, conditional on a noisy internal representation of the evidence presented to them, but optimized for a prior regarding the true value of p that subjects bring to the experiment from their prior experience with probabilities, as proposed by Zhu *et al.* (2020).[2]

We have already tested (and presented evidence against) the hypothesis that our subjects form correct Bayesian estimates, starting from an incorrect prior but conditioning on a

[2] We tell our subjects that p is drawn from a uniform distribution. But Bayesian models of perception should not be understood as assuming that people consciously apply Bayes' rule in order to form the judgments that they express, and the implicit prior for which perceptual processing has been optimized, under a hypothesis of this kind, is not generally supposed to be modifiable by information about stated probabilities supplied in an experiment rather than learned from experience (Feldman, 2015).

perfect record of the ring draws that they have seen. We wish now to consider the more general hypothesis of Bayesian inference, starting from a possibly incorrect prior *and* not necessarily based on a perfect record of the ring draws. Even under these weaker assumptions, Bayesian inference requires a certain consistency between a subject's estimate after a certain amount of evidence and the estimates that they will make after another ring draw (depending which color of ring is drawn). We can seek to determine whether successive estimates of our subjects are consistent in this way, regardless of the objective correctness of the probability beliefs that they hold before the additional ring is drawn.

Let r_t be the cognitive state of a subject at a given trial t, $f(p|r_t)$ the Bayesian belief distribution it implies (possibly derived from an incorrect prior), and let $\hat{p}(r_t)$ be the subject's estimate, derived from the cognitive state r_t. In addition to allowing for a subjective prior, we do not necessarily wish to assume that the observer's subjective belief just before observing a new ring, x_{t+1}, is necessarily the same as the one upon which the estimate $\hat{p}(r_t)$ is based; that is, we do not wish to assume that the cognitive state r_{t+1} includes a perfect record of the previous cognitive state r_t, given the evidence above suggesting that the cognitive state is merely one-dimensional. Thus we assume that r_{t+1} consists of an imprecise record $\tilde{r}_{t+1}$ of the previous cognitive state, together with the new observation, x_{t+1}, but we do not assume that r_t can be precisely reconstructed from knowledge of $\tilde{r}_{t+1}$. (For example, both $\tilde{r}_t$ and $\tilde{r}_{t+1}$ may be one-dimensional, while $r_t = (\tilde{r}_t, x_t)$ is two-dimensional.) We nonetheless assume that $\tilde{r}_{t+1}$ contains sufficient information about the previous cognitive state to allow the value of the posterior mean conditional on r_t to be inferred from $\tilde{r}_{t+1}$; specifically, $\mathrm{E}[p|\tilde{r}_{t+1}] = \mathrm{E}[p|r_t]$. Note that this allows $\tilde{r}_{t+1}$ to be one-dimensional (a record of the previous posterior mean). Moreover, under the hypothesis that the subject's response on each trial is her subjective posterior mean, i.e., $\hat{p}(r_t) = \mathrm{E}[p|r_t]$, then it makes sense that it should still be possible to condition upon the value of the previous posterior mean when choosing $\tilde{r}_{t+1}$, since the previous response $\hat{p}(r_t)$ will still be visible on the screen under our experimental interface.

Let $f(p|\tilde{r}_{t+1})$ be the subjective density over p conditional on an imperfect record $\tilde{r}_{t+1}$ before observing the next ring draw, and let $f^x(p|\tilde{r}_{t+1})$ be the subjective posterior density if a ring draw $x \in \{R, G\}$ is observed along with the imperfect record $\tilde{r}_{t+1}$. Correct Bayesian updating further requires that

$$f(p|\tilde{r}_{t+1}) = P(R|\tilde{r}_{t+1})f^R(p|\tilde{r}_{t+1}) + P(G|\tilde{r}_{t+1})f^G(p|\tilde{r}_{t+1}), \tag{3.4}$$

where $P(x|\tilde{r}_{t+1})$ is the subjective probability of drawing the ring x, under the belief implied by the imprecise record, $\tilde{r}_{t+1}$. In other words, Bayesian updating requires that the 'prior', $f(p|\tilde{r}_{t+1})$, be the 'expectation of the posterior', where the expectation is taken over the possible observations, x. Furthermore, given an imprecise record, $\tilde{r}_{t+1}$, the subjective probability of observing a red ring is the mean of the posterior over p implied by $\tilde{r}_{t+1}$, i.e., $P(R|\tilde{r}_{t+1}) = \int pf(p|\tilde{r}_{t+1})dp \equiv \mathrm{E}[p|\tilde{r}_{t+1}]$, a quantity that we have assumed to be equal to $\mathrm{E}[p|r_t] = \hat{p}(r_t)$; similarly, the subjective probability of observing a green ring is $1 - \hat{p}(r_t)$. If the response after observing the additional ring draw x, which we denote by $\hat{p}^x(\tilde{r}_{t+1})$, is here also the mean of the subjective posterior (i.e., $\hat{p}^x(\tilde{r}_{t+1}) = \int pf^x(p|\tilde{r}_{t+1})dp$), then from the equation above we derive the relation

$$\hat{p}(r_t) = \hat{p}(r_t)\hat{p}^R(\tilde{r}_{t+1}) + (1 - \hat{p}(r_t))\hat{p}^G(\tilde{r}_{t+1}), \tag{3.5}$$

i.e., the response of a Bayesian observer at a given trial is the expectation of the response at the next trial, where the expectation is taken over the possible outcomes. An equivalent view on this property is obtained by rearranging the terms of the equation, as

$$\hat{p}(r_t)\big(\hat{p}^R(\tilde{r}_{t+1}) - \hat{p}(r_t)\big) + (1 - \hat{p}(r_t))\big(\hat{p}^G(\tilde{r}_{t+1}) - \hat{p}(r_t)\big) = 0. \tag{3.6}$$

Bayesian inference thus prescribes, in our task, a relation between the adjustment of the response should the new ring be red, $\hat{p}^R(\tilde{r}_{t+1}) - \hat{p}(r_t)$, and the adjustment of the response should the new ring be green, $\hat{p}^G(\tilde{r}_{t+1}) - \hat{p}(r_t)$. If the estimate of the proportion of red rings, $\hat{p}(r_t)$, is large, then the adjustment should be small if a red ring is observed and it should be large if a green ring is observed; and vice-versa if $\hat{p}(r_t)$ is small instead.

We call Eq. 3.6 the 'consistent-updates' property, and we emphasize that it is implied by Bayes' rule and the structure of our task. We implicitly assume that the subjects understand this structure, and in particular that they believe the true probability p to be constant (in Methods we examine, and reject, alternative hypotheses). The only additional assumption is that the imprecise record of the previous cognitive state allows for a precise recall of the previous posterior mean, i.e., of the previous response—an assumption that we deem reasonable, as the previous response is visible on screen. Unfortunately, in all trials we observe either the response following a red ring, or the response following a green ring, but we do not observe both simultaneously; neither do we observe the cognitive state. In order to be able to investigate and test the consistent-updates property in subject's data, we consider the two subjective-probability-weighted revisions

$$\begin{aligned} \delta^R(r, \tilde{r}) &= \hat{p}(r)\big(\hat{p}^R(\tilde{r}) - \hat{p}(r)\big) \\ \text{and } \delta^G(r, \tilde{r}) &= -(1 - \hat{p}(r))\big(\hat{p}^G(\tilde{r}) - \hat{p}(r)\big). \end{aligned} \tag{3.7}$$

The consistent-updates property (Eq. 3.6) implies that these are equal, i.e.,

$$\delta^R(r, \tilde{r}) = \delta^G(r, \tilde{r}). \tag{3.8}$$

We take the expectation over r and $\tilde{r}$ of both sides of Eq. 3.8 conditional on the observed sequence s, and obtain two functions of s, $\mathrm{E}[\delta^R|s]$ and $\mathrm{E}[\delta^G|s]$, which are predicted to be equal, i.e.,

$$\mathrm{E}[\delta^R|s] = \mathrm{E}[\delta^G|s]. \tag{3.9}$$

We can estimate these quantities from a subject's overt behavior, and test that they are equal. (We note that although we can only observe δ^R (and δ^G) when the sample at the next trial is red (and green, respectively), both quantities are independent of the outcome at the next trial, thus their averages conditional on the outcome are equal to their unconditional averages.) Alternatively we can take the expectation of Eq. 3.8 conditional on another observable which we have shown to be a variable of interest: the preceding response, $\hat{p}$. For the reasons explained in the previous section, we consider instead the interval I_i to which belongs the preceding response, and obtain the prediction

$$\mathrm{E}[\delta^R|\hat{p} \in I_i] = \mathrm{E}[\delta^G|\hat{p} \in I_i]. \tag{3.10}$$

The scatterplots of the left- and right-hand sides of Eqs. 3.9 and 3.10 in subjects' data suggest that these relations are not verified in many instances (Fig. 8). For each relation, we

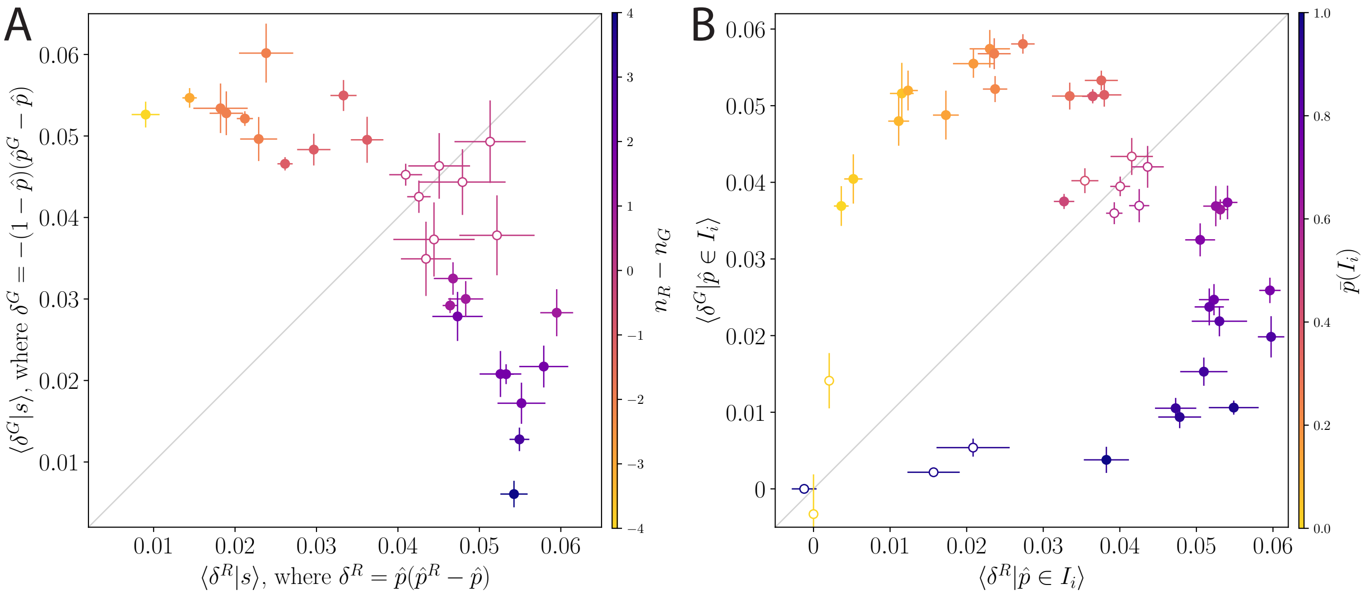

Figure 8: Subjects' revisions of their estimates are not consistent with Bayesian updating. Left-hand side vs. right-hand side of Eq. 3.9 (**A**) and of Eq. 3.10 (**B**). Each point corresponds to a different sequence of ring draws (**A**), or to a different interval containing the response $\hat{p}$ (**B**), and has as coordinates the means of the quantities $\delta^R = \hat{p}(\hat{p}^R - \hat{p})$ and $\delta^G = -(1-\hat{p})(\hat{p}^G - \hat{p})$, where $\hat{p}$ is the response at a given trial, $\hat{p}^R$ the response at the following trial in cases where the ring is red, and $\hat{p}^G$ the response at the following trial in cases where the ring is green. The length of the vertical and horizontal error bars equal twice the standard errors of the means. Bayesian inference predicts the two quantities to be equal, and thus the points to align on the first bisector. Filled points indicate where the t-test of equality of the means is rejected at the .01 level. The colors indicate (A) the difference $n_R - n_G$ between the numbers of red and green rings in each sequence, and (B) the center of the interval containing the preceding response $\hat{p}$, from $n_R - n_G = -4$ and $\hat{p}$ close to 0 (yellow), to $n_R - n_G = 4$ and $\hat{p}$ close to 1 (dark blue).

test the hypothesis that it is verified against the alternative hypothesis that the two sides of the equation can be different. More precisely, for Eq. 3.9, we test the restricted hypothesis that there is a function $m(s)$ such that $\mathrm{E}[\delta^R|s] = \mathrm{E}[\delta^G|s] = m(s)$, against the unrestricted hypothesis that the two functions $m^R(s) \equiv \mathrm{E}[\delta^R|s]$ and $m^G(s) \equiv \mathrm{E}[\delta^G|s]$ are not necessarily equal. Similarly, for Eq. 3.10, we test the restricted hypothesis that there is a function $m(I_i)$ such that $\mathrm{E}[\delta^R|\hat{p} \in I_i] = \mathrm{E}[\delta^G|\hat{p} \in I_i] = m(I_i)$, against the unrestricted hypothesis that the two functions $m^R(I_i) \equiv \mathrm{E}[\delta^R|\hat{p} \in I_i]$ and $m^G(I_i) \equiv \mathrm{E}[\delta^G|\hat{p} \in I_i]$ are not necessarily equal.

When pooling subjects' responses, the restricted hypotheses of the two tests are strongly rejected. With individual tests, the hypothesis of equality conditional on the sequence (Eq. 3.9) is rejected at the .01 level for 95% of the subjects, while the hypothesis of equality conditional on the interval to which belongs the preceding response (Eq. 3.10) is rejected at the .01 level for 71% of subjects (Table 3). We conducted, in addition, a series of additional t-tests, which strengthen these results (see Methods). We conclude that the responses of a

Restricted hypothesis	*Unrestricted hypothesis*	*Subj. pooled* p-value	*Individual tests* % $<$ .01	median p-v.
$\mathrm{E}[\delta^R \vert s] = \mathrm{E}[\delta^G \vert s], \forall s$	$\mathrm{E}[\delta^C \vert s] = m^C(s)$, $C \in \{R, G\}$	0.0***	.95	4.9e-37
$\mathrm{E}[\delta^R \vert I_i] = \mathrm{E}[\delta^G \vert I_i], \forall I_i$	$\mathrm{E}[\delta^C \vert I_i] = m^C(I_i)$, $C \in \{R, G\}$	4.9e-250***	.71	1.1e-13

Table 3: Statistical tests reject the hypothesis that subjects' updates of their estimates satisfy the Bayesian-consistency property. F-tests of the hypotheses that the responses of the subjects satisfy the equalities in Eqs. 3.9 and 3.10, predicted for Bayesian observers.

majority of subjects do not verify the consistent-updates property. Hence, their responses do not appear to be compatible with Bayesian reasoning.

Moreover, there is a clear pattern to the way in which subjects' adjustments of their estimates differ from what would be required by Bayesian updating. Bayesian updating requires that the upward adjustment of the probability estimate in the case that a red ring is drawn be different in size from the downward adjustment in the case that a green ring is drawn; the size of the adjustment in each case should be proportional to the extent to which that ring draw is a surprise (under the subjective beliefs indicated by the previous estimate $\hat{p}$). Instead, our subjects adjust their estimates by a similar amount (though in opposite directions) regardless of whether a red ring or a green ring is drawn. This results in $\delta^R > \delta^G$ (points below the diagonal in either panel of Figure 8) when a red ring is more likely to be drawn (the cases in which $n_R > n_G$, or $\hat{p} > 0.5$), and in $\delta^G > \delta^R$ (points above the diagonal in the figures) when a green ring is more likely (when $n_R < n_G$, or $\hat{p} < 0.5$). This suggests a simple interpretation of the bias: subjects pay attention to the color of the new ring that has been drawn, when deciding how to adjust their estimate following each ring draw, but pay little attention to either previously observed evidence or their previously expressed estimate. This suggests a simple model of the cognitive process reflected in our subjects' estimates, which we explain in section 5. Before exploring this model, however, we note that our finding that subjects substantially depart from Bayesian optimality calls for an examination of a prominent hypothesis regarding human inference: that people are not perfectly Bayesian, but apply instead a distorted version of Bayes' rule. We examine this hypothesis in section 4.

4 Can a "quasi-Bayesian" model of belief updating explain the biases?

In the preceding sections we have shown that subjects' responses exhibited systematic departures from optimal Bayesian inference, that could not be accounted for by a transformation of the Bayesian estimate, by Bayesian updating based on an incorrect prior, nor by a Bayesian inference process that allows for imprecise recalls of the previous ring draws. We thus turn towards a family of non-Bayesian models of inference, that nonetheless retain much of the structure of correct Bayesian inference. Indeed the model subjects in these 'quasi-Bayesian' models of inference make use of their prior and of the likelihood of observations, as does the

Bayesian observer, but they assign too much or too little importance to either or both of these quantities in their updating of their belief. Specifically, we consider a quasi-Bayesian subject who holds a belief distribution $\tilde{f}(p)$ over the proportion of red rings, and who observes one or several samples randomly drawn from the box; we denote by d this observed data, and by $P(d|p)$ the likelihood, i.e., the probability of the data given a proportion p. The belief of the quasi-Bayesian subject about p after observing d is the posterior density

$$\tilde{f}_{\rho\lambda}(p|d) = \tilde{f}(p)^{\rho} P(d|p)^{\lambda} Z_d, \tag{4.1}$$

where $\rho > 0$, $\lambda > 0$, and Z_d is a normalization constant. This update rule resembles Bayes' rule, but the prior and the likelihood are given 'weights' that distort the subject's posterior, in comparison to the Bayesian posterior. The parameter ρ is the weight of the prior in this update rule, and λ is the weight of the likelihood; correct Bayesian inference corresponds to $\rho = \lambda = 1$.

This specification is largely inspired by the literature, and several existing models belong to this quasi-Bayesian family, including the model of conservatism proposed by Phillips and Edwards (1966), the "diagnostic expectations" model of Bordalo *et al.* (2019), the representativeness model of Grether (1980) and the model of base-rate neglect of Benjamin *et al.* (2019). We examine, below, the ability of these models to capture the behavioral data. We note, however, that in this model family the posterior will be a deterministic function of the presented data, and more precisely, of a two-dimensional statistics of the data. Thus we do not expect these models to reproduce the unidimensionality of the statistics used to provide a response — except in degenerate cases where the parameters take extreme values — nor the autocorrelation in subjects' responses. But the prominence of this modeling approach in the literature prompts us to investigate how it fares in capturing at least some aspects of the data, such as the conservatism after the first few trials followed by an overreaction at later trials. Besides, after having rejected a Bayesian account of subjects' responses, we would like to examine the extent to which a more general model is more successful in capturing the behavioral data.

4.1 Representativeness and the hypothesis of an incorrect weighting of the likelihood

We start by looking at the case in which the prior is assigned the weight that it receives in correct, Bayesian inference, i.e., $\rho = 1$, but the likelihood may be under- or over-weighted, i.e., $\lambda \neq 1$. A model of this kind was first introduced by Phillips and Edwards (1966) as an account of conservatism in 'bookbag-and-poker-chips' inference tasks (in which subjects are asked to estimate the probability that either of two bags is the one from which are drawn the poker chips presented to them.) Their results, along with those of many subsequent studies, suggest that $\lambda < 1$, i.e., that the likelihood is underweighted (see Benjamin (2019) for a review). A more recent proposal, the model of "diagnostic expectations" proposed by Bordalo *et al.* (2019), conversely implies an overweighting of the likelihood. This theory is of particular relevance for our experiment because it is explicitly proposed for settings in which probability judgments are constantly revised in response to a stream of sequential evidence (such as economic or financial forecasting,) and because it allows for over-reaction of the kind that we observe in the case of longer sequences of ring draws. We introduce a model that is inspired by this theory, and which belongs to our family of quasi-Bayesian models, with $\rho = 1$ and an overweighting of the likelihood ($\lambda \geq 1$). We then examine the responses of quasi-Bayesian observers with $\rho = 1$ (but no specific restriction on λ other than $\lambda > 0$).

The theory of diagnostic expectations is motivated by Kahneman and Tversky's representativeness heuristic (Kahneman and Tversky, 1972; Tversky and Kahneman, 1974), and proposes that estimates are based not on the correct Bayesian posterior, given the stream of evidence observed, but rather on a posterior density distorted by a measure of representativeness. Here, we consider a more flexible hypothesis, in that it allows people to start from an incorrect prior, as assumed (for example) by Zhu *et al.* (2020). Our hypothesis thus nests both the case of diagnostic expectations starting from a correct prior (as in the exposition of Bordalo *et al.*), and the hypothesis of correct Bayesian updating starting from an incorrect prior.

Let $f(p|x_{1:t})$ be the posterior of a Bayesian observer starting from a prior $f(p)$ and who observes a sequence of rings $x_{1:t}$. We define the representativeness of the proportion p as the ratio

$$\frac{f(p|x_{1:t})}{f(p)}. \tag{4.2}$$

A proportion is thus more representative if it is more probable, when given the evidence $x_{1:t}$, in comparison with its probability in the absence of evidence. Following Bordalo *et al.* (2019), we consider the hypothesis that the posterior is distorted by representativeness; specifically, that subjects derive their responses from the density

$$\tilde{f}_\theta(p|x_{1:t}) = f(p|x_{1:t})\left(\frac{f(p|x_{1:t})}{f(p)}\right)^\theta Z, \tag{4.3}$$

where Z is a normalization constant, and $\theta \geq 0$ is a parameter controlling the degree of the distortion (if $\theta = 0$, there is no distortion). The Bayesian posterior is proportional to the product of the prior and of the likelihood, i.e., $f(p|x_{1:t}) \propto f(p)P(x_{1:t}|p)$, and thus the posterior distorted by representativeness is

$$\tilde{f}_\theta(p|x_{1:t}) = f(p)P(x_{1:t}|p)^{1+\theta} Z_{x_{1:t}}, \tag{4.4}$$

where $Z_{x_{1:t}}$ is a normalization constant. The model subject under this theory is thus a 'quasi-Bayesian observer' with $\rho = 1$ and $\lambda = 1 + \theta$, and it follows that $\lambda \geq 1$, i.e., the likelihood is overweighted (or it has the correct weight, if $\theta = 0$). However we do not restrict our analysis to this case, and below we allow λ to be smaller or larger than 1.

In order to examine the responses of this model subject (with $\rho = 1$ and $\lambda \neq 1$), we assume, for the reasons exposed in section 1.3, that the subject's prior is a symmetric Beta prior with parameter α. The resulting distorted posterior, in the context of our task, is a Beta distribution with parameters $\lambda n_R + \alpha$ and $\lambda n_G + \alpha$, where n_R and n_G are the number of red and green rings in the sequence $x_{1:t}$. The subject provides as a response the mean of the distorted density, which we can also formulate as a function of the optimal estimate, as

$$p_{\alpha\lambda} = \frac{\lambda n_R + \alpha}{\lambda t + 2\alpha} = p^* + (2p^* - 1)\frac{1 - \alpha/\lambda}{t + 2\alpha/\lambda}. \tag{4.5}$$

This hypothesis is thus equivalent to the erroneous-prior hypothesis, with the parameter α replaced by the ratio α/λ (see Eq. 1.5). If this ratio is greater than 1 (for instance, if $\lambda < 1$ and α is correctly set to 1, as in Phillips and Edwards' model,) then the model subject exhibits 'conservatism', or 'underinference'. If conversely the ratio is lower than 1 (for instance, if $\lambda > 1$ and $\alpha = 1$, as in the model of Bordalo *et al.*), then there is overreaction to the evidence presented. But this model cannot reproduce the conservatism followed by overreaction that we observe in behavioral data; and the hypothesis that subjects' estimates are consistent with this model is rejected for all the subjects (Table 1, line 6). We conclude that assigning an incorrect weight to the likelihood in the updating of beliefs is insufficient to account for subjects' behavior, even if we also allow the prior to differ from the uniform distribution used in the experiment.

4.2 Base-rate neglect and the hypothesis of an incorrect weighting of the prior

The representativeness heuristic proposed by Tversky and Kahneman (1974) is not only concerned with the way people take into account the likelihood of observations in the formation of their beliefs: it also predicts an 'insensitivity' to the prior probabilities of outcomes, a behavioral pattern known as 'base-rate neglect'. In the formalizations of this reduced sensitivity proposed by Grether (1980) and by Benjamin *et al.* (2019), people are quasi-Bayesian observers who assign an incorrect weight on the prior in their update rule, i.e., $\rho \neq 1$. For Benjamin *et al.*, following the conclusions of the extensive review, conducted by Benjamin (2019), of the literature on 'bookbag-and-poker-chip' experiments, base-rate neglect results from the weight on the prior being lower than 1, i.e., $\rho < 1$, while in Grether's approach, the weight on the prior is lower than the weight on the likelihood, i.e., $\rho < \lambda$.

Before we examine the case in which $\rho \neq 1$, we must first specify how the quasi-Bayesian update rule is applied when several samples are sequentially presented. A first theory is one in which the subject's original prior (before any sample is presented) is exponentiated by the exponent ρ, while the likelihood of the entire sequence of pieces of evidence is exponentiated by the exponent λ; i.e., after observing the last element of the sequence $x_{1:t}$, the subject applies the update rule (Eq. 4.1) with $d = x_{1:t}$, and with $\tilde{f}(p)$ equal to its initial prior. In a second theory, the subject applies the update rule recursively: each time an additional piece of evidence is observed, a posterior is formed using the update rule, and then this posterior becomes the new prior for the application of the update rule when the next piece of evidence is observed; i.e., after observing the last element of the sequence $x_{1:t}$, the subject applies the update rule (Eq. 4.1) with $d = x_t$ and with $\tilde{f}(p)$ equal to the posterior $\tilde{f}_{\rho\lambda}(p|x_{1:t-1})$ derived after the preceding observation.

Under the first theory, if the initial prior of the subject is the correct, uniform density over $[0, 1]$, then the exponentiated prior is also the uniform density, and thus the value of ρ has no impact. If the prior is incorrect, then we note that the behavior of a model subject who holds the initial prior $\tilde{f}(p)$ and assigns to it the weight ρ when updating its beliefs is identical to the behavior of a model subject whose initial prior is the same function raised to the power ρ and normalized, but who assigns to it the correct weight when updating its beliefs. Thus in any case, under the first theory, the behavior of the model subject is essentially already addressed in the previous section (in which we have assumed $\rho = 1$). Thus henceforth we focus on the second theory, by which the quasi-Bayesian update rule is applied recursively.

If $\rho \neq 1$, the recursive model has different implications than the first theory. In particular, while in the first theory the effect of a given set of observations of red and green rings should be independent of the order in which they are observed (so that n_R, n_G continue to be sufficient statistics for the sequence s), this is no longer true in the recursive model when $\rho \neq 1$. Instead we should observe 'order effects': two sequences that have the same length, $t \geq 2$, and the same number of red rings, n_R (with $0 < n_R < t$), will not result in the same estimates if the red and green rings appear at different positions in the two sequences. We conduct an array of tests to look for such order effects in our subjects' data, and find that a majority of subjects do not exhibit significant order effects (for details see Methods), contrary to the results obtained by authors such as Ashinoff *et al.* (2022) using the more familiar "balls and urns" paradigm. And indeed, when we estimate the parameters of the recursive quasi-Bayesian model (as discussed below), the value of the parameter ρ that best fits subjects' data is close to 1. (When fitting individual subjects, we find that the best-fitting parameter ρ for 81% of the subjects is between 1 and 1.06.) Nonetheless, allowing for a value of ρ different from 1 allows the model to better fit some aspects of subjects' behavior.

Simply allowing ρ to differ from 1 is not enough, however. If we allow $\rho \neq 1$, but start from the correct, uniform prior (we consider the case of an incorrect prior in the next section), and if we assume that the likelihood is correctly weighted ($\lambda = 1$), then the model cannot account for the observed pattern of bias. After the first ring draw ($t = 1$), the prior will still be the correct, uniform one even if raised to a power ρ different from 1; thus after only one ring draw the quasi-Bayesian inference with $\lambda = 1$ will be the same as correct Bayesian inference, so that the subject's estimate should equal p^*. The strong conservatism of subjects after the first ring draw is thus inconsistent with this specification. Similarly, we show in Methods that even after more than one ring ($t > 1$), this model predicts either consistent over-reaction to the evidence (if $\rho > 1$) or consistent under-reaction (conservatism, if $\rho < 1$). The model cannot reproduce the subjects' initial conservatism followed by eventual reversal of the bias, illustrated in Fig. 4. Hence the sole hypothesis of a misweighting of the prior in the application of Bayes' rule is insufficient to account for the main patterns in our behavioral data.

The model can predict a wider range of types of behavior if both the prior and the likelihood can be assigned incorrect weights ($\rho \neq 1$ and $\lambda \neq 1$). But even in this case, it cannot account for the observed pattern of estimation biases if we assume 'base-rate neglect' in a way compatible with the experimental findings of Grether (1980), i.e., with $\rho < \lambda$; or in a way consistent with the conclusions of the meta-analysis of Benjamin (2019), i.e., with $\rho < 1$. Starting from the correct prior, the model reproduces the conservatism of subjects at the first trial ($t = 1$) only if the likelihood is underweighted, i.e., if $\lambda < 1$. But then 'base-rate neglect,' under either formulation, requires that $\rho < 1$ as well. In this case, a model subject necessarily under-reacts to evidence for all t, since an observation k trials earlier than the current one receives a multiplicative weight $\lambda\rho^k < 1$ rather than the weight of 1 that it would receive under correct Bayesian updating. The model is capable of predicting conservatism for low t together with over-reaction for larger values of t (as observed in our subjects) only if we suppose that $\lambda < 1$ but $\rho > 1$. (This combination of inequalities represents a necessary but not a sufficient condition for reversal of the sign of the estimation bias as n increases; for details see Methods.) Thus the biases observed in our experiment are inconsistent with typical parameterizations of quasi-Bayesian updating. The model can reproduce the pattern we observe in data if we assume $\lambda < 1 < \rho$. We further comment on this parameterization, below; but a significantly better account of the data is obtained if an incorrect prior is assumed, and thus we first turn to this case.

4.3 Quasi-Bayesian updating from an incorrect prior

We can further generalize the class of models considered by also allowing the model subject's prior to differ from the correct, uniform prior. As above, we assume that the prior is a symmetric Beta distribution with parameter α. Under this assumption, the subject's posterior after observing t samples, $x_1, \ldots, x_t$, is a Beta distribution with parameters $\tilde{n}_R + 1$ and $\tilde{n}_G + 1$, where $\tilde{n}_R$ and $\tilde{n}_G$ are exponentially-weighted counts of the red and green rings

$$\tilde{n}_R = \lambda \sum_{i=0}^{t-1} \rho^i x_{t-i} + \rho^t(\alpha - 1) \text{ and } \tilde{n}_G = \lambda \sum_{i=0}^{t-1} \rho^i (1 - x_{t-i}) + \rho^t(\alpha - 1). \tag{4.6}$$

The subject's estimate is then the mean of the posterior,

$$p_{\alpha\rho\lambda} = \frac{\tilde{n}_R + 1}{\tilde{n}_R + \tilde{n}_G + 2}. \tag{4.7}$$

We note that to keep track of $\tilde{n}_R$ and $\tilde{n}_G$ over the course of several successive trials, it is not necessary to remember the whole sequence of ring draws: these quantities can be updated recursively after each new observation, x, by replacing $\tilde{n}_R$ by $\rho\tilde{n}_R + \lambda x$, and $\tilde{n}_G$ by $\rho\tilde{n}_G + \lambda(1-x)$. (With $\rho = 1$ and $\lambda = 1$, this amounts to a simple count of the red and green rings.) The responses of the quasi-Bayesian observer are thus functions of a bidimensional summary statistics of the evidence presented, just as with the Bayesian observer, with the simple counts n_R and n_G replaced by their counterparts $\tilde{n}_R$ and $\tilde{n}_G$.

If we assume that the weight on the likelihood is correct ($\lambda = 1$), and allow α and ρ to be different from 1, then we show in Methods that it is possible to obtain conservatism followed by overreaction to evidence. A necessary (though not sufficient) condition is the combination of inequalities $\alpha > 1$ and $\rho > 1$, i.e., that the initial prior be more concentrated around middle values than the uniform prior, and that the prior be overweighted in the inference. So far we have thus seen two specifications of the model of quasi-Bayesian inference that are able to reproduce qualitatively the pattern we find in the behavioral data: $\{\alpha = 1, \rho > 1, \lambda < 1\}$, as seen in the previous section, and $\{\alpha > 1, \rho > 1, \lambda = 1\}$ (in both cases, the two inequalities are necessary but not sufficient conditions). However, the model comparison presented in the next section shows that the model of quasi-Bayesian inference with $\alpha = 1$, and the one with $\lambda = 1$, with in each case the other two parameters being allowed to differ from 1, both yield much poorer fits than the one in which the three parameters, α, ρ, and λ, are allowed to be different from 1. In addition, with the subjects' best-fitting parameters, these two models do not in fact reproduce the overreaction to evidence observed in subjects' data (see Fig. 13 in Methods). For these reasons, we focus here on the version of the model in which α, ρ, and λ are all free parameters.

We fit these three parameters to our subjects' data, under the additional assumption that the responses include Gaussian noise (the model-fitting procedure is detailed in section 5). We find that the best-fitting value of ρ is relatively close to 1 (1.011), while those of α and λ are relatively small: 0.067 for α, and 0.022 for λ. With these values, the average response to sequences featuring n_R red rings and n_G green ring can be approximated (see Methods) as

$$\bar{p}_{\alpha\rho\lambda} \approx \frac{1}{2} + \frac{\lambda}{4\alpha}(n_R - n_G). \tag{4.8}$$

Note that this approximate average response does not depend on $t = n_R + n_G$, the total number of rings drawn, but only on $n_R - n_G$, the difference between the numbers of red and green rings in the sequence. In other words, when fitted to subjects' data, the model of a quasi-Bayesian observer equipped with an incorrect prior approaches a model in which responses are determined by a unidimensional summary statistic of the presented sequence—precisely the kind of behavioral pattern that we have exhibited in section 2.2.

Furthermore, we can write the (approximate) average response to sequences that include n_R red rings and n_G green rings as a function of the optimal estimate, as

$$\bar{p}_{\alpha\rho\lambda} \approx p^* + (2p^* - 1)\frac{1}{2}\Big[\frac{\lambda}{2\alpha}(t+2) - 1\Big]. \tag{4.9}$$

If the term in the brackets is negative, the model subject exhibit conservatism, while if it is positive the model subject overreacts to the evidence, as compared to the optimal subject. We note that this term increases with the number of samples, t; in particular, it is negative up to $t = 3$ (and thus the model subject exhibits conservatism up to the third sample) if the ratio λ/α is lower than $2/5 = 0.4$; while it is positive for $t \geq 5$ (and thus the model subject exhibits overreaction from the fifth sample on) if this ratio is greater than $2/7 \approx 0.29$. The values of λ and α that best fit subjects' data satisfy both of these conditions ($\lambda/\alpha = 0.33$). In short, this model, with the subjects' best-fitting parameters, is able to reproduce the conservatism followed by overreaction that we observe in subjects' data.

Although the quasi-Bayesian model captures these aspects of subjects' behavior, it remains unsatisfactory in other respects. First, it does not capture the autocorrelation in responses documented above. Second, the test of the restricted hypothesis that the average response equals $p_{\alpha\rho\lambda}$ against the unrestricted hypothesis that it is a more general function $m(s)$ of the presented sequence is rejected for 81% of subjects (Table 1, line 7). Third, the values of the parameters α and λ that best fit subjects' data are relatively extreme, suggesting that they correspond to a degenerate case of the model: $\alpha = 0.067$ corresponds to a prior on $[0, 1]$ in which 74% of the probability mass is either below 0.01 or above 0.99 (versus 2% under the correct, uniform prior). We deem it unlikely that the subjects hold such an extreme belief. And the value $\lambda = 0.022$ implies that new observations have only a very small impact on estimates, which also seems implausible.

To the extent that the quasi-Bayesian model can fit the data, this seems to reflect the fact that it can be parameterized so as to imply that subjects respond mainly to the net difference between red and green rings, a simple heuristic that would already account for both the subjects' initial conservatism and its subsequent reversal. Above, we concluded section 3 by noting that subjects seemed to decide on the adjustment of their estimate mostly by paying attention to the color of the new ring, but neglecting other kind of information, such as the total amount of evidence presented, t. Together, our results suggest a simple "noisy-counting" model of the subjects' cognitive process, which we detail in the next section.

5 A "Noisy-Counting" Model of Probability Estimation

We have shown that our subjects' responses do not represent optimal (Bayesian) responses, conditional on the information contained in the noisy cognitive state on the basis of which the response is generated. Instead, our results indicate that subjects' adjustments of the position of the slider following a new ring draw are largely insensitive to the information about previous evidence that would instead modulate the size of the adjustment in the case of an ideal Bayesian statistician.

5.1 Informational insensitivity of the adjustment decision

As mentioned above (Eq. 2.3), the ideal Bayesian response rule (Eq. 1.1) requires that the amount by which the slider should be adjusted following the observation of an additional red ring will equal

$$\hat{p}^R - \hat{p} \;=\; \frac{1}{t+3}\,(1-\hat{p}), \tag{5.1}$$

where t is the number of ring draws upon which the estimate $\hat{p}$ was based. This implies that for any given current slider position $\hat{p}$, the size of the upward adjustment should be smaller the larger the number of rings t that had already been observed; and also that for any number of ring draws, the size of the upward adjustment should be smaller the larger the existing estimate $\hat{p}$. Instead, in our data there is very little sensitivity of the adjustment size to the values of either t or $\hat{p}$.

We first consider the dependence of subjects' adjustments on t. Our results suggest that the average of a subject's response $\hat{p}_{t+1}$ obtained after the presentation of a new ring, $x_{t+1} = R$ or G, can be predicted by the preceding response, $\hat{p}_t$, in a way that is independent of the sequence length, t (see Fig. 6, top-right panel). This is contrary to the prediction of Eq. 5.1, but consistent with our previous observation that subjects' cognitive states appear to evolve along a line; a unidimensional cognitive state suffices to determine a current estimate $\hat{p}$, but cannot differentiate between different experiences that lead to the same estimate $\hat{p}$ after sequences of observations of different lengths.

This is shown more clearly in panel A of Figure 9, where we plot the distribution of estimates $\hat{p}^R_{t+1}$ following observation of another red ring as a function of the previous estimate $\hat{p}_t$, for each of the different values of t. The solid lines show the mean value $\bar{p}^R_{t+1}$ as a function of $\hat{p}_t$ for each value of t, while the dotted lines indicate the range between the 5th and 95th percentiles of the response distributions, as a function of $\hat{p}_t$, for each value of t. It is evident that there is very substantial overlap between the distributions associated with different values of t.

We statistically test the dependence of the adjustment size on the number of rings drawn in the following way. For each value of $\hat{p}$, we can compute an average value of the subsequent response following a red ring, $\bar{p}^R(\hat{p})$, averaging over trials in which $\hat{p}$ is based on different amounts of evidence. On any individual trial, we can then compute the deviation of $\hat{p}^R_{t+1}$

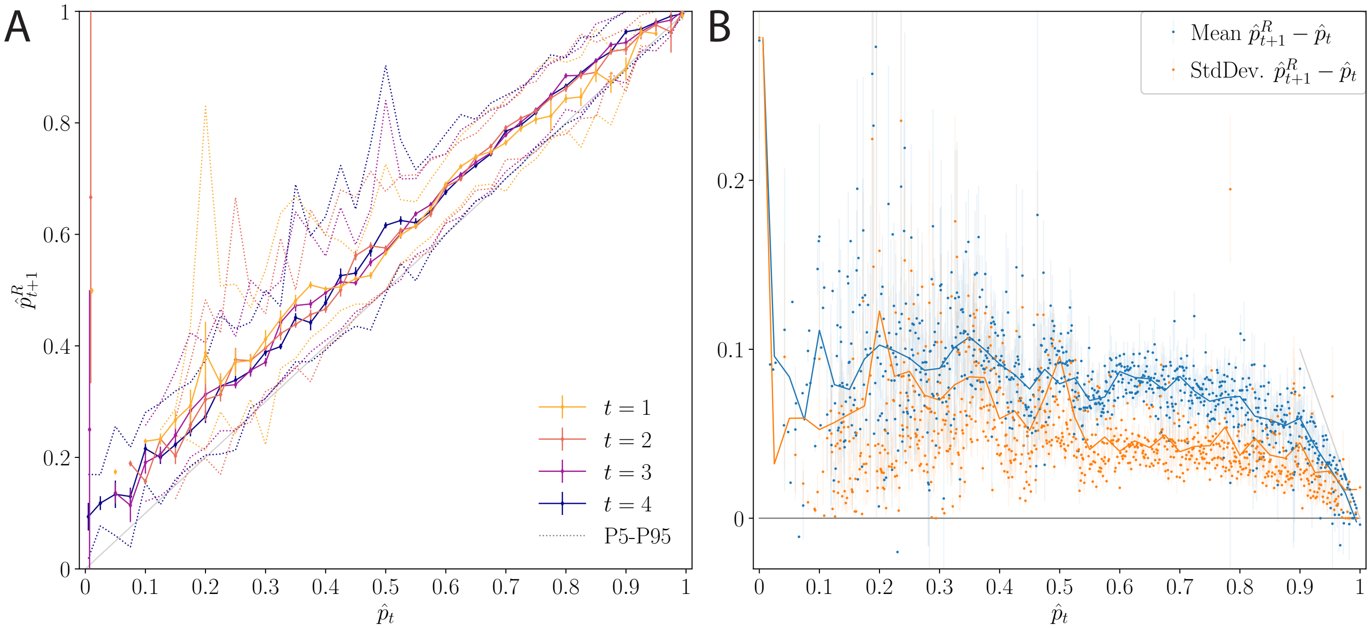

Figure 9: The adjustments in subjects' responses following the observation of a red ring do not vary strongly with the sequence length, nor with the preceding response, except for preceding responses close to 1. (A) Average responses after a red ring, $\hat{p}^R_{t+1}$ (solid lines), and fifth and 95th centiles of the distributions of responses following a red ring (dotted lines), for sequence lengths from $t = 1$ to 4, as a function of the preceding response, $\hat{p}_t$. The error bars equal twice the standard error of the mean. Averages are taken over intervals of length .025. (B) Averages (blue) and standard deviations (orange) of the adjustments in responses following a red ring, $\hat{p}^R_{t+1} - \hat{p}_t$, as a function of the preceding responses, $\hat{p}_t$. Dots correspond to single responses, $\hat{p}_t$, and lines to intervals of length .025.

t	p-value	$\% < .01$
1	.26	.19
2	.17	.14
3	.24	0
4	.25	.29

Table 4: Average responses do not depend on the number of rings that have been observed, for most subjects. Student's t-tests of the hypotheses that, for each sequence length, t, the deviations conditional on the sequence length, $\langle \hat{p}^R_{t+1} - \bar{p}^R(\hat{p}) | t \rangle$, are on average zero, where $\bar{p}^R(\hat{p})$ is the unconditional (not conditional on t) average response following a response $\hat{p}$ and the observation of an additional red ring. A large p-value indicates that we cannot reject the hypothesis that the average response conditional on the sequence length is equal to the average unconditional response, thus suggesting that the sequence length does not influence subjects' responses.

from this average; and we can obtain a sample distribution of deviations from the average adjustment, for each of the different values of t. We wish to test the hypothesis that these different distributions all have the same mean, zero. An F-test of the joint hypothesis that the means are zero in all four cases $t = 1, 2, 3, 4$ has a p-value of 0.117, so that we fail to reject the null hypothesis of equality at the ten percent level of significance. Table 4 (middle column) similarly shows the p-values for t-tests of the null hypothesis that the mean deviation is zero for each of the individual values of t; again, we fail to reject any of these null hypotheses at even the ten percent level.

We conduct the same analyses at the individual level. For the F-test, the null hypothesis is rejected at the .05 level for 52% of subjects, and at the .01 level for 43% of subjects. As for the t-tests, the p-values indicate that for up to 29% of subjects the null hypothesis that the average conditional deviation is zero is rejected at the .01 level (Table 4, last column). These results suggest that there are subjects for whom the number of rings observed, t, does have some effect on the new response, $\hat{p}_{t+1}$. To obtain a measure of the magnitude of this influence, we compute for each subject the ratio of the variance of the means of the deviations, $\mathrm{Var}[\langle \hat{p}^R_{t+1} - \bar{p}^R(\hat{p}_t) | t \rangle]$, divided by the total variance of the deviations, $\mathrm{Var}[\hat{p}^R - \bar{p}^R(\hat{p})]$. Across the subjects, the median of this ratio is 1.3%, and 90% of subjects have a ratio below 5%, suggesting that the quantitative impact of the number of rings observed is limited.

Next we consider the way in which the average adjustment size $\hat{p}^R(\hat{p}) - \hat{p}$ depends upon a subject's existing estimate $\hat{p}$. In the case of a Bayesian ideal observer, Eq. 5.1 requires that this should decrease linearly with increasing values of $\hat{p}$. Instead, panel B of Figure 9 shows that the mean adjustment is of about the same size (approximately 0.08 to 0.09) over the entire range of values of $\hat{p}$ between 0.1 and 0.9. (Equation 5.1 would instead require it to be 9 times as large at one end of that interval as at the other.) The adjustments are smaller in the case of values of $\hat{p}$ close to 1, but this is largely a mechanical consequence of the fact that the slider cannot be moved much higher when it is already near its upper bound. (There is also some evidence of larger adjustments when $\hat{p}$ is close to zero, but the number of observations is not large enough for this to be estimated very precisely.) As explained at the end of the previous section, the large departures from the diagonal in both panels of Figure 8 also result from the fact that the average adjustment size varies very little with the subject's previous estimate of the probability of drawing another red ring.

Panel B of Figure 9 also plots the standard deviation of the distribution of adjustment sizes as a function of $\hat{p}$, and this is also relatively constant over much of the range of possible values of $\hat{p}$. (Again, it is smaller for values of $\hat{p}$ near 1, but this can largely be attributed to the fact that the amount that the estimate can be increased is mechanically bounded in these cases.) Thus our data indicate that the distribution of adjustments following observation of an additional red ring is much the same, regardless of the estimate $\hat{p}$ prior to this observation, except for the fact that the distribution is necessarily truncated by the subject's inability to express a probability estimate higher than 1. The same is true for the distribution of adjustments following observation of an additional green ring, except that the sign of the adjustments is reversed in this case, and the distribution is truncated by the inability to express a probability estimate lower than 0.

5.2 A quantitative model of "noisy counting"

This suggests that to a reasonable approximation, our subjects' responses are consistent with a model of "noisy counting." According to this model, a subject keeps a running count of the net amount of evidence in favor of a higher rather than a lower value of p, adding an additional positive increment to the count whenever a red ring is observed, and an additional negative increment whenever a green ring is observed. (We suppose that the process starts from a neutral belief — the value of the count corresponding to an estimate $\hat{p} = 0.5$ — before the first ring is drawn.) This is like a rule of simply counting the net excess of red rings over green rings, in that if the size of each increment were a constant, the cognitive state would correspond to (a linear transformation of) the quantity $n_R - n_G$. However, we posit "noisy counting," because the size of the increment on any given trial is a draw from a probability distribution of possible adjustment sizes. From this hypothesis we derive several versions of the noisy-counting model, which we now present.

In a first version, a subject's estimate of the value of the hidden probability evolves according to a law of motion

$$\hat{p}_{t+1} \;=\; [\hat{p}_t \,+\, x_{t+1}y_{t+1}]_{[0,1]}, \tag{5.2}$$

where $\hat{p}_t$ is the estimate after t rings have been observed, x_{t+1} is the sign of the $t+1$st observation (taking the value $+1$ in the case of a red ring, and -1 in the case of a green ring), and y_{t+1} is an independent draw from a distribution $F(y)$ of possible adjustment sizes. Note that the distribution $F(y)$ is assumed to be the same for adjustments of either sign, and independent of both the previous estimate $\hat{p}_t$ and all of the previous observations $x_{1:t}$. For any real number x, we use the notation $[x]_{[0,1]}$ to indicate the value of x truncated to remain within the interval $[0, 1]$. Thus Eq. 5.2 indicates that the probability estimate is increased or decreased (depending on the color of ring that is drawn) by the random amount y_{t+1}, except when this would take the estimate outside the range $[0, 1]$; in the latter case, the new estimate $\hat{p}_{t+1}$ is given by the most extreme feasible value. A complete specification of the model requires that we specify the distribution $F(y)$ from which adjustment sizes are drawn. One simple choice is to assume a normal distribution,

$$y \;\sim\; N(\mu, \sigma^2), \tag{5.3}$$

with parameters μ, σ^2 to be estimated. (This is the model considered on line 8 of Table 5.)

In this version of the noisy-counting model, the randomly evolving cognitive state (on the basis of which the subject's response is chosen) is identified with the response itself; the cognitive state r_t on any trial is simply the response $\hat{p}_t$ itself. (Thus in Equation 5.2, we have directly written a stochastic law of motion for the estimate.) Alternatively, we might suppose that the cognitive state r_t is a noisy count, which must then be converted into an estimate using a response rule. In this case, we need not assume that the noisy count itself is truncated to remain within the interval $[0, 1]$; the truncation may instead be part of the response rule. We might also assume additional stochastic noise in the response rule that determines $\hat{p}_t$, over and above the noise in the counting process.

Thus we might alternatively assume that the cognitive state r_t is a noisy count, updated according to a law of motion

$$r_{t+1} \;=\; r_t \,+\, x_{t+1}y_{t+1}, \tag{5.4}$$

starting from an initial condition $r_0 = 0.5$, and that the subject's estimate on any trial is then given by a noisy truncation rule

$$\hat{p}_t \;=\; [r_t \,+\, \varepsilon_t]_{[0,1]}, \tag{5.5}$$

where ε_t is a draw from the distribution $N(0, \nu^2)$, independent of the evolution of the cognitive state and of the response noise on any other trials. If, for example, we assume that y_{t+1} is drawn from a Gaussian distribution (Eq. 5.3), the model is completely specified by values for the parameters μ, σ^2, and ν^2. (This is the model considered on line 10 of Table 5.) Note that although we assume normal distributions for both the update of the cognitive state (Eqs. 5.3-5.4) and the response noise (Eq. 5.5), the two types of additive Gaussian noise do not have equivalent effects, so that their respective variances can be estimated. Indeed, while the effects of the realization of y_t on the cognitive state are propagated to later periods, the size of the response noise ε_t has no consequences for responses in later periods.

5.3 Comparison of behaviors favors noisy counting over response noise

If (on grounds of parsimony) only one type of noise were to be assumed, it is clearly more important to allow for noise in the evolution of the cognitive state, in order to account for the features of our data. Figure 10 recalls some of the key regularities discussed above, and compares the ability of four models to account for them. The top row of the figure reproduces figures illustrating five different aspects of our subjects' data, already discussed above. The other rows show the corresponding plots for data from simulations of four different theoretical models, where in each case the free parameters of the theoretical model are chosen so as to maximize the likelihood of the subjects' data.

The second row considers a model in which subjects are assumed to correctly observe the ring draws and to perform correct Bayesian inference, but they start from an incorrect prior Beta(α, α). (Truncated Gaussian response error of the kind modeled in Eq. 5.5 is added to the Bayesian estimate, in order for the model to have a well-defined likelihood function.) In the third row, we consider a counting model in which the cognitive state evolves in accordance with Eq. 5.4, but there is no noise in the count ($y_{t+1} = \mu$ with certainty; i.e., we assume that $\sigma = 0$), and the subject's response is assumed to be given by Eq. 5.5, where we allow ν^2 to be positive. The fourth row instead considers a model in which the noise and the truncation are assumed to occur at the level of the cognitive state, while there is assumed instead to be no response noise. In this case, the cognitive state can be identified with the subject's estimate $\hat{p}_t$, and the law of motion for the cognitive state is given by Eq. 5.2, in which y_{t+1} is assumed to be drawn from a Gaussian distribution (Eq. 5.3). In short, in the third row of Fig. 10, there is noise in the model subject's responses, but not in the cognitive states, while in the fourth row there is noise in the cognitive states, but not in the responses. We note that the latter model of counting noise is not the one that best fits our data (see Table 5); but considering this case allows the most direct comparison with the alternative model in the third row of Fig. 10 (with noise in responses), since the number of free parameters is then the same in both models (two: μ and a single noise variance), and both models assume additive Gaussian noise prior to the truncation. We comment on the fifth row of Fig. 10 further below.

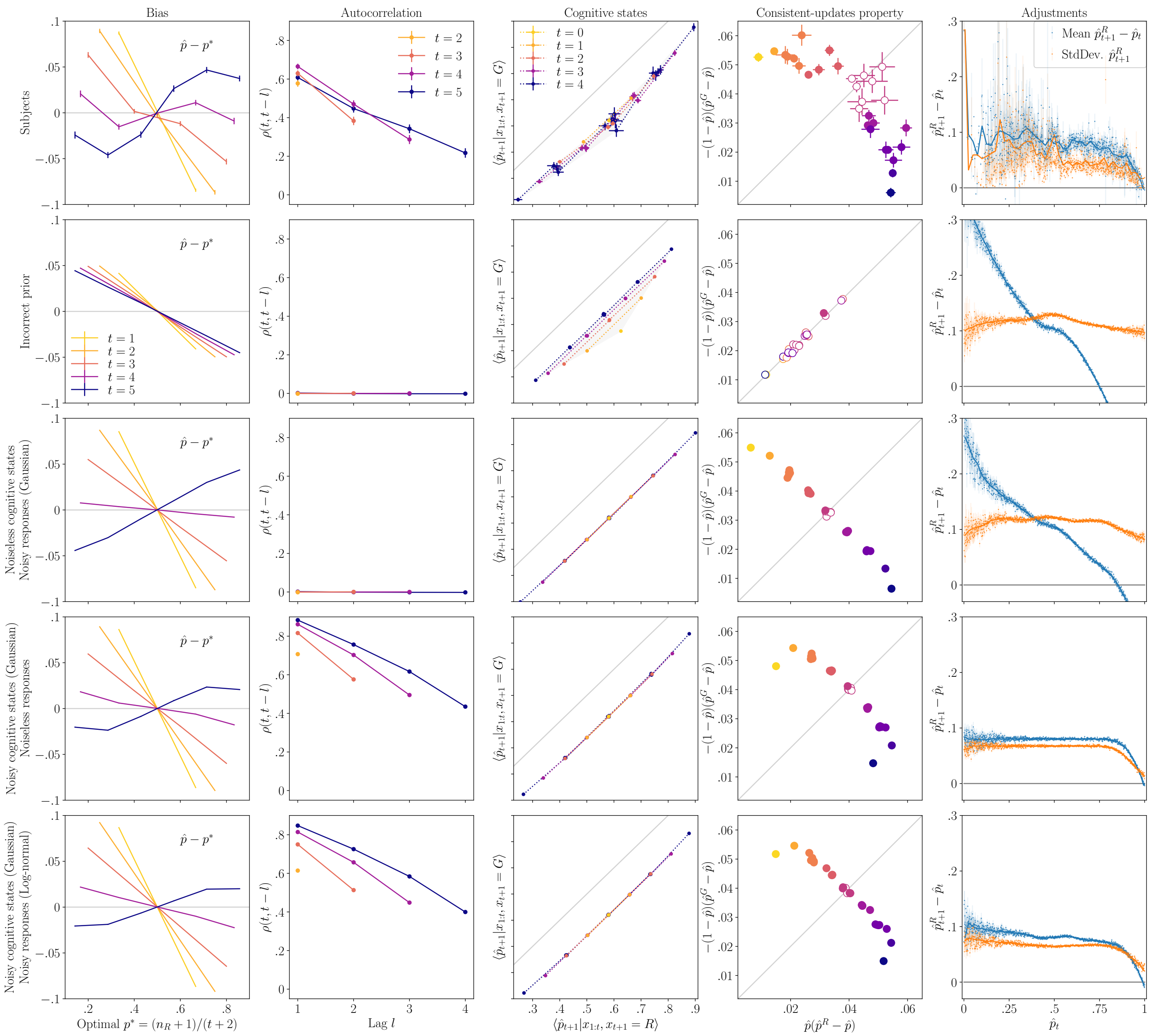

Figure 10: The model in which a scalar cognitive state is updated with noise qualitatively reproduces subjects' behavioral patterns. Behavior of the subjects (first row), of a Bayesian observer equipped with an incorrect prior (second row), of the model featuring noiseless updates of a scalar cognitive state, and noisy (Gaussian) responses (third row), of the model featuring noisy (Gaussian) updates of a cognitive state, and noiseless responses (fourth row), and of the model with noisy (Gaussian) updates of a cognitive state, and noisy (log-normal) responses (last row). *First column:* biases in responses as a function of the optimal estimate, for different sequence lengths (as in Fig. 4B). *Second column:* serial correlation in the responses at two trials in a sequence, as a function of the distance between the two trials (as in Fig. 5B). *Third column:* average responses after observing a given sequence followed by a green ring, vs. average responses after observing the same sequence followed by a red ring (as in Fig. 6, second row). *Fourth column:* Bayesian consistent-updates property, with the quantities δ_R and δ_G averaged over sequences (as in Fig. 8A). *Last column:* response adjustment following a red ring, $\hat{p}^R_{t+1} - \hat{p}_t$, as a function of the preceding response, $\hat{p}_t$ (as in Fig. 9B). The models were simulated on 100 times more trials than the subjects faced.

We observe that the model of Bayesian inference based on an incorrect prior (second row) fails to capture a number of salient features of the data: it fails to predict the pattern of under-reaction to evidence when $t \leq 3$ combined with over-reaction after more rings are observed (the first column of Fig. 10); it fails to predict positive serial correlation in subjects' responses (the second column); it predicts that the cognitive state should be two-dimensional rather than unidimensional (the third column); it fails to predict the pattern of systematic deviation from the Bayesian-consistency property in subjects' updates (the fourth column); and it predicts counterfactually that the mean adjustment after observing a red ring should be a sharply decreasing function of $\hat{p}_t$ (the fifth column).

Both variants of the noisy-counting model (third and fourth rows) do better on several of these counts: they predict the pattern of under-reaction switching to over-reaction as t increases; they predict the existence of a unidimensional cognitive state; and they get the nature of the departures from the Bayesian-consistency property of estimate updates broadly correct. However, the counting model with only response noise (third row) still fails to predict positive serial correlation of responses (the second column), and still predicts that the mean adjustment after observing a red ring should be a sharply decreasing function of $\hat{p}_t$ (the fifth column). The noisy-counting model in which the noise is in the evolution of the cognitive state instead (fourth row) makes the right qualitative predictions in all five columns.

In addition to the five behavioral patterns just examined, we showed above that the responses of subjects were not 'well calibrated' (Fig. 7). This sixth pattern in the behavioral data is also well reproduced by the model with noise in the cognitive states. However, we find that the responses of most of our models exhibit as well a similar pattern of deviations from the calibrated-responses Bayesian property. This analysis is thus not strongly differentiating, and we have not included it in this comparison; but the reader can find it in Methods.

Finally, we have also run simulations of the quasi-Bayesian models. As noted in section 4.3, the responses of the model in which the prior may be incorrect ($\alpha \neq 1$) and the prior and the likelihood may be assigned incorrect weights ($\rho \neq 1$ and $\lambda \neq 1$), and which is able to reproduce more behavioral patterns than the other quasi-Bayesian models, are well approximated by a linear function of the difference $n_R - n_G$ (Eq. 4.8). The behavior of this model is thus not very different from that of the model of counting with response noise, shown in the third row of Fig. 10. We report in Methods on the behavior of this model and of the other quasi-Bayesian models.

5.4 Model fitting

We can quantify the relative fit of these models, as well as a number of other alternatives, using the Bayes Information Criterion (BIC; Schwarz, 1978) as a basis for model comparison. (This allows us to compare models that differ in the number of free parameters, by penalizing the use of additional parameters.) Table 5 compares the fit of twelve alternative models. The first three lines of Table 5 consider models in which the cognitive state is assumed to be two-dimensional and consisting of the quantities n_R and n_G, or alternatively t and n_R (which suffices for implementation of the estimation rules that we consider here, as these are sufficient statistics for the information about the value of p contained in the history of ring draws.) In the models of the next four lines of Table 5, the cognitive state is also two-dimensional, but consists of the exponentially-filtered counts $\tilde{n}_R$ and $\tilde{n}_G$ that determine the responses of quasi-Bayesian observers (see Eq. 4.6). In each of the seven models considered in this part of the table, a deterministic estimate is computed on the basis of the values of n_R and n_G, or of $\tilde{n}_R$ and $\tilde{n}_G$, and truncated Gaussian response noise is added (as in Eq. 5.5) in order to allow the BIC to be finite.

In the first three models, we consider three classes of deterministic estimation rules. The model in line 1 corresponds to the correct Bayesian estimate, but with response error added. In line 2, Bayesian estimation is again assumed, but starting from a Beta(α, α) prior (the model simulated in the second row of Figure 10). Finally, the model in line 3 assumes a linear-in-log-odds transformation of the correct estimate (Eq. 1.3), as in line 3 of Table 1. (In Table 1, we considered the consistency of these models with subjects' average estimates in different evidentiary states; we now consider the ability of versions of these models that have been augmented with additive Gaussian response error to fit subjects' trial-by-trial responses.) No subject has their responses best-fitted by any of these full-information models (Table 5, last column).

The next four models also correspond to deterministic estimation rules, but those of quasi-Bayesian observers with a weight on the prior, ρ, allowed to be different from 1. In line 4, the prior and the weight on the likelihood are correct (i.e., $\alpha = 1$ and $\lambda = 1$). In line 5 the weight on the likelihood, λ, is allowed to differ from 1; in line 6 the prior is a Beta(α, α) distribution, with α a free parameter (and $\lambda = 1$); and in line 7 both α and λ are free parameters. The responses of this last model are based on the estimate $p_{\alpha\rho\lambda}$ (see Eq. 4.7), with the addition of truncated Gaussian noise. We denote by p_ρ, $p_{\rho\lambda}$, and $p_{\alpha\rho}$ the estimates on which the responses of the models in line 4, 5, and 6 are based, and which are equal to $p_{\alpha\rho\lambda}$ with, respectively, $\alpha = \lambda = 1$, $\alpha = 1$, and $\lambda = 1$. As for the quasi-Bayesian models with $\rho = 1$ and $\lambda \neq 1$, and with α either set to 1 or allowed to differ from 1, we have seen that they produce the same responses as the Bayesian, incorrect-prior model of line 2 (if α is replaced by $1/\lambda$ or by α/λ; see Eq. 4.5), and thus they yield the same likelihood. The model with α set to 1 has the same number of parameters as the Bayesian, incorrect-prior model, and thus the same BIC; but the BIC of the model with α as an additional parameter is necessarily higher. For these reasons, we do not show these models in Table 5.

The remaining five models considered are versions of the noisy-counting model. The model in line 8 assumes that the cognitive state is a precise, deterministic count (a linear transformation of $n_R - n_G$), but allows for truncated Gaussian response error; this is the model simulated in the third row of Figure 10. The model in line 9 instead assumes noisy counting, with a Gaussian distribution of cognitive-state updates, and no response error; this is the model simulated in the fourth row of Figure 10. The model in line 10 nests the previous two models, by allowing both noisy counting (with a Gaussian distribution of cognitive-state updates) and a Gaussian response noise. This model corresponds to Eqs. 5.4–5.5 above, with y_{t+1} drawn from a Gaussian distribution (Eq. 5.3), and with $\sigma \neq 0$ and $\nu \neq 0$.

In the last three models just presented, we have assumed Gaussian noise in the responses, in the cognitive states, or in both. However, it is evident from Fig. 9A that in subjects' responses, the adjustment is almost always positive in sign when a red ring is observed (and similarly, the adjustment is almost always negative when a green ring is observed.) While subjects' adjustments are largely independent of both their previous observations and their previous estimates, they are quite informative about the sign of the most recent ring draw;

the distributions of adjustments associated with red rings and green rings are almost entirely non-overlapping. And the degree to which the sign of the adjustment is predictable from the sign of the ring draw is somewhat greater than one would expect under the hypothesis of a Gaussian distribution of adjustment sizes.

The empirical distribution of adjustment sizes is asymmetric, and more similar to a log-normal distribution than to a normal distribution. This skewness may originate either in the stochastic updates to the internal cognitive state, or in the errors in response selection. This leads us to consider two additional models: first, a noisy-counting model in which cognitive states follow the law of motion in Eq. 5.4, where the random update, y, is log-normally distributed, i.e.,

$$\log y \ \sim \ N(\log \mu, \sigma^2), \tag{5.6}$$

again with parameters μ, σ^2 to be estimated, and we assume in addition that Gaussian noise and truncation occur at the moment of the response (Eq. 5.5). This is the model in line 11 of Table 5. Second, we consider (in line 12) a different model in which the cognitive state, r_t, again follows the law of motion of Eq. 5.4, but here with Gaussian updates y (Eq. 5.3), and with log-normal noise in the choice of a response. More precisely, given a preceding response $\hat{p}_t$ and a new cognitive state r_{t+1}, we assume that the new response is chosen following the noisy truncation rule

$$\hat{p}_{t+1} = [\hat{p}_t + (r_{t+1} - \hat{p}_t)e^{\varepsilon_t}]_{[0,1]}, \tag{5.7}$$

where ε_t is a draw from the normal distribution $N(0, \nu^2)$, and where we assume the initial condition $p_0 = 0.5$. In other words, the response $\hat{p}_t$ is adjusted, prior to truncation, with a log-normally-distributed adjustment whose median is $r_{t+1} - \hat{p}_t$. Thus we assume that subjects' slider movements are subject to a law of "scalar variability," of the kind observed in case of many tasks where a subject must reproduce or estimate a physical magnitude (Petzschner *et al.,* 2015) or a number (Whalen *et al.,* 1999; Dehaene and Marques, 2002) that they have previously been shown.

We fit each model, first with identical parameters for all the subjects, and second, with subject-specific parameters (Table 5). For all the models we obtain a better fit (lower BIC) in the latter case, and thus we focus on the results of this fit. The model-fitting results substantiate the results obtained in the previous sections. First, the models on lines 2-7, which allow the optimal Bayesian inference process to be distorted in one way or another, all fit better than the model on line 1, which departs from full optimality only by allowing for truncated Gaussian response noise. Among these seven models, the model in line 7 has the lowest BIC, by a sizable amount (more than 9000); this model corresponds to the quasi-Bayesian observer with free parameters α, ρ, and λ, whose responses we have seen are well approximated by a linear function of the net difference $n_R - n_G$ (Eq. 4.8). The model in line 8, precisely, assumes that responses are a linear function of $n_R - n_G$ (with the addition of truncated Gaussian noise), and its BIC is also lower than those of the first six models. (It is however higher than that of the quasi-Bayesian model of line 7. We note that the model in line 8 has two parameters, μ and σ, while the one in line 7 has four; presumably this enables a larger flexibility in capturing subjects' behavior.) However, these eight models with deterministic cognitive states and noise only in the response (lines 1-8) are the ones with the highest BICs in Table 5; they all have BICs that are higher than those of the four counting models with noisy, unidimensional cognitive states (lines 9-12), and higher than that of our

	Cognitive state	Noise		Response selection	Homog.	Heterog.	% subj.
(1)	$r = (n_R, n_G)$		•	$\hat{p} = [p^* + \varepsilon]_{[0,1]}$, where $p^* = \frac{n_R+1}{t+2}$	250,098	246,664	0%
(2)	$r = (n_R, n_G)$		•	$\hat{p} = [p_\alpha + \varepsilon]_{[0,1]}$, where $p_\alpha = \frac{n_R+\alpha}{t+2\alpha}$	247,107	237,050	0%
(3)	$r = (n_R, n_G)$		•	$\hat{p} = [\mathrm{Lo}^{-1}\big(a\mathrm{Lo}(p^*) + b\big) + \varepsilon]_{[0,1]}$	248,240	238,915	0%
(4)	$r = (\tilde{n}_R, \tilde{n}_G)$		•	$\hat{p} = [p_\rho + \varepsilon]_{[0,1]}$	250,099	244,710	0%
(5)	$r = (\tilde{n}_R, \tilde{n}_G)$		•	$\hat{p} = [p_{\rho\lambda} + \varepsilon]_{[0,1]}$	246,049	234,592	0%
(6)	$r = (\tilde{n}_R, \tilde{n}_G)$		•	$\hat{p} = [p_{\alpha\rho} + \varepsilon]_{[0,1]}$	246,950	235,943	0%
(7)	$r = (\tilde{n}_R, \tilde{n}_G)$		•	$\hat{p} = [p_{\alpha\rho\lambda} + \varepsilon]_{[0,1]}$	243,370	224,613	4.8%
(8)	$r_{t+1} = r_t + x_{t+1}\mu$		•	$\hat{p} = [r + \varepsilon]_{[0,1]}$	243,359	227,242	0%
(9)	$r_{t+1} = [r_t + x_{t+1}\mu + \xi]_{[0,1]}$	•		$\hat{p} = r$	232,098	220,982	0%
(10)	$r_{t+1} = r_t + x_{t+1}\mu + \xi$	•	•	$\hat{p} = [r + \varepsilon]_{[0,1]}$	231,602	219,528	14.3%
(11)	$r_{t+1} = r_t + x_{t+1}\mu e^\xi$	•	•	$\hat{p} = [r + \varepsilon]_{[0,1]}$	230,909	217,279	33.3%
(12)	$r_{t+1} = r_t + x_{t+1}\mu + \xi$	•	•	$\hat{p}_{t+1} = [\hat{p}_t + (r_{t+1} - \hat{p}_t)e^\varepsilon]_{[0,1]}$	229,305	209,676	47.6%

Table 5: Comparison of the fit of alternative stochastic models. Mathematical formulations of the models, and their BICs. n_R and n_G denote the number of red and green rings presented, and t the total number of rings. $\tilde{n}_R$ and $\tilde{n}_G$ denote the exponentially-filtered counts with weight ρ (Eq. 4.6). $p_{\alpha\rho\lambda}$ is the estimate of the quasi-Bayesian model with free parameters α, ρ, and λ (Eq. 4.7), while p_ρ, $p_{\rho\lambda}$, and $p_{\alpha\rho}$ correspond to the cases in which, respectively, $\alpha = \lambda = 1$, $\alpha = 1$, and $\lambda = 1$. ξ and ε are two mean-zero Gaussian random variables (thus e^ξ and e^ε are two log-normal random variables with median 1). For models 8-12, $r_0 = 0.5$, and for model 12, $p_0 = 0.5$. The "Noise" column indicates whether each model includes noise in the cognitive states (left dots) and in the response selection (right dots). Models are fitted either by requiring the same parameters for all the subjects ("Homog." column) or by allowing for different parameters for each subject ("Heterog." column). The last column shows the proportion of subjects whose responses are best-fitted by the model in the row. The horizontal lines marks the separation between the models that do not have noise in the cognitive state (above the line) and those that do (below the line).

best-fitting model (line 12) by more than 14,900. This indicates very substantially worse fit; even in the case of the best of these models (line 7, with subject-specific parameters), the Bayes factor in favor of the noisy-counting model is larger than 10^{3243}.

Second, the model in which we assume that the response is a noisy report based on a cognitive state that counts the net number of red rings precisely (line 8) results in a higher BIC (by more than 6000) than the model that assumes truncated Gaussian noise in the counting process, but no response noise (line 9); the latter thus yields a better fit of the data. This supports the conclusions of the visual comparison of the predictions of these two models in rows 3 and 4 of Figure 10. Allowing for both noisy counting and response noise (line 10) results in an even lower BIC (despite the penalty for the additional free parameter); but it is clear that the noise in the evolution of the cognitive state is the more important of the two types of noise to include.

Finally, the two best-fitting models assume both noisy counting and noisy responses, but with an asymmetric distribution for either of these two types of noise. The responses of 81% of our subjects are best-fitted by one of these two models. The one with log-normal noise in responses (line 12) yields a lower BIC than the model with log-normal updates of cognitive states (line 11). We note that the large difference in BIC between these two models (more than 7,600) is mainly driven by the responses of two subjects. These two subjects have a strong propensity to repeat their responses, in successive trials: more than 40% of their responses are repetitions (while the other subjects have a median of 2.3% of repetitions). These frequent repetitions are well captured by a model of log-normal responses (Eq. 5.7), when the noise parameter, ν^2, is large: in this case, almost half of the adjustments result in repetitions, because they are smaller than the resolution of the response scale (0.1%, on a scale that ranges from 0% to 100%).

We show, in the last row of Fig. 10, the behavior of the best-fitting model (with log-normal noise in responses): it reproduces the behavioral patterns that we have identified in subjects' responses. The second-best model (with log-normal updates of the cognitive states and truncated Gaussian response noise,) and the third-best model (in which both types of noise are Gaussian) also reproduce these patterns (see Methods). Overall, the responses of nearly half of our subjects are best-fitted by the model with log-normal noise in the responses (and Gaussian updates of the cognitive states), while one third is best-fitted by the model with log-normal updates of the cognitive states (and Gaussian noise in responses; Table 5, last column). As 14.3% of subjects are best-fitted by the model in which the two types of noise are Gaussian, a total of 95.2% of subjects are best-fitted by models in which there is noise in the evolution of the cognitive states *and* in the selection of a response. Moreover, if we allow these subjects to have only one type of noise (either in the updates of the cognitive states, or in the response selection; i.e., if we only allow for the models in line 8 and 9 of Table 5), we find that 90% are better fitted by the model with noise in the cognitive states.

In summary, our results point to an inference mechanism that is not Bayesian, but not either approximately Bayesian; and which relies instead on a unidimensional cognitive state that is updated (increased or decreased) upon each observation of a new ring, and which determines the response (no subject is best fitted by any of the full-information models); in addition, noise occurs both in the cognitive-state updates and in the selection of a response, but the former accounts for a larger share of the response variability. This simple mechanism accounts for the major behavioral patterns found in our data: first, the Bayesian properties that we have identified are not verified; second, the updates of the unidimensional cognitive states account for the reversal of the conservatism bias; and third, the presence of noise in the cognitive-state updates produces autocorrelation in successive responses.

6 Discussion

6.1 Summary

We have documented patterns of both bias and variability in the probability estimates of our subjects, that are fairly consistent across subjects, and that indicate not only that they fail to accurately produce the correct Bayesian estimate, but that their estimates depart predictably from the Bayesian benchmark on average. The subjects' average responses do not seem to be monotonic transformations of the optimal Bayesian responses, and we reject both the hypothesis that responses result from a process of Bayesian updating that optimally takes into account the evidence but on the basis of an erroneous prior, and the hypothesis that subjects misweight the prior and the likelihood in their application of Bayes' rule; the combination of these two hypotheses is rejected as well (Fig. 3, Table 1). We also reject two hypotheses under which subjects conduct Bayesian inference, but on the basis of erroneous beliefs: a belief in sudden changes of the underlying probability, and a belief in sequential dependency in the ring draws. Finally, we consider, and reject, the hypothesis of random shocks in the prior occurring at the beginning of each block, and a hierarchical Bayesian model with learning across blocks (see Methods).

The responses of a Bayesian observer who would have only access to imprecise (noisy) representations of the decision situations, in our inference task, would verify two properties implied by Bayesian inference (the 'calibrated-responses' property, Eq. 3.2, and the 'consistent-updates' property, Eq. 3.6). In contrast to many studies of Bayesian models of human cognition, which usually aim at showing that given some evidence presented, subjects' responses are approximately consistent with those of a Bayesian observer, we note that these two properties are a kind of implication of Bayesian inference that imposes restrictions on the distributions of responses of the Bayesian observer, across various evidentiary states. We find that these properties are not verified in behavioral data, and subjects' responses are in fact markedly different from them (Figs. 7-8, Table 3). In short, the behavior of subjects is not compatible with Bayesian inference.

Moreover, we find that subjects' estimates under-react to the evidence presented (conservatism) in the first few ring draws of a new regime, but then eventually over-react to the evidence presented, once a sufficient number of rings have been observed (Fig. 4). Our subjects exhibit, in addition, autocorrelation in their responses, i.e., for a given sequence of rings, a large response at some trial is likely to be followed by a large response at subsequent trials (Fig. 5), which suggests that the noise in the subjects' representation of the current situation propagates through successive trials. What is more, the responses of subjects appear not to be based on two statistics obtained from the observed evidence, as would be necessary to provide an optimal response, but instead seem to derive from a unidimensional cognitive state that imperfectly reflects the sequence of evidence (Fig. 6). Indeed, the size of the adjustments chosen by the subjects seem to be relatively insensitive to the total amount of evidence presented, although an optimal observer would adopt smaller adjustments as more evidence is accumulated (Table 4). In fact, the adjustments do not either vary strongly with the preceding response, in contradiction with optimal behavior (Fig. 9; this also explains why the consistent-updates Bayesian property is not verified in subjects' data). In sum, the adjustments seem to depend mostly on only one kind of information: the color of the ring drawn. If it is red, the subjects adjust the slider to the right by some distance, and if it is green they adjust it to the left, by roughly the same distance.

These various patterns are consistent with a "noisy-counting" model of the way in which running probability estimates are adjusted in response to the sequential arrival of evidence. We show that assuming that subjects keep an imprecise running count of the net number of red over green rings that is updated, with noise, on each presentation of a new ring, is sufficient to qualitatively reproduce the main patterns identified in data, such as the conservatism after short sequences of evidence and its reversal after longer sequences, the autocorrelation in responses, the violations of the identified Bayesian properties, and the independence of adjustments from the length of the sequence and from the preceding response (Fig. 10).

Finally, model fitting favors noisy-counting models, with noise in both the cognitive state and in response selection, with a degree of skewness in either of the types of noise (Table 5).

We have also run a variant of the experiment in which the length of each block of trials is a geometrically-distributed random variable (instead of being fixed at five draws). We present these additional results in Methods. The subjects in this variant exhibit similar behavioral patterns than in the original experiment, thus substantiating our results. One difference, however, is that in this variant the subjects seem to make smaller adjustments in their responses (another finding that is not accounted for by traditional models of Bayesian inference). Further below, we discuss how this is in fact consistent with an efficient adaptation of the decision-making process to the statistics of the sequence lengths.

6.2 In what ways are subjects Bayesian?

Our finding that human behavior is not consistent with Bayesian inference seems at odds with the literature on inference from sequential data, in which models of the optimal, Bayesian observer have been shown to capture several aspects of human behavior (Wilson *et al.,* 2013; Khaw *et al.,* 2017a; Prat-Carrabin *et al.,* 2021). In particular, the result of Khaw *et al.* (2017a) reproduced in Fig. 1B (and obtained in a probability-estimation task,) suggests that the average response of subjects equals the Bayesian estimate; but in our study this hypothesis is significantly rejected for all our subjects (Fig. 3 and Table 1, line 1). An important difference, however, between these results is that in the analysis of Khaw *et al.* (2017a) the responses of subjects are averaged over groups of *different* sequences of observations that all yield identical or close-to-identical Bayesian estimates. This 'pooling' masks the diversity of the evidentiary states in which subjects are asked to make a decision. By contrast, the restricted number of possible sequences, in our task design, allows to examine the distribution of subjects' responses in *each* separate evidentiary state, and thus to exhibit that their average differs from the Bayesian estimate, in most cases (Fig. 3). Hence, while a natural interpretation of the results in Fig. 1 is that human subjects are well approximated by the Bayesian observer once one averages out the imprecision in responses, these findings suggest that subjects' responses appear Bayesian when averaged instead over the different possible sets of presented evidence. Our results show that at least in the case of short sequences of evidence, the average response, when looking at one specific sequence, is not Bayesian.

This raises the questions so as to why human behavior should be suboptimal in response to each one of many short sequences of evidence (as in our study), but close to optimal when averaged over many long sequences of evidence (as in the studies of Gallistel *et al.,* 2014 and Khaw *et al.,* 2017a); and as to what mechanism may give rise to such a behavioral pattern. A conjecture is that one's environment typically changes over time (and in fact it did in the experiments of Gallistel *et al.,* 2014 and Khaw *et al.,* 2017a), such that one usually faces a series of inference problems that are roughly identical, but with different underlying parameters. If the brain is subject to cognitive limitations that prevent it from forming the optimal response to one given set of evidence, it may be advantageous, in such a changing environment, to optimize the average response to many different sets of evidence. Another possibility is that after long sequences of evidence, a given inferential mechanism is used, while when no or very little evidence has been observed, human inference relies on a different mechanism, perhaps because of the novelty of the situation. In any case, these questions call for a closer examination of human inference in the context of long sequences, so as to more clearly appreciate whether (or when) human responses to sequences of evidence are close to Bayesian optimality.

6.3 Conservatism and over-reaction as byproducts of limited attention

One well-known discrepancy from Bayesian optimality is conservatism, i.e., the insufficient reaction of human inferences to the evidence presented, as compared to the prescriptions of Bayes' rule. First documented in the 1960s (Phillips *et al.*, 1966; Phillips and Edwards, 1966), conservatism has since then been the object of a considerable number of studies (see Benjamin (2019) for a review). We uncover a more subtle pattern: although we do find significant conservatism in the responses of subjects after the presentation of up to three rings, this behavior is reversed after longer sequences, and subjects then excessively react to the evidence, in comparison to the Bayesian observer (Fig. 4). This finding is all the more surprising as a meta-analysis of the evidence on belief-updating found that conservatism was "more severe the larger the sample size" (Benjamin, 2019), although this was obtained in the context of simultaneous presentations of the pieces of evidence, and with 'bookbag-and-poker-chips' tasks, in which subjects have to estimate the probability that one of two bags is the one being drawn from — which is the kind of tasks used in the vast majority of studies that report conservatism.

These tasks present similarities and differences with ours. In both paradigms, subjects are presented with instructions that clearly imply, in formal terms, the prior probabilities of several hypotheses, and a likelihood function (i.e., the probability of data, under each hypothesis); they are then presented with data, and asked about their posterior beliefs. The experiments differ in that in the typical bookbag-and-poker-chips task, there are only two hypotheses (e.g., either 30% of red chips, or 70% of red chips), while in our task there is a continuum of hypotheses (all the proportions between 0% and 100%). The latter may be more ecologically relevant (e.g., when estimating the probability that it will rain, there is usually no reason to restrict the support of one's prior to just two values). Furthermore, in bookbag-and-poker-chips tasks, subjects are asked for the posterior probability of one of the two possible hypotheses, while in our task subjects are asked for a point in the continuum of hypotheses (their estimate of the proportion), and not for a posterior probability. This could be an important difference, as subjects might find it more natural to report a 'concrete' hypothesis (a proportion of red rings), which could actually have generated the observed data, than to report an abstract, Bayesian probability. In any case, the reward function in our task implies that the response provided should be the mean of the posterior: thus in short, we ask for the first moment of the posterior, while in bookbag-and-poker-chip experiments subjects are asked for the posterior itself. In the inference task studied by Phillips, Hays, and Edwards (1966), there are four hypotheses (instead of just two), and subjects are asked to report the posterior probabilities of the four of them, i.e., the full posterior. Here also, the authors find conservatism in subjects' responses. The conservatism in the updates of the posterior mean, which we obtain for $t \leq 3$, could be interpreted as a natural corollary of the conservatism in the updates of the posterior, exhibited in these studies, insofar as the latter further extends to a continuum of hypotheses.

Reports of over-reaction to evidence, as we find, are relatively rare in the literature, although it is already visible in the data reported by Peterson and Miller (1965) and Phillips and Edwards (1966). In Brown and Bane (1975), subjects' estimates of a probability are more extreme than the Bayesian estimates, but in this task the true probability increases as more samples are drawn (when it does not increase, the authors find conservatism). In economics, expectations about macroeconomic and financial quantities (e.g., stock returns, interest rates, etc.) have been shown to overreact to new information (Bordalo *et al.,* 2020; Afrouzi *et al.,* 2020). Benjamin (2019) points to several studies in which over-reaction to evidence has been found, with bookbag-and-poker-chips tasks or similar tasks in which subjects are asked to infer the probability of a hidden state that can take only two values (Peterson and Miller, 1965; Phillips and Edwards, 1966; Donnell and Du Charme, 1975; Griffin and Tversky, 1992). A common finding in these studies is that if the distribution of the evidence conditional on one state (the likelihood, or 'diagnosticity') is very different from the distribution conditional on the other state (such that the evidence is very informative and distinguishing the two states is easy), then subjects tend to underreact to the evidence, as compared to the Bayesian observer (i.e., they exhibit conservatism); while if the likelihoods are close (such that the evidence is not very informative and distinguishing the two states is difficult), then subjects tend to overreact to the evidence. Augenblick, Lazarus, and Thaler (2023) closely examine this point and consistently find, in experimental and empirical evidence, overreaction to weak signals and underreaction to strong signals.

We note that in the former case (if the likelihoods are dissimilar), then a new piece of evidence brings a lot of information, thus the Bayesian posterior is very different from the prior, and the update of the Bayesian estimate is large. In the latter case (if the likelihoods are close), then a new piece of evidence brings little information, and the Bayesian update is small. In our task, the first ring drawn brings a lot of information: the Bayesian observer thus adjusts its estimate by a large amount (from 1/2 to 1/3 or 2/3, i.e., an adjustment of 16.67%). The last drawn ring brings less information, and consequently the update of the Bayesian observer is smaller (5.5% on average). In both cases, subjects adjust their estimates by about 8%. Thus subjects seem to underreact to the evidence when it brings a lot of information, and overreact to the evidence when it brings little information — similarly to the studies mentioned above.

Our account of these effects, however, is very different from the explanation extensively studied in the literature, according to which people, when updating their beliefs, misweight the likelihood (as in the original proposal of Phillips and Edwards, 1966), or misweight the prior, or both, amounting to a distortion of the optimal Bayesian procedure that we have dubbed 'quasi-Bayesian inference' — and which does not provide a satisfying explanation of the behavioral data. The account we propose is not grounded on the principles of Bayesian inference. We show that the subjects seem to pay attention to the color of the new ring, and to choose their adjustment on the basis of that information alone, thus neglecting many pieces of information that would be useful to a Bayesian observer, such as the total number of rings observed, the number of red rings among them, and even their current estimate of the probability.

In spite of this great frugality in the attention paid to relevant information, subjects perform reasonably well in the task (a point we further discuss below). What our results suggest is that one should be careful in modeling human inference (and probably other decisions), as subjects may use only an unexpectedly small fraction of the information presented to them. A typical modeling approach founded on the Bayesian paradigm starts with the ideal information-processing procedure that makes optimal use of all the available information; the resulting behavior is then compared to experimental data; and finally some of the procedure's assumptions are relaxed in order to accommodate for the suboptimal patterns found in data. It would perhaps be fruitful to adopt a different approach, and take instead as a starting point a total lack of information of the decision-maker about the decision situation, and then ask what piece of information would be useful in reaching a decision. The rationale for this approach is that subjects seem to considerably economize on the information upon which they base their decisions. At least in our task, subjects' decisions about how they will move the slider seem to depend mostly on just one binary variable, the color of the new ring.

Presumably, paying attention to more variables would necessitate more cognitive effort. Human decision-makers have been shown to be sensitive, indeed, to the cognitive demands associated with different tasks, and tend to prefer the less demanding ones, even if it decreases their rewards (Kool *et al.,* 2010; Westbrook *et al.,* 2013). This suggests the existence of a trade-off between rewards and cognitive costs, as formalized in the theory of 'expected value of control' (Shenhav *et al.,* 2013). How to best characterize the cost of acquiring actionable information remains uncertain, and all the more so in the context of our task, as paying attention to the color of the ring (displayed on screen) presumably involves mechanisms of a different nature than those allowing to remember some information about the sequence of preceding rings (which are not displayed on screen anymore).

The theory of 'rational inattention' puts forth, as a candidate formalized cost, the mutual information between the relevant variable and the signal obtained about it (Sims, 2003; Maćkowiak *et al.,* 2020). This framework is typically used to model the processing of information that is external to the observer, although Azeredo da Silveira *et al.* (2020) use it in a model of imprecise recall from memory. A different cognitive constraint appears in some encoding-decoding models of perception, which posit a cost (or a bound) proportional to a measure of the encoding capacity of the perceptual system, resulting in imprecise representations (Ganguli and Simoncelli, 2010; Morais and Pillow, 2018; Prat-Carrabin and Woodford, 2021). In any case, we surmise that changing the presentation of the evidence (e.g., leaving on screen the sequence of past rings) would change the cognitive cost of paying attention to it, and result in different behavioral patterns. We leave, however, for future studies a fuller theoretical and experimental investigation of the mechanisms by which subjects neglect or pay attention to the information presented to them.

6.4 Implications of the imprecision in cognitive states

An inference task with short sequences of binary stimuli, very similar to ours, was conducted by Shanteau (1970), who also concluded that the Bayesian benchmark was inadequate in reproducing human behavior. He considered, in addition, a model based on information integration theory (Anderson, 1991). This theory proposes that in sequential decision-making problems, new information is integrated in an existing decision state, although this integration does not have to be constrained by Bayes' rule. The resulting 'additive' decision model considered by Shanteau (1970) is very similar to the version of our counting model in which no noise perturbs the cognitive state, i.e., in which the model subject maintains a deterministic, precise count of the net excess of red rings over green rings. However, both the model-selection procedure and the qualitative examination of subjects' and models' behavioral patterns, in our task, suggest that the presence of noise in the cognitive states is crucial in reproducing subjects' behavior. Noise in response selection alone does not capture, in particular, the autocorrelation in responses, and if only one type of noise is allowed then the responses of 90% of subjects are best-fitted by the model that features noise in cognitive states.

The autocorrelation in responses is not predicted by the model in Shanteau (1970), nor is it predicted by many models of sequential inference in the literature, in which beliefs are, often implicitly, deterministic functions of the presented evidence, and noise is sometimes added in the selection of a response. In the model of Gallistel *et al.* (2014), each response is a function of the preceding response and of the new evidence, and thus ultimately it is a function of all the evidence presented. Nassar *et al.* (2010) and Wilson *et al.* (2013) model belief updating in a changing environment using 'delta-rule' approximations to Bayesian inference that are deterministic. In Khaw *et al.* (2017a), a model of 'rational inattention' describes at each trial the response distribution, which itself depends deterministically on the evidence history.

Several recent studies, however, emphasize the role of computational imprecision in behavioral variability (Renart and Machens, 2014; Wyart and Koechlin, 2016; Findling *et al,* 2019, 2021). Hilbert (2012) shows how a cognitive model based on an information channel with noisy memory accounts for eight deviations from optimality commonly found in human decisions. In a decision task requiring the accumulation of sensory evidence, Drugowitsch *et al.* (2016) find that a dominant fraction of choice suboptimality results from random fluctuations in inference, while only a minimal fraction originates in sensory and response-selection noise. In Azeredo da Silveira *et al.* (2020), over-reaction to new information is accounted for by imprecision in the memory of past observations. In a sequential inference task with binary states, Glaze *et al.* (2018) find that a 'stochastic learning algorithm' outperforms other (deterministic) models, and in another inference task, Prat-Carrabin *et al.* (2021) show that subjects' behavior is better captured by models in which an approximation of the posterior is stochastically updated upon each new observation, than by various models of noisy response selection. We note that such models of stochastic inference, by nature, yield autocorrelation in responses, which decrease with the distance between the presentations of two pieces of evidence, as observed in our subjects' behavior (Fig. 5).

Because it propagates over successive observations of evidence, the noise in cognitive states is a source of response variability that is qualitatively different than response-selection noise. It implies that the cognitive state (and thus the belief) at a given time is contingent not only on all the evidence observed up until that time, but also on the idiosyncrasies of the preceding cognitive state. In sequential inference tasks, for a given sequence of evidence, cognitive states may thus take many different 'paths'; and one cannot understand a subject's belief at a given time just by looking at the sequence of evidence that this subject has observed. In other words, the sequence observed represents only a minor fraction of the information content of a subject's response, while the subject's preceding response represents a major fraction (Table 2). Similarly, in perceptual identification tasks, subjects' responses have been shown to be influenced by their preceding responses, when no feedback is provided (Ward and Lockhead, 1971; Mori and Ward, 1995), a behavior reproduced by the relative judgment model of Stewart, Brown, and Chater (2005). (We note in addition that this last model is one in which subjects are sensitive to the difference in successive stimuli, rather than to their absolute magnitudes).

In the context of inference tasks, a consequence of noisy cognitive states is that when several pieces of evidence are presented simultaneously, the resulting cognitive state depends (possibly stochastically) on the initial state and on the whole set of evidence, while when the evidence is presented sequentially, it is the result of a succession of noisy updates of the cognitive state. This may provide an explanation to the fact that human subjects make different inferences in these two cases (Shanteau, 1970; Benjamin, 2019). The closer examination of how human inference depends on the sequential or simultaneous presentation of evidence may shed light on the role of noise in the updating of beliefs; we leave this investigation to future work.

6.5 Subjects' behavior is adapted to their limited attention

Our noisy-counting model might seem to show that people use a heuristic that reflects an incorrect understanding of the rules of probability, or even misunderstanding of the nature of the task in our experiment. Instead, we find that the most notable feature of the cognitive process that subjects appear to use is the degree to which it allows them to perform the task (producing the required judgments, in a sequential fashion) while requiring very little information about the specific situation in which each new judgment is selected. Subjects appear to approach the task not by considering afresh on each trial which slider position represents an appropriate estimate given the cumulative evidence revealed to that point, but instead by considering what size of change in the slider position is appropriate in light of the new evidence observed since their last response. This adjustment decision is made in a way that takes little account of either the previously observed evidence or their previous responses (including the existing location of the slider); it depends mostly on the color of the latest ring draw. Finally, while it seems that subjects pay attention to the amount by which they change their reported estimate (i.e., how much they move the slider) on each trial, they do not exert close control of this, so that there is considerable trial-to-trial random variation in the exact size of the adjustment (which our model treats as pure response error), though our model assumes perfect control of the direction of adjustment.

Subject to these limits on the degree of attention paid to the specific situation and the degree of control exercised over the subject's precise response, the subjects' responses are reasonably well-adapted to the task and its reward structure. The decision to approach the task as one of decision on a direction and size of slider adjustment, rather than making a fresh decision about where to place the slider after each new ring draw, is in fact adaptive, assuming the above constraints on the information that is to be used in action selection. If the subject were instead to make a fresh decision about slider placement on each trial, but subject to the informational constraints just summarized, they would select an estimate $\hat{p}_t$ on each trial as an independent draw from a distribution of possible slider positions associated with the ring draw x_t. Thus there would be only two possible distributions from which the estimate would be randomly drawn on any trial, and the distribution of possible estimates would depend on the current ring draw in the same way for all values of t. Choosing a change based on the same limited information instead allows the subject's estimate to be highly correlated with the cumulative excess number of red rings, $n_R - n_G$, rather than being correlated only with the most recent ring draw, and this provides a better approximation to the optimal estimates across sequences of differing length.

The average size of subjects' adjustments in response to each ring draw is also somewhat reasonable. Suppose that we take as given that the update of the cognitive state will be chosen from a Gaussian distribution with variance σ^2, independently of any information about ring draws prior to the current one or the subject's previous estimates, as well as the fact that given the subject's choice of a direction of adjustment, the size of the adjustment, prior to truncation, will necessarily be drawn from a log-normal distribution (Eq. 5.7), where the value of ν (reflecting the subject's degree of control over their action) is also taken as given, regardless of the subject's choice of the mean size of the cognitive-state update, μ. Under these constraints, the only aspects of a subject's response behavior that remain to be chosen are the sign of the adjustment following each perceived ring draw, and the value of μ. With regard to these elements of the decision rule, it is clearly optimal for the adjustment to have the same sign as x_{t+1}.

The parameter μ in turn will have an optimal value that trades off the considerations that smaller values of μ will imply under-reaction for small values of t (when an additional ring draw provides a great deal of additional information) while larger values will imply over-reaction for larger values of t (when an additional ring draw should not change one's beliefs much). One should therefore expect, if μ is chosen optimally, to observe under-reaction for small values of t, together with over-reaction for larger values of t, as we do. And finally, under our log-normal model of the distribution of responses, there is a greater risk of adjustments being considerably larger than desired than of their being smaller than desired to the same extent; this makes it optimal to aim for a smaller mean size of the cognitive-state update, μ, the larger is ν; this is also a factor that makes it optimal for μ to be chosen more "conservatively."

The blue line in Fig. 11 illustrates how the predicted mean squared error (MSE) of a subject's estimates varies with different possible choices for the mean update size μ, taking as given values for σ and ν. (Here these parameters are set at the medians of the subject-specific parameter values obtained from maximum-likelihood estimation of the noisy-counting model described on line 12 of Table 5. Further details of the simulations used to produce this figure are given in Methods.) There is a clear interior minimum; for most subjects, this is when the mean size of the cognitive-state update, μ, reaches a value slightly below 0.1. Compared to these optimal values, the average sizes of the adjustments chosen by subjects are seen to be of roughly the right magnitude, and allow them to keep a mean error close to the minimum (Fig. 11, blue boxplot). Hence, although there is imprecision in the way that subjects update their cognitive states upon the observation of a new ring, along with an imperfect control in their choice of a response, the mean size of their adjustments seems appropriate, and close to the value that would maximize the accuracy of their responses (and thus their reward in the task).

Because a particular choice of μ typically results in under-reaction to the evidence up to some sequence length t, and over-reaction for longer sequences, the optimal choice of μ depends on the distribution of sequence lengths that one can expect to experience. In a

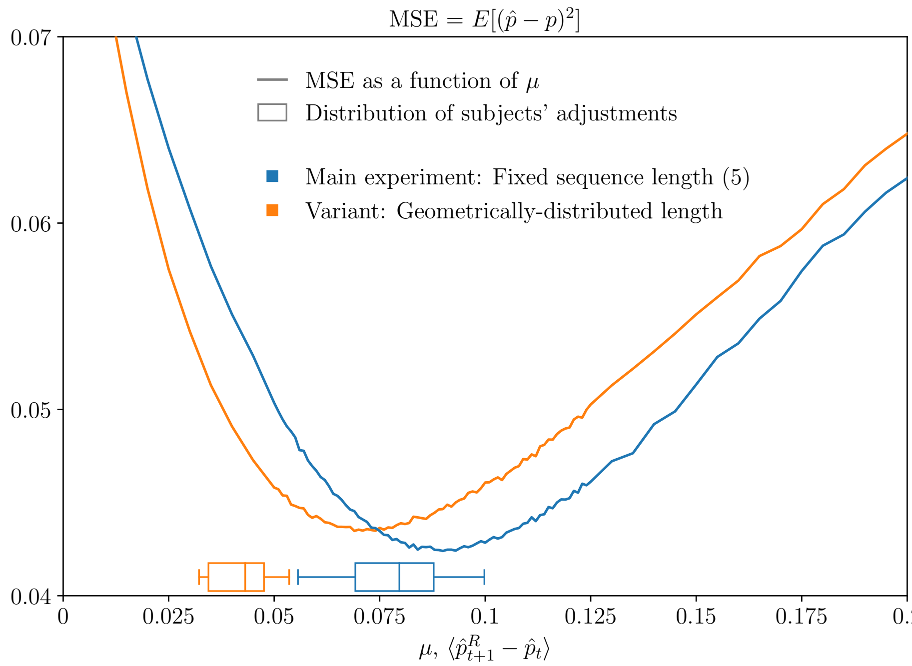

Figure 11: Subjects' mean update sizes are close to optimal, given the limited information upon which updates are conditioned, and the noise in cognitive states and in response selection. Here we consider how different values of μ would affect the mean squared error (MSE) of a model subject's probability estimates, for given values of ν and σ, assuming the model of noisy counting specified by Eqs. 5.4, 5.3, and 5.7 (Gaussian updates of the cognitive states and truncated log-normal response noise). *Lines:* MSE in simulations of the model for alternative values of the mean size of the cognitive-state update, μ, when $\sigma = 0.0344$ and $\nu = 0.194$, the median estimated values of these parameters across our subjects. *Boxplots:* Across-subjects distribution of average adjustments (vertical line: median, box: first and third quartiles, whiskers: 10th and 90th percentiles). *Blue:* Main experiment, with sequences of five draws. *Orange:* Variant experiment, with geometrically-distributed sequence lengths.

variant of the experiment that we run, the lengths of the sequences presented are not fixed to five draws, but are instead random, and geometrically-distributed. The average sequence length is here also equal to five, but sometimes the subjects see only one or two draws, for instance (in which cases larger adjustments are appropriate), and sometimes they see ten or more (in which cases smaller adjustments are better). With these different statistics, the MSE of a noisy-counting subject with the same values of σ and ν as above reaches its minimum for a different, and smaller, value of μ (Fig. 11, orange line). The subjects, consistently, make smaller adjustments in this variant of the experiment (Fig. 11, orange boxplot; see also Methods). This strengthens the hypothesis that the subjects' average adjustment size is not an arbitrary choice, but is in fact an efficient adaptation, both to the statistics of the task and to the imprecision in their decision-making process.

Methods

Behavioral Task and Subjects

The computer-based task was developed and run using Python and oTree (Chen *et al.*, 2016). After reading the instructions, the participants faced eight practice blocks of five trials (five ring draws with the same underlying probability, which was sampled from the uniform distribution on [0,1]). Feedback was provided after each practice block: the correct probability, their five responses, and the number of points obtained in each trial were shown to the subjects. The points obtained in the practice trials did not count towards the final payoff. After the practice trials, the participants faced 200 'real' blocks, also of five trials, which counted for the final payoff. In each block, the probability was sampled from the uniform distribution on [0,1]. The end of each block and the beginning of the following one were explicitly notified to the subject. No feedback was provided at the end of each 'real' block. Each estimate provided by a subject augmented her total score by $6.5 - 85(\hat{p} - p)^2$, where $\hat{p}$ is the estimate and p the true probability. At the end of the experiment, subjects received 1 cent (USD) per point accumulated, with a minimum of $10.

Twenty-five subjects, aged 19 to 50 (average 25.3), were recruited using ORSEE (Greiner, 2015) and participated in the experiment. The sample size was determined so as to be comparable to that used in similar experiments (Gallistel *et al.*, 2014; Khaw *et al.*, 2017a). All subjects gave informed consent. The experimental protocol was approved by Columbia University's Institutional Review Board (IRB; protocol number: AAAQ2255). Four subjects performed significantly less well than the other subjects. Their average absolute error, $|\hat{p}-p|$, was .263 (standard deviation: .0298), while the average absolute error of the other 21 subjects was .176 (standard deviation: .0132). We excluded these subjects from our analyses because of this difference of more than 6 standard deviations. Thus the responses of a total of 21 subjects were included in the analyses. We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Bias from Bayesian estimate: additional tests

Table 6 shows the results of t-tests of equality between the average of the responses of subjects and the optimal estimate. In most cases, the equality is rejected. Table 7 shows the results of the t-tests when pooling by the numbers of red and green rings in the sequences.

Mutual information: alternative method of estimation

Table 8 provides the result of the decomposition of the information contained in a subject's response, using a different estimation method than the one used in the main text. The results are similar and suggest the same conclusions.

Intervals I_i of responses

In the computer-based inference task, the resolution of the response scale allows for a precision of one decimal digit, where responses are expressed as percentages (e.g., "72.4%"; see

	p^*	$\hat{p}$	p-value	% < .01	median p-v.
1/7	.143	.119	4.6e-10	.62	.00017
1/6	.167	.187	3.1e-08	.67	2.7e-10
1/5	.200	.263	3.9e-76	.71	5.1e-12
1/4	.250	.339	2.2e-214	.86	1e-18
2/7	.286	.240	1.7e-39	.81	4.9e-07
1/3	.333	.391	1.5e-204	.90	4.7e-18
2/5	.400	.402	.51	.52	.0016
3/7	.429	.405	8e-12	.38	.029
1/2	.500	.498	.15	.14	.37
4/7	.571	.598	1.3e-13	.38	.023
3/5	.600	.588	7.1e-06	.52	.00014
2/3	.667	.610	4e-200	.95	1.1e-19
5/7	.714	.761	1.2e-40	.76	6e-06
3/4	.750	.663	1e-193	.81	6.2e-25
4/5	.800	.747	1.4e-57	.81	1.6e-11
5/6	.833	.824	.015	.71	9.1e-10
6/7	.857	.895	2.8e-29	.81	7.2e-08

Table 6: Test of equality of subjects' responses and optimal estimates, for each optimal estimate. Bayesian estimates (first two columns), averages of subjects' responses (third column), and p-values of the t-tests of equality between the two (fourth column). Last two columns: proportions of subjects for whom the p-value is below 0.01, and median p-value across subjects.

Fig. 2). This results in 1001 possible responses. In our analyses, we split this set of possible responses in 41 disjoints intervals I_i. Two intervals, one at each end of the response scale, contain 13 responses: $\{0, .001, \ldots, .011, .012\}$ and $\{.988, .989, \ldots, .999, 1\}$, and 39 intervals contain 25 responses, each centered on a multiple of .025, e.g., $\{.013, .014, \ldots, .036, .037\}$.

Bayesian inference with alternative structural assumptions

We consider the possibility that subjects undertake sound Bayesian inference, but on the basis of erroneous beliefs about the process underlying the observations. We examine two hypotheses: that subjects assume that the proportion of red rings suddenly changes from time to time, and that subjects believe that the probability of drawing a red ring depends on whether a green or a red ring was drawn in the preceding trial.

Belief in sudden changes in the probability parameter

In each trial of the task, the drawn ring is replaced in the box after its presentation to the subject. Therefore, within each block of five trials, the proportion of red rings in the box does not change. Nevertheless, here we consider the possibility that the subjects believe that the proportion of red rings undergoes random changes at unannounced times. Such beliefs in 'non-stationarity' have been proposed by Yu and Cohen (2008) as an account of sequential

Sequence				Pooled subjects			Individual tests	
t	n_R	n_G	p^*	$\hat{p}$	$\hat{p}-p^*$	p-value	% < .01	median p-v.
1	0	1	.333	.420	+.087	0	.95	1.5e-41
1	1	0	.667	.582	−.085	0	1	3.7e-54
2	0	2	.25	.339	+.089	2.2e-214	.86	1e-18
2	1	1	.5	.497	−.003	0.079	.05	.34
2	2	0	.75	.663	−.087	1e-193	.81	6.2e-25
3	0	3	.2	.263	+.063	3.9e-76	.71	5.1e-12
3	1	2	.4	.402	+.002	0.51	.52	.0016
3	2	1	.6	.588	−.012	7.1e-06	.52	.00014
3	3	0	.8	.747	−.053	1.4e-57	.81	1.6e-11
4	0	4	.167	.187	+.020	3.1e-08	.67	2.7e-10
4	1	3	.333	.318	−.015	2.8e-06	.57	.0043
4	2	2	.5	.500	−.000	0.91	.05	.57
4	3	1	.667	.678	+.011	0.00039	.67	.00077
4	4	0	.833	.824	−.009	0.015	.71	9.1e-10
5	0	5	.143	.119	−.024	4.6e-10	.62	.00017
5	1	4	.286	.240	−.046	1.7e-39	.81	4.9e-07
5	2	3	.429	.405	−.024	8e-12	.38	.029
5	3	2	.571	.598	+.027	1.3e-13	.38	.023
5	4	1	.714	.761	+.047	1.2e-40	.76	6e-06
5	5	0	.857	.895	+.038	2.8e-29	.81	7.2e-08

Table 7: Test of equality of subjects' responses and optimal estimates, for each pair (n_R, n_G) of numbers of red and green rings. One-sample t-tests of equality between the subjects' average estimates when presented with n_R red rings and n_G green rings, and the optimal estimates. In the first two rows ($t = 1$), the bias, $\hat{p} - p^*$, is positive when $p^* < .5$ and negative when $p^* > .5$. In the last six rows ($t = 5$), the opposite holds true: the bias is negative when $p^* < .5$ and positive when $p^* > .5$.

Panel A

	Value	Share
$H(\hat{p}_{t+1})$	6.41	100%
$= I(\hat{p}_{t+1}; x_{t+1})$	0.25	3.95%
$+I(\hat{p}_{t+1}; \hat{p}_t, x_{1:t} \mid x_{t+1})$	3.88	60.5%
$+H(\hat{p}_{t+1} \mid \hat{p}_t, x_{1:t+1})$	2.28	35.6%

Panel B

	Value	Share
$I(\hat{p}_{t+1}; \hat{p}_t, x_{1:t} \mid x_{t+1})$	3.88	100%
$= I(\hat{p}_{t+1}; x_{1:t} \mid x_{t+1})$	0.88	23%
$+I(\hat{p}_{t+1}; \hat{p}_t \mid x_{1:t+1})$	3.00	77%

Table 8: Breakdown of the entropy of a response, using alternative method of entropy estimation. Same as Table 2, but using the estimation method developed by Nemenman *et al.* (2002) (see also Wolpert and Wolf, 1995).

effects in some behavioral tasks. We examine the same model as Yu and Cohen: specifically, we consider a Bayesian subject who believes that the color of the ring drawn at some trial t is a Bernoulli random variable whose parameter, p_t, may change from one trial to the next. At the first trial, p_1 is assumed to be sampled from a prior distribution f. We do not assume this prior to necessarily be the uniform distribution on $[0,1]$; we assume however that it is symmetric around 1/2, i.e., $f(1-p)=f(p)$ (in the next section we show that most subjects' responses are indeed symmetric). At any later trial t, the probability p_t is assumed to be equal to its preceding value, p_{t-1}, with probability $1-\gamma$; and with probability $\gamma \in [0,1]$, it is sampled from the prior distribution, f. In other words, the parameter p_t is subject to 'change points', which occur at each trial with probability γ. We posit, moreover, that the subject holds a belief distribution about the parameter p_t, which she updates in a manner that is optimal under the non-stationarity assumption, i.e., through Bayes' rule and taking into account the possible change points. Finally, we assume that our model subject provides as a response the (subjective) probability of observing a red ring in the next draw. (We note that Yu and Cohen show that under some conditions this Bayesian model can be approximated by a non-Bayesian model featuring an exponential filtering of the past observations. This corresponds to the quasi-Bayesian model with $\rho < 1$ that we examine, and reject, in the main text.)

Under these assumptions, we can compute, as in the main text, the subject's expected adjustment (implied by the subject's subjective probability). We obtain

$$\hat{p}(\hat{p}^R - \hat{p}) + (1-\hat{p})(\hat{p}^G - \hat{p}) = -\gamma\Big(\hat{p} - \frac{1}{2}\Big), \tag{6.1}$$

where $\hat{p}$ is the response at some trial, and $\hat{p}^R$ and $\hat{p}^G$ are the responses at the following trial if the new ring is red and if it is green, respectively. We have used the fact that because of its symmetry around 1/2, the mean of the prior is 1/2. In the absence of change points ($\gamma = 0$), we obtain the consistent-updates property (Eq. 3.6), i.e., the expected adjustment is zero. If the probability of change points is positive ($\gamma > 0$), then the subject takes into account the possibility that the parameter p_t may change to an unknown value, sampled from the prior, f. As a result, the expected adjustment is directed towards the mean of the prior, 1/2: if $\hat{p} > 1/2$, the expected adjustment is negative, while if $\hat{p} < 1/2$, it is positive.

Using similar notations as in the main text, namely, $\delta^R = \hat{p}(\hat{p}^R - \hat{p})$ and $\delta^G = -(1-\hat{p})(\hat{p}^G - \hat{p})$, the equation above implies the relation

$$\delta^G = \delta^R + \gamma\Big(\hat{p} - \frac{1}{2}\Big). \tag{6.2}$$

Thus this model predicts that when the response $\hat{p}$ is greater than 1/2, the quantities δ^G and δ^R should verify the inequality $\delta^G > \delta^R$, which corresponds to the area above the diagonal in Figure 8; and conversely ($\delta^G < \delta^R$) when $\hat{p} < 1/2$. The subjects, however, exhibit the opposite behavior (Fig. 8). As mentioned in the main text, when they believe that a red ring is more probable than a green ring (i.e., $\hat{p} > 1/2$), the adjustment they adopt in case of a red ring is too large in comparison with that which would be implied, under the consistent-update property, by the adjustment they adopt in case of a green ring. As a result, when $\hat{p} > 1/2$ their expected adjustment is positive, i.e., the (expected) new response is closer to 1. By contrast, the expected adjustment of the model subject with a belief in non-stationarity is negative when $\hat{p} > 1/2$ (Eq. 6.1), i.e., the (expected) new response is closer to 1/2. When $\hat{p} < 1/2$, the subjects move the response slider closer to 0, in expectation, while the model subject with a non-stationarity belief moves the slider closer to 1/2, in expectation. This model is thus incompatible with the behavioral data.

We note that these predictions of the model remain valid if we allow the model subject to hold only an imprecise record of the ring draws. Moreover, a similar qualitative discrepancy between model and empirical data is obtained if we assume a different response-selection strategy, in which the model subject is assumed to provide as a response the current expected value of p_t (instead of the probability that the next ring will be red). Consistently with these results, when we fit to subjects' data this model with belief in non-stationarity (assuming a symmetric Beta prior), we find that the likelihood of the model is maximized when $\gamma = 0$, i.e., when the model subject correctly assumes that the proportion of red rings in the box is not subject to random changes within blocks of trials.

Belief in sequential dependency

Although the subjects do not seem to be assuming that the proportion of red rings suddenly changes from one trial to the next, they may believe that the probability that the next ring be red depends on the color of the preceding ring. The hypothesis that subjects, in presence of a sequence of binary stimuli, are inferring the *conditional* probabilities generating the successive observations, has also been suggested as an account of sequential effects in choice reaction-time tasks and prediction tasks (Meyniel, Maheu, and Dehaene, 2016; Prat-Carrabin, Meyniel, and Azeredo da Silveira, 2022). Thus we consider a model in which the subject believes that the statistics of the samples are determined by two parameters, q_G and q_R, which represent the probability of observing a red ring after having observed a green ring, and after having observed a red ring, respectively. The subject holds a prior belief on these parameters, $f(q_G, q_R)$, which she updates through Bayes' rule at each trial. We denote by $f(q_G, q_R|x_{1:t})$ her posterior belief after having observed the sequence $x_{1:t}$. As in the previous section, we assume that the subject provides as a response the inferred probability that the next ring be red, which is the expected value of q_G, if the previous ring is green, or the expected value of q_R, if the previous ring is red.

First, we assume a prior of the form $f(q_R, q_G) = (q_G + 1 - q_R)f_G(q_G)f_R(q_R)$. Under this assumption, after observing any number $t \geq 1$ of ring draws, the posterior distribution is one under which q_R and q_G are distributed independently of each other, i.e.,

$$f(q_G, q_R|x_{1:t}) = f_G(q_G|x_{1:t})f_R(q_R|x_{1:t}), \tag{6.3}$$

where f_G and f_R are the posterior beliefs over q_G and q_R. (We note that if this equality is verified after one observation, x_1, then it is also verified for all subsequent sequences of observations.) This is the kind of posterior assumed by Meyniel *et al.* (2016). Under this assumption, one can update separately the two marginal posteriors f_G and f_R, as

$$\begin{aligned} f_G(q_G|x_{1:t}) &\propto f_G(q_G|x_1)q_G^{n_{GR}}(1-q_G)^{n_{GG}}, \\ \text{and } f_R(q_R|x_{1:t}) &\propto f_R(q_R|x_1)q_R^{n_{RR}}(1-q_R)^{n_{RG}}, \end{aligned} \tag{6.4}$$

where n_{XY} is the number of times that the color X is followed by the color Y in the sequence $x_{1:t}$.

It follows that the response to a sequence that ends with a green ring is determined by the counts n_{GR} and n_{GG}, while the response to a sequence that ends with a red ring is determined by the counts n_{RR} and n_{RG}. Therefore, the two sequences (R, G) and (R, R, R, R, G), for instance, are predicted to result in equal responses (as in both cases $n_{GR} = n_{GG} = 0$); but the subjects provide very different responses, in these two cases (on average 0.506 and 0.777). It is obvious in the behavioral data that different sequences that have the same pair of counts (n_{GR} and n_{GG} for the sequences ending with a green ring, or n_{RR} and n_{RG} for the sequences ending with a red ring) result in different responses, in the general case. F-tests of the hypotheses that these responses are in fact equal on average in subjects' data, for each set of sequences that have the same pairs of counts, are all rejected with a p-value lower than 10^{-3}, except for the sequences for which an observer with a correct belief about the structure of the task would also provide identical responses, such as (G, G, G, R, G), (G, G, R, G, G), and (G, R, G, G, G). We conclude that this model under the independence assumption (Eq. 6.3) fails at reproducing subjects' responses.

We relax the independence assumption, and consider a more general case. We now only impose a symmetry assumption on the prior: specifically, that $f(q_G, q_R) = f(1-q_R, 1-q_G)$. (This property implies that the prior does not contain an *a priori* bias towards one color: the prior belief that a red ring is more probable after observing a red ring, for instance, is equal to the prior belief that a green ring is more probable after observing a green ring. See the next section for an analysis of the symmetry in subjects' responses.) The rules of probabilities imply a series of relations, which we now present, between the responses of a subject who is inferring the conditional probabilities q_R and q_G. First, we note that these two parameters together determine the probability of an observation x, unconditional on any preceding observation, as

$$P(x|q_G, q_R) = \frac{q_G^x (1-q_R)^{1-x}}{q_G + 1 - q_R}, \tag{6.5}$$

where $x = 1$ for a red ring, and $x = 0$ for a green ring (and thus the numerator is q_G if the ring is red, and $1 - q_R$ if it is green). The joint probability of observing the sequence $x_{1:t}$, given q_G and q_R, is then

$$P(x_{1:t}|q_G, q_R) = \frac{q_G^{x_1}(1-q_R)^{1-x_1}}{q_G + 1 - q_R} q_G^{n_{GR}} (1-q_G)^{n_{GG}} q_R^{n_{RR}} (1-q_R)^{n_{RG}}, \tag{6.6}$$

and given a prior $f(q_G, q_R)$, one obtains the subjective joint probability of the sequence, as $P(x_{1:t}) = \iint P(x_{1:t}|q_G, q_R) f(q_G, q_R) \mathrm{d}q_G \mathrm{d}q_R$. Consequently, two sequences $x_{1:t}$ and $\tilde{x}_{1:t}$ have the same subjective joint probability of occurence if

$$\begin{aligned} n_{GG} &= \tilde{n}_{GG}, \\ n_{GR} &= \tilde{n}_{GR} + \tilde{x}_1 - x_1, \\ n_{RG} &= \tilde{n}_{RG} - \tilde{x}_1 + x_1, \\ \text{and } n_{RR} &= \tilde{n}_{RR}, \end{aligned} \tag{6.7}$$

where $\tilde{n}_{XY}$ is the number of times that the color X is followed by the color Y in the sequence $\tilde{x}_{1:t}$. Finally, we note that the subjective joint probability of $x_{1:t}$ can be expressed as a function of the successive responses of the model subject when presented with the sequence $x_{1:t-1}$, as

$$\begin{aligned} P(x_{1:t}) &= P(x_1)P(x_2|x_1)P(x_3|x_{1:2}) \dots P(x_t|x_{1:t-1}) \\ &= \frac{1}{2} \prod_{i=1}^{t-1} \hat{p}(x_{1:i})^{x_{i+1}} (1 - \hat{p}(x_{1:i}))^{1-x_{i+1}}, \end{aligned} \tag{6.8}$$

where $\hat{p}(x_{1:i})$ is the subject's response after having observed the sequence $x_{1:i}$. We have used the fact that the symmetry property of the prior implies that the subjective probability of the first outcome, red or green, is 1/2.

In summary, two sequences $x_{1:t}$ and $\tilde{x}_{1:t}$ that verify the equalities in Eq. 6.7 have equal subjective joint probabilities, and these can be recovered from the successive responses to each sequence, following Eq. 6.8; thus providing testable predictions of the model. For instance, the sequences (G, R, R) and (R, R, G) verify the equalities in Eq. 6.7, which implies that their subjective probabilities are equal, i.e., $P(G, R, R) = P(R, R, G)$; and thus we obtain the prediction $\hat{p}(G)\hat{p}(G, R) = \hat{p}(R)(1 - \hat{p}(R, R))$, which relates the two successive responses for the sequence (G, R) with the two successive responses for the sequence (R, R). In subjects' data, however, these two quantities are significantly different (t-test p-value: 10^{-15}). The two sequences that 'mirror' the previous two (by inverting R and G), (R, G, G) and (G, G, R), also verify Eq. 6.7, and thus they make a prediction of the same kind, $(1 - \hat{p}(R))(1 - \hat{p}(G, G)) = (1 - \hat{p}(G))\hat{p}(G, G)$, but here also the hypothesis of this equality in subjects' data is rejected (p-v.: 10^{-18}). More generally, there are 50 pairs of sequences of length no greater than 6 that verify Eq. 6.7 and whose two elements are not 'mirrors' of each other. We test the 50 corresponding predictions. For 37 of them, the t-test of equality is rejected at the .01 level (and for 30 of them it is rejected at the .001 level). Moreover, ten of these predictions are rejected at the .01 level by more than 50% of subjects, and four are rejected at the same level by more than 85% of subjects. The responses of a majority of subjects thus do not seem to be consistent with the predictions of this model.

To complement this analysis, we fit the model to subjects' data. This requires specifying a prior, $f(q_G, q_R)$. First we consider the case in which the prior is the product of the marginal priors $f_G(q_G)$ and $f_R(q_R)$, which we assume to be Beta distributions with parameters (α, β) and (β, α), respectively, i.e., $f(q_G, q_R) \propto q_G^{\alpha}(1 - q_G)^{\beta} q_R^{\beta}(1 - q_R)^{\alpha}$. This prior is symmetric, in the sense that $f(q_G, q_R) = f(1 - q_R, 1 - q_G)$. In addition, after observing a sample x, the subject's posterior is proportional to the product of this prior and of the probability $P(x|q_G, q_R)$, as given in Eq. 6.5: we note that the resulting posterior is *not* such that q_G and q_R are independent, as in Eq. 6.3, which implied predictions that we rejected. In other words, with this prior, all observations are informative about both q_G and q_R. We fit this model by maximizing its likelihood (with the additional assumption of noise in the response). We compare the resulting maximum likelihood with that of the model of an optimal Bayesian observer who has a correct belief on the structure of the task. Although we clearly reject that observer model in the main text, we find that the model with beliefs in sequential dependency result in a much lower likelihood. The corresponding Bayes factor is about 10^{-780}, indicating a very poor fit to subjects' data, in comparison with the optimal-observer model.

This result, however, largely follows from our choice of prior. Although it may seem natural, for an observer who believes that the random draws obey different laws when the preceding ring is red and when it is green, the form of prior we have assumed posits that the laws in these two cases are completely unrelated: for all choices of α and β, there is no correlation between q_G and q_R. But the true structure of the random draws, in our task, is precisely the opposite extreme case, as q_G and q_R should in fact be equal (and thus perfectly correlated). In other words, this functional form of the prior does not nest any prior that implies the correct assumption about the lack of sequential dependency in the task. A possibility, instead, is that the belief of subjects is not so different from the truth, and that their prior on q_G and q_R implies that these two parameters are correlated, although not necessarily equal, thus allowing from some amount of deviation from the correct belief. To that end, we find it convenient to reparameterize the (putative) structure of the draws, and replace q_G and q_R by two other parameters: p_R, which determines the unconditional probability of observing a red ring (unconditional on any preceding observation), and ω, a parameter that indicates the extent to which the structure of the random draws (under these parameters) deviates from the true structure (in which the probability of a red ring does not depend on the color of the previous ring). Specifically, given p_R and ω, the unconditional probability of a red ring and the probability of a red ring conditional on the previous ring being red, respectively, are

$$P(x = R|p_R, \omega) = p_R, \text{ and } P(x_{t+1} = R|x_t = R, p_R, \omega) = p_R + \frac{\omega}{p_R}. \tag{6.9}$$

Implied is the probability of a red ring conditional on the previous ring being green, as $P(x_{t+1} = R|x_t = G, p_R, \omega) = p_R - \frac{\omega}{1-p_R}$. If $\omega = 0$, the probability of a red ring does not depend on the color of the previous ring, while if $\omega \neq 0$, the probability of a red ring following a red ring is different from the probability of a red ring following a green ring.

We now choose a functional form for the prior, $f(p_R, \omega)$, on these two parameters. We assume that the marginal prior on p_R is a symmetric Beta distribution with parameter α. As for ω, we note that the range of permitted values of ω depends on the value of p_R. For each p_R, we assume that within this range the prior over ω is proportional to a Gaussian function, $e^{-\omega^2/\tau}$, where $\tau > 0$, which we extend to a Dirac delta function, $\delta(\omega)$, in the case where $\tau = 0$. In this case, the subject believes that there is no dependency across successive samples ($\omega = 0$), and the prior is equivalent to the 'incorrect prior' examined in Section 1.3. With a positive parameter τ, the prior assigns a positive probability on structures of the random draws that feature sequential dependency ($\omega \neq 0$).

When fitting this model to subjects' data, we find that its likelihood is maximized when $\tau = 0$, i.e., when the prior posits that with probability 1 there is no dependency in successive samples ($\omega = 0$). Looking at the individual data, we find that for all the subjects the likelihood of this model is also maximized when $\tau = 0$. Thus there does not seem to be evidence that subjects hold erroneous beliefs about the structure of the random ring draws, and the model with belief in sequential dependency does not capture subjects' responses better than the model of Bayesian inference with incorrect prior of Section 1.3, which we have rejected. In conclusion, this model-fitting analysis, in addition to the results above regarding predictions of the model that are not observed in the behavioral data, suggests that the hypothesis of a belief in sequential dependency does not provide a satisfying account of subjects' responses.

Symmetry hypothesis

We examine in the behavioral data the hypothesis that the responses of subjects are symmetric around 1/2, i.e., that their average response to a sequence is equal to 1 minus their average response to the corresponding 'mirror' sequence (in which all the observations R and G are inverted). Specifically, for each of the 31 sequences that start with a red ring, we conduct a t-test of the equality between the subjects' response to this sequence and their response to the mirror sequence. For all the sequences except one, the equality hypothesis is *not* rejected at the 0.01 level. With a p-value of 0.009, the hypothesis is rejected at the 0.01 level for the sequence that contains five rings, all of them red. The average response to this sequence is 0.8946, which implies under the symmetry hypothesis an average response to the mirror sequence (five green rings) of 0.1054, but the subjects' average response is 0.1186, i.e., larger by 0.0132. In other words, the subjects on average seem to have a slight bias for larger values, or to the right-hand side of the response slider. When running the same tests individually for each subject, we find that for all the sequences, 10% or less of the subjects reject the hypothesis; and for 15 sequences (out of 31) no subject rejects the hypothesis. Overall, we conclude that the symmetry hypothesis holds for most subjects, and that deviations from this symmetry do not constitute a major feature of the behavioral data.

Block-level priors and sequential effects across blocks

Hierarchical Bayesian inference

In all the Bayesian and quasi-Bayesian models that we have considered, we have assumed that the subjects' prior is always the same at the beginning of all the blocks of five trials. This is consistent with the task (in which the proportion of red rings is chosen from an unchanging prior, the uniform distribution), and with the instructions given to the subjects. Here we examine the hypothesis that the subjects are uncertain about the prior, but that they hold a belief about it, which they update throughout the experiment. In other words, we consider a hierarchical Bayesian model of learning of the prior across the blocks of trials.

Specifically, we consider a model in which priors are Beta distributions with parameters α and β, and subjects hold a 'hyperprior' belief, $h(\alpha,\beta)$, about these two parameters. Subjects update this hyperprior on the basis of the observed ring draws. It is useful to define the function $g(m,t;\alpha,\beta)$, as the probability of a sequence of t draws, $x_{1:t}$, of which m are red, conditional on the prior specified by the parameters α and β:

$$\begin{aligned} g(m,t;\alpha,\beta) &\equiv P(x_{1:t}|\alpha,\beta), \quad \text{for any } x_{1:t} \text{ such that } \sum x_{1:t} = m \\ &= \prod_{i=1}^{t} P(x_i|\alpha,\beta,x_{1:i-1}), \end{aligned} \tag{6.10}$$

i.e.,

$$g(m,t;\alpha,\beta) = \frac{\prod_{i=0}^{m-1}(\alpha+i)\prod_{i=0}^{t-m-1}(\beta+i)}{\prod_{i=0}^{t-1}(\alpha+\beta+i)}. \tag{6.11}$$

(We use the convention $x_t = 1$ when the ring at trial t is red, and $x_t = 0$ when it is green.) In our experiment, each block of trials comprised $T = 5$ ring draws. After the last ring draw of a given block i, let $N_m^{(i)}$ be the number of blocks in which the subject has observed m red rings out of the T draws, for each $m \in \{0,1,2,3,4,5\}$. Applying Bayes' rule, the model subject updates its belief about the parameters α and β, and obtains the 'hyperposterior' h_i, as

$$h_i(\alpha,\beta) = \frac{1}{Z} h(\alpha,\beta) \prod_{m=0}^{T} g(m,T;\alpha,\beta)^{N_m^{(i)}}, \tag{6.12}$$

where Z is a normalization constant. This hyperposterior will be the hyperprior at the beginning of the next block.

Given the hyperprior h_i, the optimal response, $p_i^*(x_{1:t})$, to a sequence of draws $x_{1:t}$, is the expected value of the proportion p of red rings conditional on the sequence (and given the hyperprior h_i), $\mathbb{E}[p|x_{1:t}]$, i.e., the (subjective) probability of observing a red ring in the next draw ($x_{t+1} = 1$), conditional on the sequence $x_{1:t}$, which we denote by $P_i(x_{t+1} = 1|x_{1:t})$. Thus we have

$$\begin{aligned} p_i^*(x_{1:t}) &\equiv P_i(x_{t+1} = 1|x_{1:t}) \\ &= \frac{P_i(x_{1:t}, x_{t+1} = 1)}{P_i(x_{1:t})} \\ &= \frac{\iint P(x_{1:t}, x_{t+1} = 1|\alpha, \beta) h_i(\alpha, \beta) \mathrm{d}\alpha \mathrm{d}\beta}{\iint P(x_{1:t}|\alpha, \beta) h_i(\alpha, \beta) \mathrm{d}\alpha \mathrm{d}\beta}, \end{aligned} \tag{6.13}$$

where we have used the definition of conditional probabilities, and marginalized over α and β. Thus, by definition of the function g,

$$p_i^*(x_{1:t}) = \frac{\iint g(\alpha, \beta; m+1, t+1) h_i(\alpha, \beta) \mathrm{d}\alpha \mathrm{d}\beta}{\iint g(\alpha, \beta; m, t) h_i(\alpha, \beta) \mathrm{d}\alpha \mathrm{d}\beta}. \tag{6.14}$$

Given a hyperprior h, we are thus able to compute the optimal response (Eq. 6.14) to the sequences presented in the successive blocks of the task, using the function g (Eq. 6.11) and the hyperprior update rule (Eq. 6.12). Finally, as in several of our other models, we assume that there is a degree of noise in the subject's selection of a response, which we model as a Gaussian variable with standard deviation σ, truncated so that the response falls in the interval $[0, 1]$.

As for the hyperprior $h(\alpha, \beta)$, we choose it in the family of bivariate normal distributions over (α, β), truncated to the first quadrant (i.e., negative α and β have zero probability). The two components of the non-truncated normal distribution are assumed to have equal means, denoted by α_0. The covariance matrix has two elements on its diagonal, which we also assume to be equal, and which we denote by $\tilde{\sigma}^2$, and we denote by ρ the Pearson correlation coefficient between the two components (this determines the off-diagonal elements of the covariance matrix, as $\rho\tilde{\sigma}^2$).

When fitting this model to each subject, we find that the best-fitting values of α_0 are very close to the best-fitting values of the parameter α of the Bayesian model with a (fixed) incorrect prior of the form Beta(α, α) (line 2 in Table 5; this model is nested by the hierarchical Bayesian model). For 52% of subjects, the best-fitting values of these two parameters differ by less than 1%, and for 90% of subjects they differ by less than 10%. As for the best-fitting values of $\tilde{\sigma}$, which can be understood as representing the degree of uncertainty in the subjects' hyperpriors, we find that they are relatively small: for 52% of subjects, $\tilde{\sigma}$ is smaller than 0.05; for 90% of subjects it is smaller than 0.25; and the ratio $\tilde{\sigma}/\alpha_0$ is lower than 0.17 for 95% of subjects. As a consequence, the responses of this model are very close to that of the (non-hierarchical) Bayesian model with fixed incorrect prior (Fig. 10, second row), and thus it does not reproduce the various behavioral patterns we have found in subjects' data, although it has two more parameters. Consistently, the BIC for this model is higher than that of the model with fixed incorrect prior (with a difference of 20 if all subjects are required to have the same parameters, and of 173 with individual parameters). This suggests that the hierarchical Bayesian model does not provide a much better fit than the model with incorrect prior; in any case, it provides a much worse fit than our best-fitting model.

Sequential effect across blocks

Although the hierarchical Bayesian model just examined does not reproduce the behavioral patterns found in subjects' responses, it remains possible that the sequence of rings drawn in one block influences the subjects' responses in the following block. For instance, a hypothesis is that a subject starts a new block with a prior that depends on whether the sequence of rings drawn in the preceding block featured a majority of red rings (i.e., three or more red rings, out of five draws), or a majority of green rings. This could stem from learning across blocks (perhaps similar to that in the previous section, but not necessarily Bayesian, or not necessarily with the kind of prior we have assumed), in which case a red-dominant block would result in a prior shifted towards higher proportions of red rings. But subjects might also believe that after a red-dominant block, green rings are 'due', as in the well-known gambler's fallacy, in which case a red-dominant block would be followed by responses biased towards smaller red proportions. We thus test whether subjects' responses differ when they come after a red- or a green-dominant block. Pooling together the responses of all the subjects, we run, for each of the 62 possible sequences in the task, a t-test of equality of the responses to the sequence when they occurred after a red-dominant block, and when they occurred after a green-dominant block. In only one such test is the null hypothesis of equality rejected at the .01 level. Under the null hypothesis, the binomial probability of rejecting one or more test out of 62, at the .01 level, is .13, suggesting that there is no strong evidence that the hypothesis of equality should be rejected. We also run these tests at the individual level. We note that for each subject we do not have enough data to run the 62 tests just mentioned; on average, we are able to run 42 tests per subject. For 81% of subjects, no more than one of these tests is rejected (for 52%, no test is rejected), and no subject rejects more than 5 tests. We conclude that there is no strong evidence that the behavior of subjects significantly changes with the color that dominated the preceding block.

Random block-level priors

In the previous two sections, we have examined the possibility that the subjects start new blocks of trials with a prior that depends on the outcomes of the preceding block. Here we consider the hypothesis that the prior is different in each block, but not for reasons related to the preceding block: instead, we model the block priors as independent and identically distributed random variables. This stochasticity could originate, for instance, in the imprecise encoding of the prior. Specifically, we assume here that the prior at the beginning of a block is a Beta distribution, whose pair of parameters (α, β) is sampled from a bivariate normal distribution truncated to the first quadrant. The means of the two components of the non-truncated normal distribution are assumed to be equal, and denoted by α_0. The two elements of the diagonal of its covariance matrix are also assumed equal, and denoted by $\tilde{\sigma}^2$, and the Pearson correlation coefficient between the two components is denoted by ρ (this determines the off-diagonal elements of the covariance matrix, as $\rho\tilde{\sigma}^2$). In

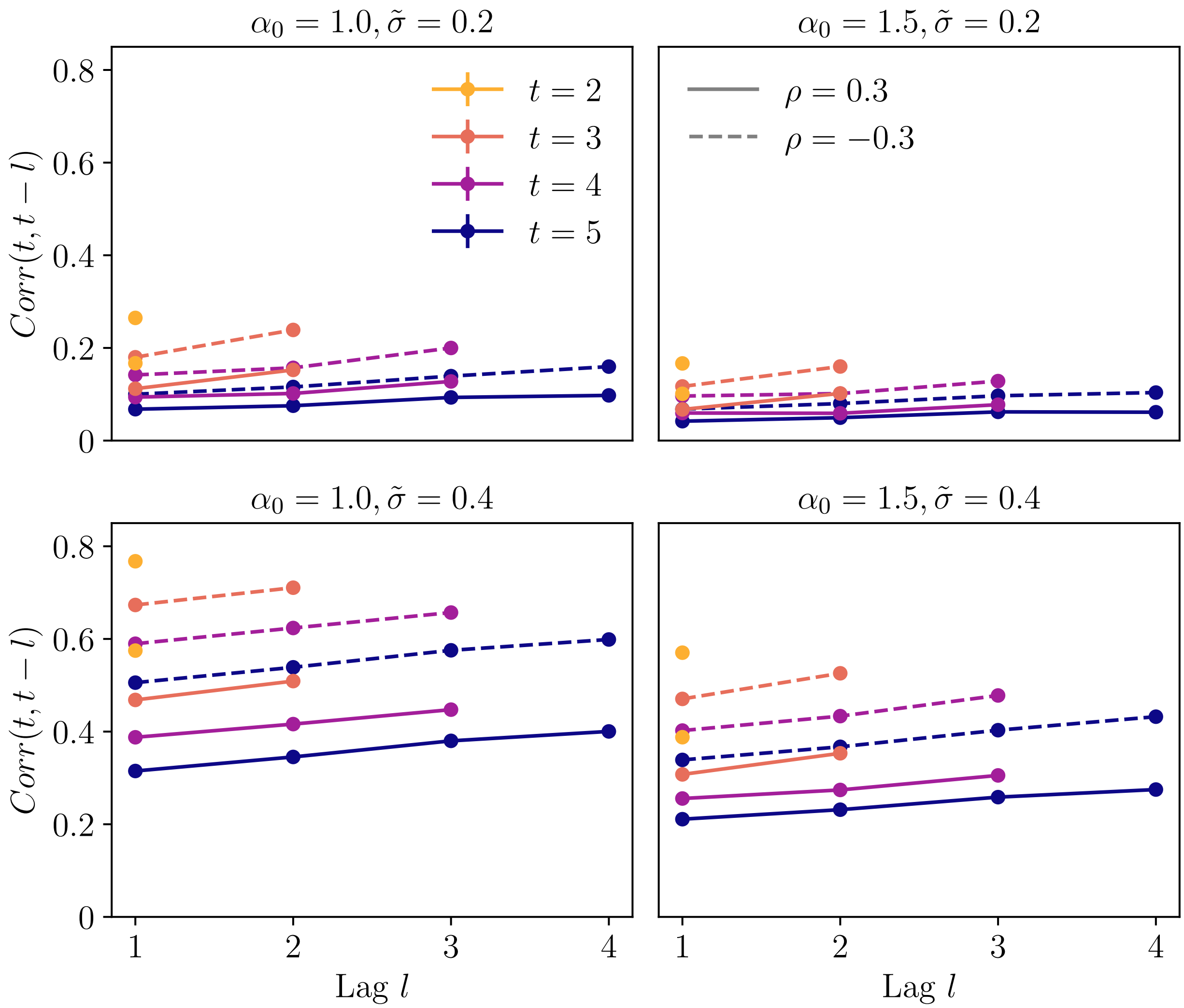

Figure 12: Block-level random priors: Autocorrelation in responses.

other words, the prior's pair of parameters (α, β) is a noisy readout of the pair (α_0, α_0), with $\tilde{\sigma}$ and ρ specifying the structure of the noise. Our model subject performs Bayesian inference, using this Beta(α, β) prior, and provides as a response the mean of the Bayesian posterior, with the addition of some response-selection noise, modeled as a centered Gaussian variable with standard deviation σ, truncated to the $[0, 1]$ interval. As all the responses within one block depend on the random prior for that block, this model predicts a positive correlation between successive responses; but we seek to examine whether it reproduces the decreasing correlation between distant trials of the same block, as observed in subjects' data (Fig. 5B).

To this end, we run simulations of this model, with the parameter α_0 taking the value 1 (which corresponds to the correct prior) or 1.5 (which is approximately the best-fitting parameter of the Bayesian model with incorrect prior, when the responses of the subjects are pooled together, and it is close to 1.6, the median of the best-fitting parameters of this same model when fitting subjects independently). As for the correlation parameter, we run simulations with $\rho = -0.3$ and $\rho = 0.3$. As for the response-selection noise parameter, σ, we choose it as a function of $\tilde{\sigma}$ (which parameterizes the noise in the prior), and such that the overall variability of the model subject's responses matches that of the subjects (i.e., a standard deviation of about .088). We note that aiming at reproducing the subjects' variability puts an upper bound on the choice of $\tilde{\sigma}$; the value of this upper bound is about 0.7. We simulate the model with $\tilde{\sigma} = 0.2$ and $\tilde{\sigma} = 0.4$, two values chosen to roughly cover the range $[0, 0.7]$. In all our simulations, we find that the correlation between the responses in a block is positive, as expected (Fig. 12). When the prior-noise parameter, $\tilde{\sigma}$, increases, the autocorrelation increases (compare first and second rows in Fig. 12), which is readily intuited, as the noise in the prior is the source of the autocorrelation. The autocorrelation with $\rho = -0.3$ is larger than with $\rho = 0.3$ (compare dashed lines to solid lines in Fig. 12), presumably because anti-correlated values of α and β yield more skewed priors than correlated values of α and β, resulting in larger variations in the subsequent responses. Finally, the autocorrelation in responses does not decrease with the distance between two trials in a block, as it does with the subjects, and in fact it seems to increase with the distance between two trials. Therefore, this model of block-level random priors does not seem to be able to reproduce the decreasing autocorrelation in responses found in subjects' data.

Bayesian properties: additional tests

Calibrated-responses property

Analysis-of-variance (ANOVA) F-tests of the equality $\mathrm{E}[p-\hat{p}|\hat{p} \in I_i] = 0$ for all intervals I_i For each interval I_i of length .025, we consider the distribution of the differences between the true probability and the response, $p - \hat{p}$. Equation 3.3 predicts the difference to be on average zero for all intervals I_i. We run an ANOVA F-test of whether the means of these distributions of differences are all equal. (The F-test does not test whether they are equal to zero, but whether they are all equal to each other. However, if the null hypothesis is rejected, we can conclude that in at least one occurrence the mean is different from zero.) When pooling the responses of all the subjects, we obtain a p-value of 4.2e-219, and the test is rejected at the .01 level for 81% of the subjects (median p-value: 3.5e-8). These results support the conclusion, in the main text, that the responses of a large majority of subjects do not satisfy the Bayesian calibrated-responses property.

***t*-tests** We consider a different strategy, based on t-tests, to test Eq. 3.3. For each of the 41 intervals I_i, we run the one-sample t-test of equality between $\hat{p}-p$, where $\hat{p} \in I_i$, and zero. We reject the null hypothesis for 35 tests at the .05 level and for 34 tests at the .01 level (Fig. 7). Instead of a series of individual tests, we seek to reach a conclusion, instead, about the one hypothesis that $\mathrm{E}[p-\hat{p}|\hat{p} \in I_i] = 0$ for all intervals I_i. Under this null hypothesis and for a significance level α, the number of rejections (among the 41 tests) follows a binomial distribution with parameters 41 and α. Thus we expect, under this hypothesis, to obtain 5% of 41, i.e., an expectation of 2.05 rejections, at the 0.05 level, and 0.41 rejections at the 0.01 level (we obtain, instead, 35 and 34). We look at the probability of obtaining a number of rejections equal to, or greater than, the obtained numbers. With $\alpha = 0.05$, the probability of 35 or more rejections is 8.5e-42. With $\alpha = 0.01$, the probability of 34 or more rejections is 4.2e-64. These results substantiate our conclusion that the calibrated-responses property is not verified in subjects' data.

Consistent-updates property

We run two series of t-tests of the consistent-updates property in subjects' data (Eqs. 3.9 and 3.10). First, for each possible sequence of ring draws, we pool the obtained responses and run a t-test of Eq. 3.9 (Fig. 8A). The p-values of the t-tests of equality are all below 1e-3, except for the sequences in which the red and green numbers are balanced (Table 9). We run the tests for each subject, and find that for nine sequences the null hypothesis (Eq. 3.9) is rejected at the .01 level for more than half of the subjects, and for four sequences it is rejected by more than 80% of the subjects (Table 9, last column). Second, we test whether the relation predicted by Eq. 3.10 is verified on average for each response interval I_i (Fig. 8B). The null hypothesis is rejected at the .05 level for 76% of subjects, and at the .01 level for 73% of subjects. The binomial probabilities of obtaining proportions equal or higher than these, under the null hypothesis (see previous section), are 5.2e-34 and 1e-53. We conclude, as in the main text, that the Bayesian consistent-updates property is not verified in subjects' data.

Quasi-Bayesian models

This section provides more details on the results regarding the quasi-Bayesian models. In this class of model, the response, $p_{\alpha\rho\lambda}$, is determined by the sequence of rings presented and by the parameters α, ρ, and λ (see Eqs. 4.6 and 4.7; note also that with $\rho = 1$ it is immediate that $p_{\alpha\rho\lambda}$ equals $p_{\alpha\lambda}$ given in Eq. 4.5).

Average response to sequences of length t with n_R red rings

Given a number of red rings, n_R, and a number of green rings, n_G, we find that the average response of a quasi-Bayesian observer — averaged over all the sequences that contain n_R red rings and n_G green rings — is

$$\bar{p}_{\alpha\rho\lambda} = \frac{1}{2} + \frac{\lambda}{2} \frac{(n_R - n_G)\frac{1}{t}\sum_{i=0}^{t-1}\rho^i}{\lambda\sum_{i=0}^{t-1}\rho^i + 2\rho^t(\alpha - 1) + 2}, \tag{6.15}$$

where $t = n_R + n_G$. Alternatively this can be expressed as a function of the optimal estimate, p^*, as

$$\bar{p}_{\alpha\rho\lambda} = p^* + (2p^* - 1)\frac{\lambda\frac{1}{t}\sum_{i=0}^{t-1}\rho^i - \rho^t(\alpha - 1) - 1}{\lambda\sum_{i=0}^{t-1}\rho^i + 2\rho^t(\alpha - 1) + 2}. \tag{6.16}$$

Noting the equality

$$\sum_{i=0}^{t-1}\rho^i = t + (\rho - 1)\sum_{i=0}^{t-1} i\rho^{t-1-i}, \tag{6.17}$$

we can also write the average response as

$$\bar{p}_{\alpha\rho\lambda} = p^* + (2p^* - 1)\frac{\lambda - 1 + \lambda(\rho - 1)\frac{1}{t}\sum_{i=0}^{t-1} i\rho^{t-1-i} - \rho^t(\alpha - 1)}{\lambda\sum_{i=0}^{t-1}\rho^i + 2\rho^t(\alpha - 1) + 2}. \tag{6.18}$$

Sequence	p-value		N_R	N_G	$\% < .01$
0	5.3e-51	***	712	1388	.71
1	1.4e-31	***	1366	685	.52
00	4.9e-64	***	317	1071	.86
01	.99		332	380	.19
10	.074	*	333	352	.10
11	8.5e-68	***	1017	349	.86
000	7.9e-59	***	231	840	.81
001	.00056	***	127	190	.24
010	6.9e-10	***	146	234	.38
011	7.5e-18	***	202	130	.57
100	3.2e-15	***	147	205	.33
101	2.6e-08	***	203	130	.33
110	3.7e-05	***	218	131	.38
111	2.9e-62	***	784	233	.86
0000	1.3e-29	***	134	706	.62
0001	5e-10	***	62	169	.33
0010	3e-08	***	52	138	.10
0011	.76		62	65	.05
0100	1.1e-14	***	77	157	.33
0101	.82		75	71	.00
0110	.037	**	60	70	.00
0111	3.3e-13	***	135	67	.33
1000	1.1e-09	***	67	138	.24
1001	.12		75	72	.00
1010	.3		68	62	.00
1011	7.9e-16	***	129	74	.33
1100	.61		73	58	.05
1101	8.7e-12	***	155	63	.14
1110	8.3e-05	***	155	78	.19
1111	9.5e-32	***	660	124	.67

Table 9: Bayesian update property: t-tests of Eq. 3.9 for each sequence of observations.

If the prior is correct ($\alpha = 1$), the likelihood is correctly weighted ($\lambda = 1$), and only the prior is misweighted ($\rho \neq 1$), then the average response is

$$\bar{p}_\rho = p^* + (2p^* - 1)(\rho - 1)\frac{1}{t}\frac{\sum_{i=0}^{t-1} i\rho^{t-1-i}}{\sum_{i=0}^{t-1} \rho^i + 2}. \tag{6.19}$$

The ratio of sums, in the right-hand-side of this equation, is zero for $t = 1$ and positive for $t > 1$. Thus after one ring ($t = 1$) the model subject's estimate is the optimal estimate p^*, and after more than one ring ($t > 1$), the model subject either overreacts, on average, to the evidence if $\rho > 1$, or underreacts (conservatism), if $\rho < 1$; but the model cannot reproduce the subjects' conservatism followed by its reversal.

If the prior is correct ($\alpha = 1$) but both the prior and the likelihood are allowed to be assigned incorrect weights ($\rho \neq 1$ and $\lambda \neq 1$), then the average estimate in response to the sequences of t rings among which n_R are red is

$$\bar{p}_{\rho\lambda} = p^* + (2p^* - 1)\frac{\lambda - 1 + \lambda(\rho - 1)\frac{1}{t}\sum_{i=0}^{t-1} i\rho^{t-1-i}}{\lambda\sum_{i=0}^{t-1}\rho^i + 2}. \tag{6.20}$$

Whether the responses of the model subject are under- or over-reactions to the evidence, as compared to the optimal estimates, is determined by the sign of the numerator in the right-hand-side of this equation (e.g., if it is positive, the subject provides an estimate larger than optimal if the optimal estimate is greater than .5). With $t = 1$, the numerator is equal to $\lambda - 1$, thus the model reproduces the conservatism of subjects at the first trial only if $\lambda < 1$. For $t > 1$, if $\lambda < 1$ and $\rho < 1$ then the numerator is negative and the model predicts conservatism at all trials. Only with $\lambda < 1$ and $\rho > 1$ can the numerator start negative (for $t = 1$) and become positive (at some $t > 1$). In this case after a few ring draws the model subject underreacts to the evidence, while after longer sequences it overreacts to the evidence.

If the weight on the likelihood is correct ($\lambda = 1$) but the prior and the weight on the prior are allowed to be incorrect ($\alpha \neq 1$ and $\rho \neq 1$), then the average estimate is

$$\bar{p}_{\alpha\rho} = p^* + (2p^* - 1)\frac{\frac{1}{t}\sum_{i=0}^{t-1}\rho^i - \rho^t(\alpha - 1) - 1}{\sum_{i=0}^{t-1}\rho^i + 2\rho^t(\alpha - 1) + 2}. \tag{6.21}$$

The subject exhibits conservatism if the sign of the factor to $(2p^* - 1)$, in the right-hand-side of this equation, is negative, and they exhibit overreaction if it is positive. With $t = 1$, this factor is $-\rho(\alpha - 1)/(3 + 2\rho(\alpha - 1))$. It is negative if $\alpha > 1$, or if $\alpha < 1$ and $\rho > \frac{3}{2(1-\alpha)}$, and thus only in these two cases does the model subject exhibits conservatism after the first ring, as do the subjects. If $\alpha < 1$ and $\rho > \frac{3}{2(1-\alpha)}$, then, noting that $\rho > \frac{3}{2}$, it is easy to show that the numerator in the equation above is positive, for any t, and the denominator is negative, for any t, and thus the factor is negative for any t, i.e., the model subject exhibits conservatism at all trials, contrary to the subjects. In the other case, $\alpha > 1$, it is immediate to see that the denominator is positive for any t. As for the numerator, we find, using the equality in Eq. 6.17, that it is positive if $\alpha < 1 + \frac{1}{t}\frac{\rho-1}{\rho}\sum_{i=0}^{t-1} i\rho^{-i}$. This provides an upper bound on α, in order for the model subject to exhibit overreaction at some trial. We note that with the constraint $\alpha > 1$, this inequality implies that ρ should be greater than 1. In sum, this inequality and the condition $\alpha > 1$ constitute necessary and sufficient conditions for the model subject to exhibit conservatism at the first trial and the opposite of conservatism at some later trial. An implied pair of necessary conditions is $\alpha > 1$ and $\rho > 1$; these are not sufficient conditions, and in fact the subjects' best-fitting values of these parameters, $\alpha = 1.47$ and $\rho = 1.07$, verify this pair of conditions, but not the bound on α just presented, for $t \leq 5$; thus this model with these parameters results in conservatism at all five trials (Fig. 13, third row).

Finally, in the general case (in which all three parameters are allowed to differ from 1), we note that if ρ is close to 1, then $\sum_{i=0}^{t-1}\rho^i$ is approximately t, and $\rho^t \approx 1 + t(\rho - 1)$. Hence we approximate the average response $\bar{p}_{\alpha\rho\lambda}$ as

$$\bar{p}_{\alpha\rho\lambda} \approx \frac{1}{2} + \frac{1}{4}\lambda\frac{n_R - n_G}{\alpha + t[\frac{1}{2}\lambda + (\alpha - 1)(\rho - 1)]}. \tag{6.22}$$

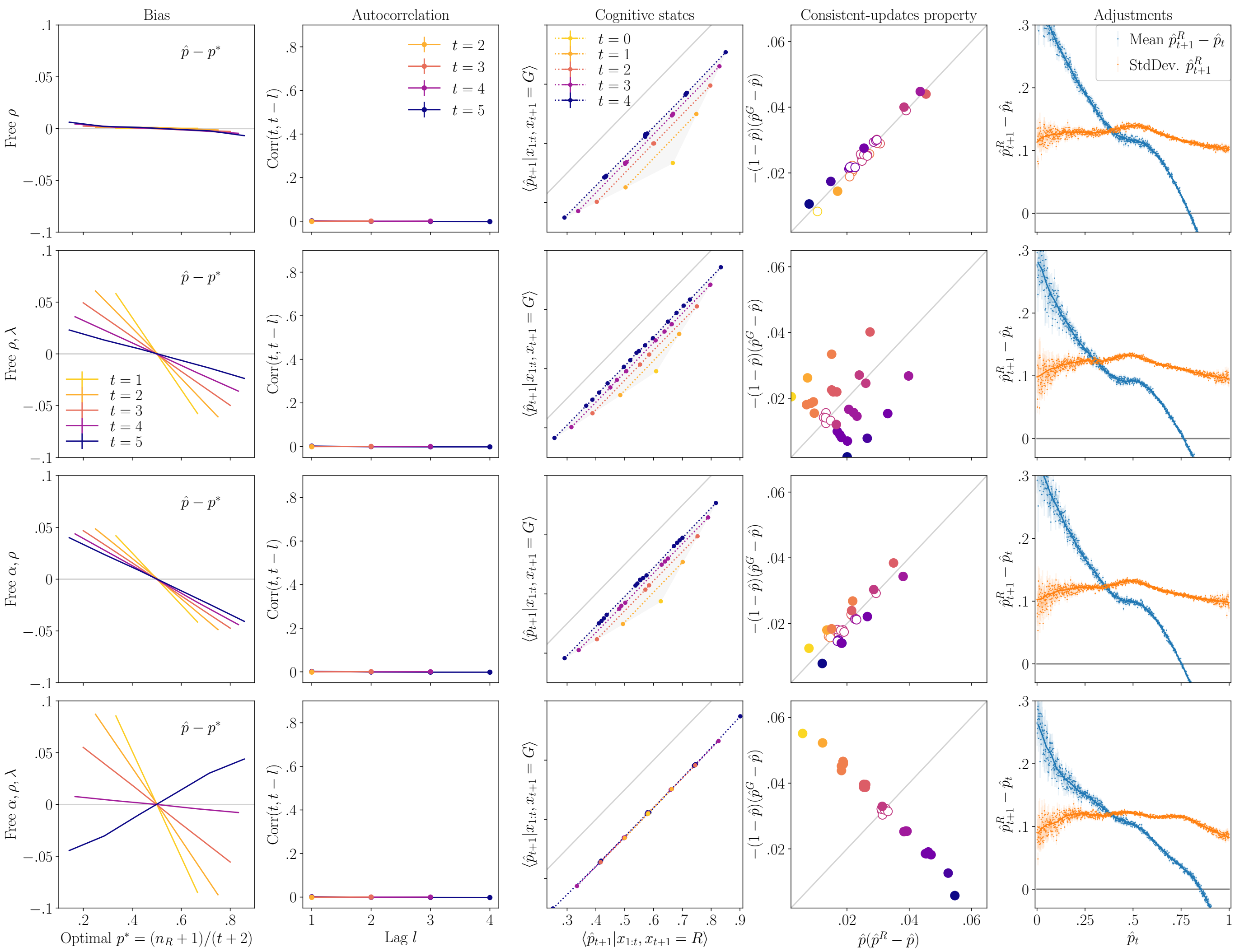

Figure 13: Behavior of the quasi-Bayesian models. As in Figure 10, with the quasi-Bayesian models with $\rho \neq 1$, and $\alpha = \lambda = 1$ (p_ρ, first row, best-fitting $\rho = 0.986$); $\alpha = 1$ ($p_{\rho\lambda}$, second row, best-fitting $\rho = 1.19$, $\lambda = 0.56$); $\lambda = 1$ ($p_{\alpha\rho}$, third row, best-fitting $\alpha = 1.47$, $\rho = 1.07$); and with the three parameters allowed to differ from 1 ($p_{\alpha\rho\lambda}$, last row, best-fitting $\alpha = 0.067$, $\rho = 1.011$, $\lambda = 0.022$).

Keeping $n_R - n_G$ constant, variations in t have a small impact on the average estimate if

$$\left| \frac{\lambda}{2} + (\alpha - 1)(\rho - 1) \right| \ll \alpha. \tag{6.23}$$

With the values best-fitted to subjects' data, the left-hand-side of the equation above is 0.0003, which indeed is well below the right-hand-side $\alpha = 0.067$ (justifying the approximation in Eq. 4.8). This suggests that the subjects' average responses vary strongly with the net difference between red and green rings, $n_R - n_G$, but weakly with the total evidence, t. The behaviors of the quasi-Bayesian models are shown in Fig. 13.

n_R	t	% $p$val < .01	median $p$val
1	2	.29	.14
1	3	.33	.19
2	3	.29	.05
1	4	.19	.11
2	4	.24	.29
3	4	.29	.08
1	5	.05	.44
2	5	.19	.15
3	5	.24	.38
4	5	.24	.20

Table 10: ANOVA F-tests of the equality of the mean responses to sequences featuring different orders of red and green rings. A low p-value suggests that one should reject the hypothesis that the mean response does not depend on the sequence order.

Order effects

Equations 4.6 and 4.7 with $\rho \neq 1$ suggest that responses may depend not only on the number of rings drawn, t, and on the number of red rings among them, n_R, but also on the order in which the red and green rings appear in the sequence. To examine whether this is the case in the responses of subjects, we run, for each pair (n_R, t) for which exist several sequences that differ only by their order, an ANOVA F-test of the null hypothesis that the mean responses to each of these sequences are all equal. We find that in most cases (characterized by n_R and t), this hypothesis is rejected for less than one third of subjects. In other words, for a majority of subjects we do not find that the order of the red and green rings in a sequence has a significant effect on the response (Table 10). In comparison, most of the effects that we exhibit in the paper are found in the responses of more than 70% of the subjects (e.g., the positive autocorrelation in responses is significant for 100% of the subjects). We note, in addition, that as the sequence length increases, so does the number of the possible orders of the red and green rings in it; as a result, the sample size available for each sequence decreases, and in many cases a subject sees a given sequence only once or twice. Small samples, in ANOVA F-tests, tend to inflate the rate of 'Type I' errors (i.e., erroneous rejections of the null hypothesis) if the variances of each population are not equal (Brown and Forsythe, 1974; Büning, 1997). We run Bartlett's tests of equality of variances for each of the groups of samples corresponding to the ten ANOVA F-tests shown in Table 10. An average of 34% of subjects reject the null hypothesis of equal variances at the 0.01 level (with a minimum, across these ten cases, of 19% of subjects). Therefore, the fractions of significant tests reported in Table 10 are likely to be overestimated.

As some subjects, although a minority, exhibit significant order effects, we examine the influence, on the responses of subjects, of the order of the rings in a sequence. For a given sequence, s, containing t samples among which n_R are red, we consider the average difference between the response to this sequence and the average response to the t-long sequences that contain n_R red rings, $\langle \hat{p}(s) - \bar{p}(n_R, t) \rangle$. We look at this quantity as a function of the average position of the red rings in the sequence (if there is only one red ring, this is simply the position of this ring). So as to compare sequences containing different number of red rings, we rescale this average position to a relative average position that ranges from -1 to 1 (if this relative position is -1, all the red rings in the sequence are at the beginning of the sequence, while if it is 1, they are all at the end). Pooling all the subjects' responses together, we find that for most lengths t and numbers of red rings n_R, the order of these red rings has no significant effect (Fig. 14, first row, dotted lines). When it has a significant effect (Fig. 14, first row, solid lines), the subjects' responses seem to exhibit a primacy effect, i.e., when the red rings in a sequence occur towards the beginning of the sequence, the responses are larger, on average, than when the red rings occur towards the end of the sequence.

Pooling together the responses of all the subjects, however, masks the heterogeneity in subjects' behaviors. First, for a majority of subjects there are no significant effects. Second, we find both primacy effects and recency effects in subjects' data. Indeed, among the 29% of subjects who exhibit a significant effect after two rings (Table 10), half exhibit a primacy effect (Fig. 14, middle row), while the other half exhibit a recency effect (i.e., the responses are larger when the red rings occur towards the end of the sequence; Fig. 14, last row).

One might conjecture that when fitting the quasi-Bayesian model with incorrect initial prior to the responses of each subject, the parameters best fitting the subjects that exhibit a recency effect or a primacy effect should result in model subjects that correspondingly exhibit a recency effect or a primacy effect. In particular, one might expect subjects with a recency effect to be best fitted by a prior weight ρ lower than 1. However, the best-fitting parameter ρ is greater than 1 for all the subjects, except one, whose best-fitting ρ is lower than 1 but who has no significant order effects. The model simulations corresponding to the other subjects (with $\rho > 1$) yield no significant order effects, or a primacy effect. Besides, 76% of subjects have best-fitting parameters such that $1 < \rho < 1.06$ and the ratio of the left- to the right-hand-side of Eq. 6.23 is lower than 0.12, i.e., parameters that, similarly to those that best fit the pooled population, imply responses that depend mostly on the difference $n_R - n_G$, and more weakly on the total evidence, t. Thus although the quasi-Bayesian model is able to capture both recency and primacy effects, it is not this aspect of data that seems the most important to capture in order to closely reproduce subjects' behavior, even for the subjects for which these effects are significant. Instead, a more important aspect of data seems to be the way that responses depend on the net count of red vs. green rings.

Behavior of the second- and third-best models

The second- and third-best-fitting models are the one with log-normal noise in the updates of the cognitive states and truncated Gaussian noise in the choice of a response, and the one in which the two types of noise are Gaussian (Table 5, lines 10 and 11). In addition, if we allow for different models for the subjects, the sizes of the cohorts of subjects whose responses are best-fitted by each of these two models are, respectively, the second and third largest (see last column of Table 5). Figure 15 shows that the five behavioral patterns identified in subjects' data are also captured by these two models.

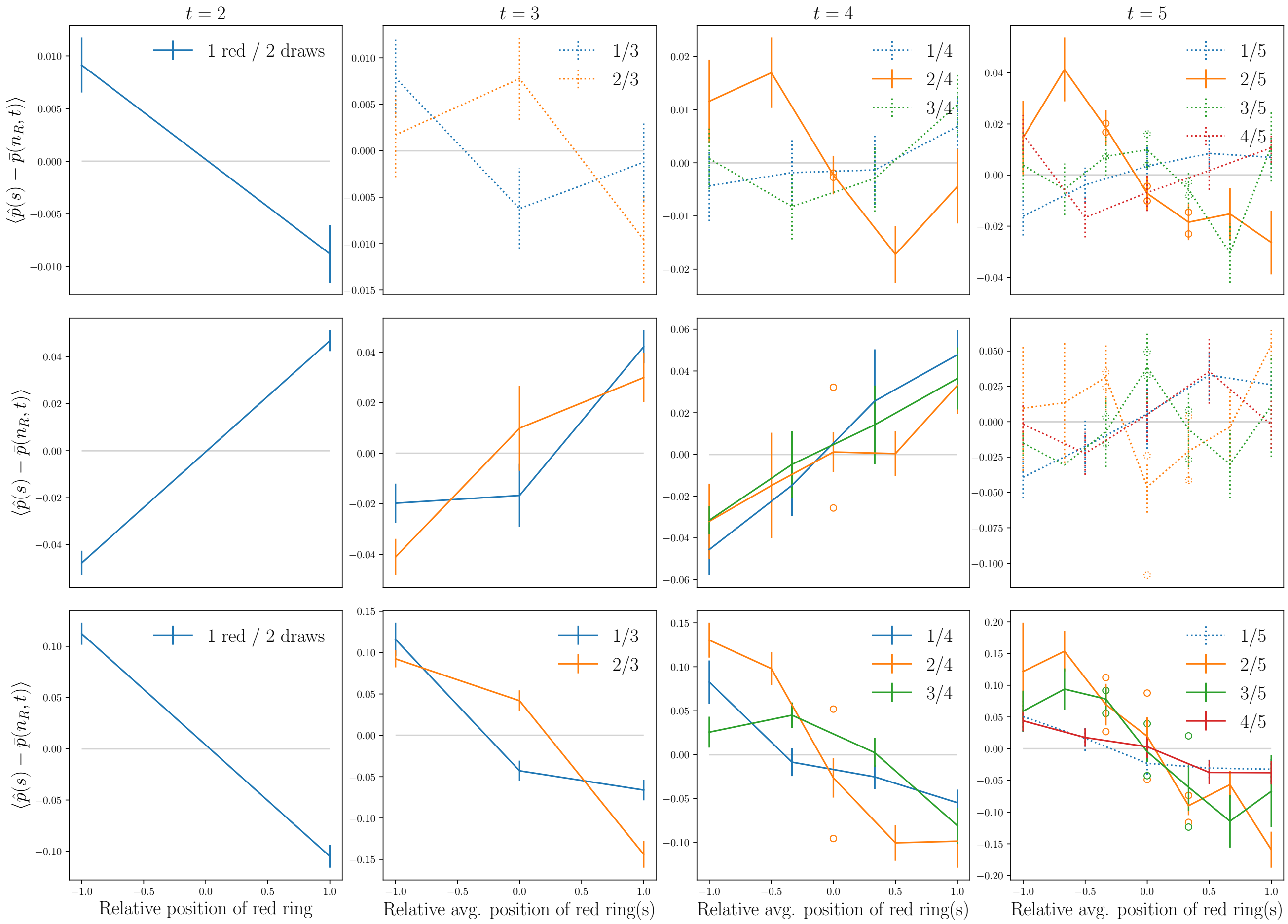

Figure 14: Order effects. Difference between the response to a sequence, $\hat{p}(s)$, and the average response to the sequences of the same length and featuring the same number of red rings, $\bar{p}(n_R, t)$, as a function of the relative average position of the red rings in the sequence; for sequences of length 2 (left column) to 5 (right column); and for all the subjects (first row), the 14.5% subjects who exhibit a significant primacy effect after two rings (middle row), and the 14.5% subjects who exhibit a significant recency effect after two rings (last row). Lines are solid when the corresponding F-test is significant to the .01 level. In some cases, sequences of the same length t and same number of red rings n_R have their red rings in a different order, but nevertheless have the same average position of red rings (and thus the same abscissa, here): in these cases, the average response to each different sequence is shown with a circle.

Calibrated-response property in models

In Figure 16, we look at whether the responses of the models, fitted to subjects' data, verify the Bayesian calibrated-response property (Eq. 3.2). To emphasize any departure from this property, we consider the empirical average of the difference between the true proportion and the response, $p - \hat{p}$, as a function of the response, $\hat{p}$ (instead of the true proportion p

as a function of $\hat{p}$, as in Fig. 7). We find that none of the models we consider verify the property, and all the models differ from the prediction of the property in ways qualitatively similar to that of the subjects.

Indeed, for extreme values of the response $\hat{p}$, the model subjects exhibit 'over-confidence' (for responses close to 1, the average true proportion is below the response, and conversely for responses close to 0), while the opposite is found for responses close to the middle value: e.g., for responses just above .5 the true proportion is on average larger than the response. The latter effect, however, is quantitatively small in the case of the model in which responses are the correct Bayesian estimates, but with response error added (Fig. 16, top middle panel). It is larger (and closer to the effect found in subjects' data) with the model of Bayesian estimation starting from an incorrect prior (and truncated Gaussian noise in responses; Fig. 16, top right panel), and it is the most similar to subjects' data in the case of the best-fitting model (which includes Gaussian updates of the cognitive state and truncated log-normal responses; Fig. 16, bottom right panel).

Experiment variant with geometrically-distributed sequence lengths

In our experiment, the subjects know that five draws will be shown in each block of trials. We run a control experiment, in which the subjects do not know in advance how many samples they will observe in a block of trials: instead, after each presentation of a new ring there is a 20% probability that the sequence ends, and an 80% probability that the sequence continues (i.e., that an additional ring is drawn). In other words, the sequence length is a geometric random variable with parameter 0.2. All other aspects of the experiment have

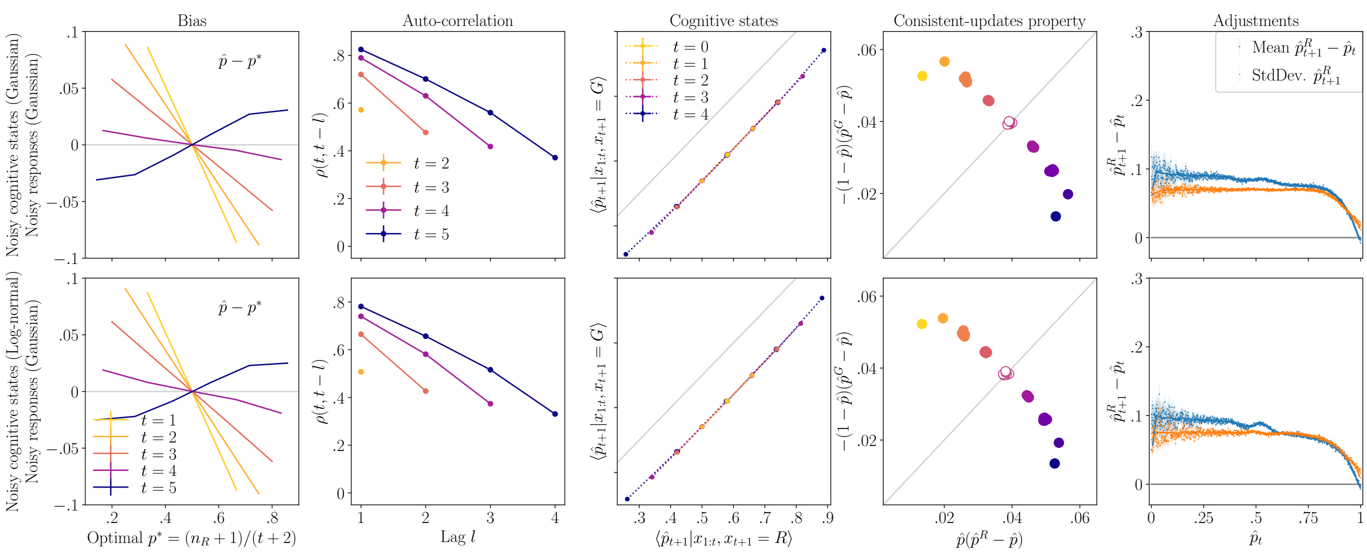

Figure 15: Behavior of the models with log-normal or Gaussian cognitive states, and truncated Gaussian response noise. As in Figure 10, with the model with Gaussian states and truncated Gaussian responses (third-best model, first row) and the model with log-normal states and truncated Gaussian responses (second-best model, second row).

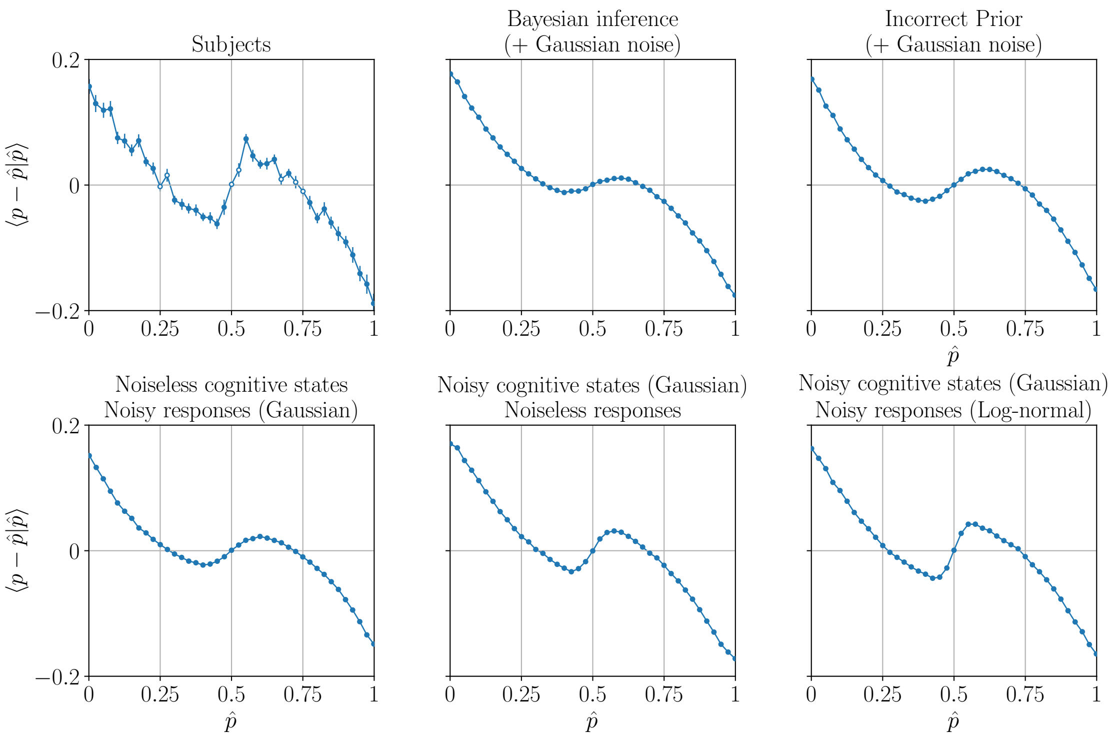

Figure 16: Subjects' and model subjects' responses do not verify the calibrated-response property. Empirical average of the difference between the true proportion and the estimate, $p - \hat{p}$, conditional on the estimate $\hat{p}$, for the subjects and for five models with the parameters best-fitting subjects' data. *Top left:* subjects' data. *Top middle:* model in which the responses derive from optimal, Bayesian inference, with in addition Gaussian truncated noise. *Top right:* model with Bayesian inference on the basis of an incorrect prior, and Gaussian truncated noise. *Bottom left:* model with a cognitive state that is a precise, deterministic count of the net excess of red over green rings, and with truncated Gaussian response error. *Bottom middle:* noisy counting model, with no response error. *Bottom right:* best-fitting model, with Gaussian updates of the cognitive states, and log-normal noise in responses.

been maintained identical. Seventeen subjects, aged 18 to 37 (average 21.9), participated in this variant of the experiment, in the same conditions as those of the main experiment (see details above). As with the main experiment, we excluded from our analyses four subjects whose average absolute error (0.226, standard deviation: 0.021) was significantly higher than that of the other subjects (0.168, standard deviation: 0.018). The analyses below are thus based on the responses of 13 subjects.

So as to be able to compare the behavior of subjects in the main experiment and in the variant, we look, first, at their responses to the sequences of up to five observations. We find that the subjects' responses are not a monotonic function of the optimal, Bayesian estimates (Fig. 17A, blue line): as in the main experiment, the average response when the optimal estimate is $5/7$ is higher than when it is $3/4$ (although $5/7 < 3/4$), and similarly for

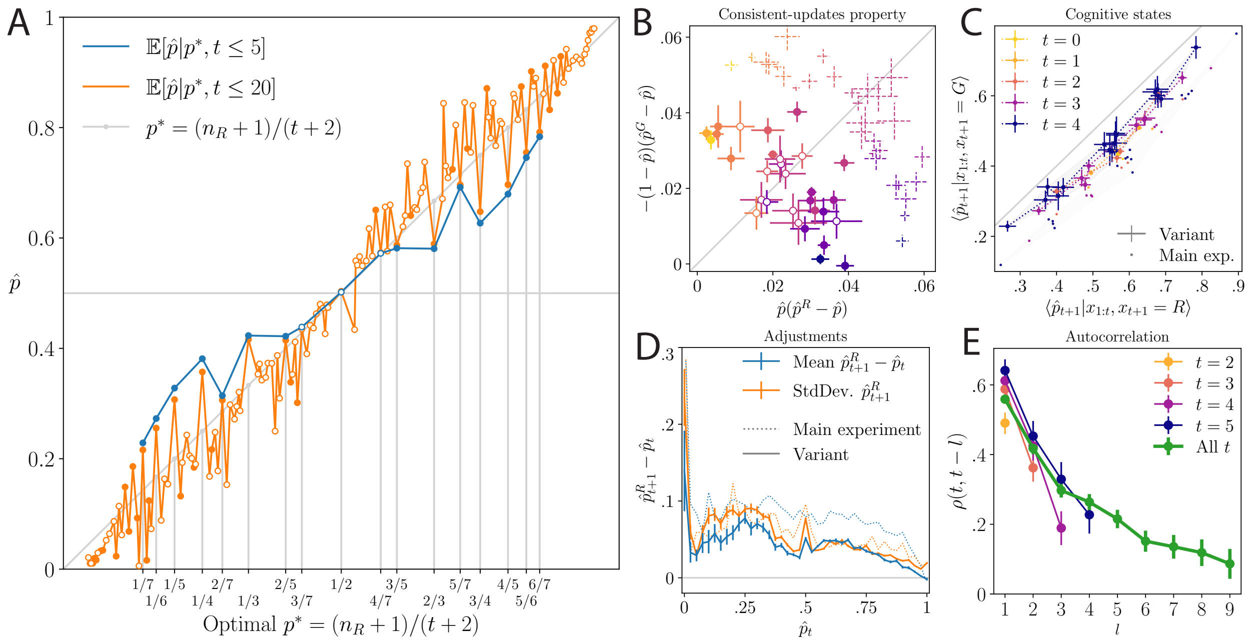

Figure 17: Subjects' behavior in the experiment variant with geometrically-distributed sequence lengths. (A) Subjects' responses as a function of the optimal estimate, with sequences of up to five observations (blue line; compare with Fig. 3), and up to 20 observations (orange line). The filled dots indicate where the p-value of a t-test of equality with p^* is below .05. (B) Consistent-updates Bayesian property: left-hand side vs. right-hand side of Eq. 3.9. Filled points indicate where the t-test of equality is rejected at the .01 level. The dashed crosses show the data from the main experiment (same as in Fig. 8A). (C) Average responses after observing the sequence $x_{1:t}$ and a green ring ($x_{t+1} = G$), vs. average responses after observing the same sequence ($x_{1:t}$) and a red ring ($x_{t+1} = R$). For comparison, the small dots show the data obtained in the main experiment (same as in Fig. 6, bottom right). (D) Averages (blue) and standard deviations (orange) of the adjustments in responses following a red ring, $\hat{p}^R_{t+1} - \hat{p}_t$, as a function of the preceding responses, $\hat{p}_t$. The bars show the standard error of the mean and of the standard deviation. The dotted lines show the data obtained in the main experiment (same as in Fig. 9B). (E) Coefficients of correlation between subjects' excursions at trial t and subjects' excursions at a preceding trial $t - l$, with t ranging from 2 to 5 (compare with Fig. 5B), and with all trials $t > l$ pooled together. Error bars indicate the 90% confidence interval.

the symmetrical estimates 1/4 and 2/7. The curves obtained in the main experiment and in the variant have very similar shapes (compare the blue lines in Fig. 3 and in Fig. 17A). A difference, however, is that in the main experiment the average responses are sometimes more extreme (closer to 0 and 1) than the optimal estimates, and sometimes less extreme (closer to .5), while in the variant the average responses (for sequences up to length 5) are never more extreme than the optimal estimates.

The similar but less-extreme responses suggest that the decision-making processes of the subjects, in the main experiment and in the variant, are identical, with the exception that in the variant the subjects choose to make adjustments of smaller sizes. And indeed, turning to the other behavioral patterns that we have examined in this study, we find that qualitatively, they all look similar (between the main experiment and the variant), while the quantitative differences all seem to mainly result from the smaller adjustments adopted by the subjects, in the variant. For instance, the consistent-update property, i.e., the equality between the probability-weighted adjustments after a green ring, $\delta^G = -(1-\hat{p}_t)(\hat{p}^G_{t+1}-\hat{p}_t)$, and after a red ring, $\delta^G = \hat{p}_t(\hat{p}^R_{t+1}-\hat{p}_t)$, is violated in the same way in the main experiment and in its variant (Fig. 17B). But in the variant, the probability-weighted adjustments are smaller (resulting in points closer to the origin in Fig. 17B). Smaller adjustments also imply that the responses after a red ring and those after a green ring are closer to each other in the variant than they are in the main experiment (resulting in points closer to the identity line in Fig. 17C). And a direct look at the subjects' adjustments, $\hat{p}^R_{t+1} - \hat{p}_t$, shows that they depend weakly on the preceding responses, in both the variant and the main experiment, but in the variant they are smaller (Fig. 17D). (We note that the standard deviations are similar, suggesting (unsurprisingly) similar degrees of cognitive noise in the two experiments.) Finally, we find in the data of the variant experiment the same autocorrelation in responses that we find in the main one (Fig. 17E).

From the comparison of the behavior of subjects in the main experiment and in the variant, we conclude, first, that the subjects process the observations and make their decisions in the same way, in the two experimental settings; and second, that they choose to make smaller adjustments in the variant than in the main experiment. As for this difference in the adjustment sizes, the analysis presented in the Discussion suggests that it results from an efficient adaptation to the different statistics of the sequence lengths in the two cases (see Fig. 11).

In the main experiment, the adjustment size chosen by the subjects results in their conservatism after a few observations, followed by their overreaction to the evidence after, for instance, five draws. The smaller adjustments, in the variant experiment, result in conservatism even after five draws (Fig. 18, solid lines). But as evidence accumulates, the Bayesian observer makes smaller and smaller adjustments to its estimate (Eq. 5.1), such that the adjustment size chosen by the subjects in the variant experiment, although small, may after a number of observations be too large, and result in overreaction to the evidence. Although for comparison purposes we first limited our analysis to sequences of up to five draws, we now examine subjects' responses after longer sequences, in order to investigate whether they eventually exhibit overreaction to the evidence. We find that the behavioral data indeed exhibits a reversal of the conservatism bias after long sequences: after 12 observations and more, there is a clear pattern of overreaction to the evidence, i.e., subjects' responses are more extreme than the optimal estimates (Fig. 18). Thus here also, we obtain in the variant experiment a behavior very similar to that obtained in the main experiment (Fig. 4), but the reversal of the conservatism bias occurs after longer sequences, because of subjects' smaller adjustments.

We note also that the pattern of decreasing autocorrelation continues after five draws (Fig. 17E). The correlation between a response at some trial and the response 9 trials later is significantly positive (Pearson correlation coefficient: 0.087, p-value: 0.001).

Finally, we look again, in Fig. 17A, at the subjects' average responses as a function of the optimal estimates, but this time including the sequences of up to 20 draws (this represents

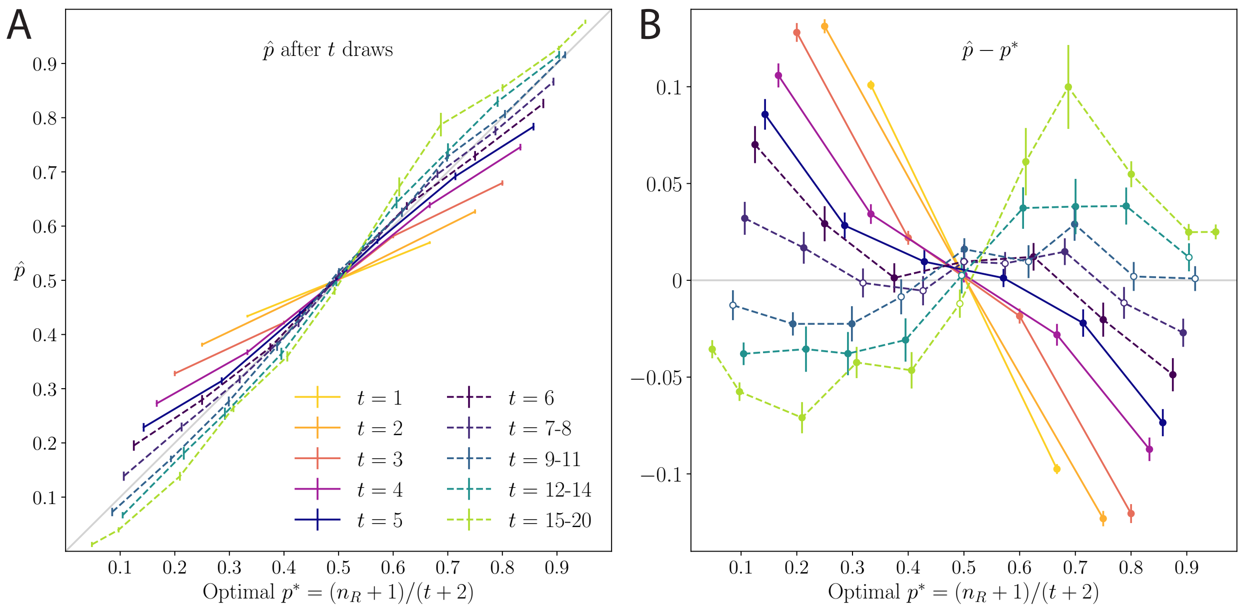

Figure 18: Bias reversal in the experiment variant with geometrically-distributed sequence lengths. (A) Subjects' response $\hat{p}$, and (B) difference $\hat{p} - p^*$ with optimal estimate, as a function of the optimal estimate, for different numbers t of rings shown. Compare to Fig. 4. As the amount of available data decreases with the sequence length, from length 7 we pool together the responses to sequences of similar lengths. In (B), the filled dots indicate where the p-value of a t-test of equality with zero is below .05.

more than 98% of the blocks of trials). The responses of subjects seem to oscillate above and below the value of the optimal estimate (Fig. 17A, orange line). These oscillations do not result from sampling noise: for many optimal estimates, the responses are significantly different from the optimal estimate (sometimes greater, sometimes lower). This emphasizes the importance of examining carefully the responses to different sequences of observations, instead of pooling together the responses to sequences that are very different but that imply optimal estimates which are close in value, as in Fig. 1B. Still, an overall view of subjects' responses, in Fig. 17A, seems to suggest that they are well aligned with the optimal estimates. We note that in the main experiment, the average responses of subjects also seemed to be oscillating around the optimal estimates, with significant differences, but on the whole following the identity line (Fig. 3). These results substantiate the idea that the subjects, in the two versions of the task, efficiently adapt their otherwise suboptimal decision process to the different statistics of the sequence lengths.

Noisy-counting model with log-normal response selection: simulations and individual best-fitting parameters

The best-fitting model for most subjects, which features log-normal response selection (line 12 in Table 5), has three parameters: μ and σ the mean and the standard deviation of the cognitive-state updates; and ν, the noise parameter of the log-normal distribution of slider adjustments. For each subject, we find the three parameters that maximize the log-likelihood of the subject's responses. The mean value of μ across subjects is 0.082 (median: 0.077, standard deviation (SD): 0.019); mean value of σ: 0.041 (median: 0.034, SD: 0.019); mean value of ν: 1.43 (median: 0.19, SD: 3.67).

Given μ, σ and ν, we compute as follows the MSE of the estimates of a model subject: for each of the 101 values of the true probability ranging from 0 to 1 by increment of 0.01, we sample M blocks of five rings, drawn with the given probability, and simulate, for a model subject parameterized by μ, σ and ν, the stochastic responses to these draws. The average of the squared errors of all these responses provides a numerical approximation to the MSE. Finally, to obtain the MSE as a function of μ in Figure 11, we repeat this procedure with an array of values of μ ranging from 0 to 0.25, and with each subject's best-fitting values of σ and ν. The resolution of the values of μ where the MSE is evaluated is higher (increments of 0.001) close to the optimum. Similarly, M ranges from 200 to 1000 or more near the optimal choice of μ.

Data availability

Requests for the data can be sent via email to the corresponding author.

Code availability

Requests for the code used for all analyses can be sent via email to the corresponding author.

References

Afrouzi, H., Kwon, S. Y., Landier, A., Ma, Y., & Thesmar, D. (2020). Overreaction and Working Memory. *NBER Working Paper*, (27947), 1–67. https://doi.org/10.3386/w27947

Anderson, N. H. (1991). *Contributions To Information Integration Theory*. Psychology Press. https://doi.org/10.4324/9781315807331

Ashinoff, B. K., Buck, J., Woodford, M., & Horga, G. (2022). Effects of base rate neglect on sequential belief updating and real-world beliefs. *PsyArXiv*. https://doi.org/10.31234/osf.io/8ey75

Augenblick, N., Lazarus, E., & Thaler, M. (2023). Overinference from Weak Signals and Underinference from Strong Signals. *MIT Sloan Research Paper*. https://doi.org/10.2139/ssrn.4315007

Azeredo da Silveira, R., Sung, Y., & Woodford, M. (2020). Optimally Imprecise Memory and Biased Forecasts. *CESifo Working Paper Series*, 85. https://doi.org/10.3386/w28075

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of behavioral economics* (pp. 69–186). Elsevier B.V. https://doi.org/10.1016/bs.hesbe.2018.11.002

Benjamin, D. J., Bodoh-Creed, A., & Rabin, M. (2019). Base-Rate Neglect: Foundations and Implications. *Working paper*, 1–62.

Best, M. J., & Chakravarti, N. (1990). Active set algorithms for isotonic regression; A unifying framework. *Mathematical Programming*, *47*(1-3), 425–439. https://doi.org/10.1007/BF01580873

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in Macroeconomic Expectations. *American Economic Review*, *110*(9), 2748–2782. https://doi.org/10.1257/aer.20181219

Bordalo, P., Gennaioli, N., Porta, R. L., & Shleifer, A. (2019). Diagnostic Expectations and Stock Returns. *Journal of Finance*, *74*(6), 2839–2874. https://doi.org/10.1111/jofi.12833

Brown, E. R., & Bane, A. L. (1975). Probability estimation in a chance task with changing probabilities. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 183–187. https://doi.org/10.1037/0096-1523.1.2.183

Brown, M. B., & Forsythe, A. B. (1974). The Small Sample Behavior of Some Statistics Which Test the Equality of Several Means. *Technometrics*, *16*(1), 129. https://doi.org/10.2307/1267501

Büning, H. (1997). Robust analysis of variance. *Journal of Applied Statistics*, *24*(3), 319–332. https://doi.org/10.1080/02664769723710

Caplin, A., Csaba, D., Leahy, J., & Nov, O. (2020). Rational Inattention, Competitive Supply, and Psychometrics. *Quarterly Journal of Economics*, *135*(3), 1681–1724. https://doi.org/10.1093/qje/qjaa011

Caplin, A., & Dean, M. (2015). Revealed Preference, Rational Inattention, and Costly Information Acquisition. *American Economic Review*, *105*(7), 2183–2203. https://doi.org/10.1257/aer.20140117

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/https://doi.org/10.1016/j.jbef.2015.12.001

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

Dawid, A. P. (1982). The Well-Calibrated Bayesian. *Journal of the American Statistical Association*, *77*(379), 605–610. https://doi.org/10.1080/01621459.1982.10477856

Dehaene, S., & Marques, J. F. (2002). Cognitive euroscience: Scalar variability in price estimation and the cognitive consequences of switching to the euro. *The Quarterly Journal of Experimental Psychology Section A*, *55*(3), 705–731. https://doi.org/10.1080/02724980244000044

Donnell, M. L., & Du Charme, W. M. (1975). The effect of Bayesian feedback on learning in an odds estimation task. *Organizational Behavior and Human Performance*, *14*(3), 305–313. https://doi.org/10.1016/0030-5073(75)90032-X

Drugowitsch, J., Wyart, V., Devauchelle, A.-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

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes. *Psychological Review*, *101*(3), 519–527. https://doi.org/10.1037/0033-295X.101.3.519

Feldman, J. (2015). Bayesian Models of Perceptual Organization. In J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization* (pp. 1008–1026). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199686858.013.007

Findling, C., Chopin, N., & Koechlin, E. (2021). Imprecise neural computations as a source of adaptive behaviour in volatile environments. *Nature Human Behaviour*, *5*(1), 99–112. https://doi.org/10.1038/s41562-020-00971-z

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., & Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature Neuroscience*, *22*(December). https://doi.org/10.1038/s41593-019-0518-9

Gallistel, C. R., Krishan, M., Liu, Y., Miller, R., & Latham, P. E. (2014). The perception of probability. *Psychological Review*, *121*(1), 96–123. https://doi.org/10.1037/a0035232

Ganguli, D., & Simoncelli, E. P. (2010). Implicit encoding of prior probabilities in optimal neural populations. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 23* (pp. 658–666). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2010/file/f9a40a4780f5e1306c46f1c8daecee3b-Paper.pdf

Glaze, C. M., Filipowicz, A. L. S., Kable, J. W., Balasubramanian, V., & Gold, J. I. (2018). A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nature Human Behaviour*, *2*(3), 213–224. https://doi.org/10.1038/s41562-018-0297-4

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125. https://doi.org/10.1007/s40881-015-0004-4

Grether, D. M. (1980). Bayes Rule as a Descriptive Model: The Representativeness Heuristic. *The Quarterly Journal of Economics*, *95*(3), 537. https://doi.org/10.2307/1885092

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*(3), 411–435. https://doi.org/10.1016/0010-0285(92)90013-R

Griffiths, T. L., & Tenenbaum, J. B. (2011). Predicting the future as Bayesian inference: People combine prior knowledge with observations when estimating duration and extent. *Journal of Experimental Psychology: General*, *140*(4), 725–743. https://doi.org/10.1037/a0024899

Hilbert, M. (2012). Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin*, *138*(2), 211–237. https://doi.org/10.1037/a0025940

Kahneman, D., & Tversky, A. (1972). Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology*, *3*(3), 430–454. https://doi.org/10.1016/0010-0285(72)90016-3

Khaw, M. W., Stevens, L., & Woodford, M. (2017a). Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*, *91*, 88–103. https://doi.org/10.1016/j.jmoneco.2017.09.001

Khaw, M. W., Stevens, L., & Woodford, M. (2017b). Forecasting the outcome of a time-varying bernoulli process: Data from a laboratory experiment. *Data in Brief*, *15*, 469–473. https://doi.org/10.1016/j.dib.2017.10.007

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719. https://doi.org/10.1016/j.tins.2004.10.007

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, *10*(7), 319–326. https://doi.org/10.1016/j.tics.2006.05.003

Laming, D. (2011). *Human judgment: The eye of the beholder* (2nd Ed.). Cengage Learning.

Maćkowiak, B., Matějka, F., & Wiederholt, M. (2020). Rational Inattention: A Review. *CEPR Discussion Papers*, (DP15408), 1–69. https://home.cerge-ei.cz/matejka/RIsurvey.pdf

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, *10*(1), 6–38. https://doi.org/10.1006/game.1995.1023

Meyniel, F., Maheu, M., & Dehaene, S. (2016). Human Inferences about Sequences: A Minimal Transition Probability Model. *PLoS Computational Biology*, *12*(12), 1–26. https://doi.org/10.1371/journal.pcbi.1005260

Morais, M., & Pillow, J. W. (2018). Power-law efficient neural codes provide general link between perceptual bias and discriminability. *Advances in Neural Information Processing Systems 31*, *2*(1), 5076–5085. http://papers.nips.cc/paper/7754-power-law-efficient-neural-codes-provide-general-link-between-perceptual-bias-and-discriminability.pdf

Mori, S., & Ward, L. M. (1995). Pure feedback effects in absolute identification. *Perception & Psychophysics*, *57*(7), 1065–1079. https://doi.org/10.3758/BF03205465

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, *29*(3), 315–335. https://doi.org/10.2307/1909635

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*, *30*(37), 12366–12378. https://doi.org/10.1523/JNEUROSCI.0822-10.2010

Nemenman, I., Shafee, F., & Bialek, W. (2002). Entropy and Inference, Revisited. In *Advances in neural information processing systems 14* (pp. 1–9). The MIT Press. https://doi.org/10.7551/mitpress/1120.003.0065

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural Computation*, *15*(6), 1191–1253. https://doi.org/10.1162/089976603321780272

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, *68*(1), 29–46. https://doi.org/10.1037/h0024722

Peterson, C. R., & Miller, A. J. (1965). Sensitivity of subjective probability revision. *Journal of Experimental Psychology*, *70*(1), 117–121. https://doi.org/10.1037/h0022023

Peterson, C. R., Ulehla, Z. J., Miller, A. J., Bourne, L. E., & Stilson, D. W. (1965). Internal consistency of subjective probabilities. *Journal of Experimental Psychology*, *70*(5), 526–533. https://doi.org/10.1037/h0022546

Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences*, *19*(5), 285–293. https://doi.org/10.1016/j.tics.2015.03.002

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, *72*(3), 346–354. https://doi.org/10.1037/h0023653

Phillips, L., Hays, W., & Edwards, W. (1966). Conservatism in Complex Probabilistic Inference. *IEEE Transactions on Human Factors in Electronics*, *HFE-7*(1), 7–18. https://doi.org/10.1109/THFE.1966.231978

Prat-Carrabin, A., Meyniel, F., & Azeredo da Silveira, R. (2022). Resource-Rational Account of Sequential Effects in Human Prediction. *bioRxiv*. https://doi.org/10.1101/2022.06.20.496900

Prat-Carrabin, A., Wilson, R. C., Cohen, J. D., & Azeredo da Silveira, R. (2021). Human inference in changing environments with temporal structure. *Psychological Review*, *128*(5), 879–912. https://doi.org/10.1037/rev0000276

Prat-Carrabin, A., & Woodford, M. (2022). Bias and variance of the Bayesian-mean decoder. *Advances in neural information processing systems 34*, (NeurIPS). https://proceedings.neurips.cc/paper/2021/file/c7c3e78e3c9d26cc1158a8735d548eaa-Paper.pdf

Renart, A., & Machens, C. K. (2014). Variability in neural activity and behavior. *Current Opinion in Neurobiology*, *25*, 211–220. https://doi.org/10.1016/j.conb.2014.02.013

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shanteau, J. C. (1970). An additive model for sequential decision making. *Journal of Experimental Psychology*, *85*(2), 181–191. https://doi.org/10.1037/h0029552

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

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

Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute Identification by Relative Judgment. *Psychological Review*, *112*(4), 881–911. https://doi.org/10.1037/0033-295X.112.4.881

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Ward, L. M., & Lockhead, G. R. (1971). Response system processes in absolute judgment. *Perception & Psychophysics*, *9*(1), 73–78. https://doi.org/10.3758/BF03213031

Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. *Nature Neuroscience*, *18*(10), 1509–1517. https://doi.org/10.1038/nn.4105

Westbrook, A., Kester, D., & Braver, T. S. (2013). What Is the Subjective Cost of Cognitive Effort? Load, Trait, and Aging Effects Revealed by Economic Preference (M.

Pessiglione, Ed.). *PLoS ONE*, *8*(7), e68210. https://doi.org/10.1371/journal.pone.0068210
Whalen, J., Gallistel, C., & Gelman, R. (1999). Nonverbal Counting in Humans: The Psychophysics of Number Representation. *Psychological Science*, *10*(2), 130–137. https://doi.org/10.1111/1467-9280.00120
Wilson, R. C., Nassar, M. R., & Gold, J. I. (2013). A Mixture of Delta-Rules Approximation to Bayesian Inference in Change-Point Problems (T. Behrens, Ed.). *PLoS Computational Biology*, *9*(7), e1003150. https://doi.org/10.1371/journal.pcbi.1003150
Wolpert, D. H., & Wolf, D. R. (1995). Estimating functions of probability distributions from a finite set of samples. *Physical Review E*, *52*(6), 6841–6854. https://doi.org/10.1103/PhysRevE.52.6841
Wyart, V., & Koechlin, E. (2016). Choice variability and suboptimality in uncertain environments. *Current Opinion in Behavioral Sciences*, *11*, 109–115. https://doi.org/10.1016/j.cobeha.2016.07.003
Yu, A. J., & Cohen, J. D. (2008). Sequential effects: Superstition or rational behavior? *Advances in neural information processing systems*, *21*, 1873–1880. http://www.ncbi.nlm.nih.gov/pubmed/26412953
Zhang, H., & Maloney, L. T. (2012). Ubiquitous Log Odds: A Common Representation of Probability and Frequency Distortion in Perception, Action, and Cognition. *Frontiers in Neuroscience*, *6*(JAN), 1–14. https://doi.org/10.3389/fnins.2012.00001
Zhang, H., Ren, X., & Maloney, L. T. (2020). The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences*, *117*(36), 22024–22034. https://doi.org/10.1073/pnas.1922401117
Zhu, J.-Q., Sanborn, A. N., & Chater, N. (2020). The Bayesian sampler: Generic Bayesian inference causes incoherence in human probability judgments. *Psychological Review*, *127*(5), 719–748. https://doi.org/10.1037/rev0000190