



Efficient stabilization of imprecise statistical inference through conditional belief updating

Julie Drevet^{1,2}✉, Jan Drugowitsch³ and Valentin Wyart^{1,2}✉

Statistical inference is the optimal process for forming and maintaining accurate beliefs about uncertain environments. However, human inference comes with costs due to its associated biases and limited precision. Indeed, biased or imprecise inference can trigger variable beliefs and unwarranted changes in behaviour. Here, by studying decisions in a sequential categorization task based on noisy visual stimuli, we obtained converging evidence that humans reduce the variability of their beliefs by updating them only when the reliability of incoming sensory information is judged as sufficiently strong. Instead of integrating the evidence provided by all stimuli, participants actively discarded as much as a third of stimuli. This conditional belief updating strategy shows good test-retest reliability, correlates with perceptual confidence and explains human behaviour better than previously described strategies. This seemingly suboptimal strategy not only reduces the costs of imprecise computations but also, counterintuitively, increases the accuracy of resulting decisions.

Efficient decision-making about the cause of noisy or ambiguous observations requires the accumulation of multiple pieces of evidence to form accurate beliefs^{1,2}, a process typically referred to as ‘statistical inference’. In stable environments, accumulating evidence across observations reduces uncertainty about their cause. There is ample experimental evidence that humans and other animals perform near-optimal inference in such conditions. However, human inference is subject to two main sources of internal noise: sensory noise, which limits the amount of evidence provided by each observation³, and computation noise, which arises during the imprecise combination of incoming sensory evidence with the current belief^{4–6}. Both sources of noise trigger trial-to-trial variability in beliefs and behaviour.

In volatile environments, the cause of noisy observations changes over time, and statistical inference requires the appropriate weighting of the current belief against incoming sensory evidence⁷. Recent findings suggest that humans and other animals are capable of performing near-optimal inference even in volatile environments⁸. However, in these conditions, the trial-to-trial variability in beliefs triggered by internal sources of noise has even larger costs for efficient decision-making. Indeed, noisy sensory observations may be mistaken for genuine changes in their latent cause. Similarly, imprecise inference may trigger unwarranted changes-of-mind when incoming sensory evidence is combined imprecisely with internal beliefs. Whether and how humans may mitigate these important cognitive costs in volatile environments remain unknown.

Here we addressed this question by studying human statistical inference in a volatile decision-making task based on noisy visual stimuli (Fig. 1). Tested participants ($n = 60$ across two experiments) were presented with sequences of marbles which could be drawn either from a light bag containing dominantly light marbles or a dark bag containing dominantly dark marbles (Fig. 1a). The light and dark areas of each marble were spatially scrambled, and the light/dark fractions of dominantly light and dark marbles were adjusted using an adaptive titration procedure (Fig. 1b), such that 20% of dominantly light marbles (that is, marbles from the light

bag) were perceived as dark and vice versa (Fig. 1c and Methods). After each marble, participants were asked to identify the bag from which it was drawn (Fig. 1d). Importantly, marbles were not drawn randomly and independently across successive trials, but rather in episodes of multiple draws from the same bag. Decision-making in the presence of such temporal structure can benefit from statistical inference that integrates uncertain visual information provided by each new marble with internal beliefs about the bag that is currently being drawn from.

To analyse human behaviour in this task, we developed a process-level model of statistical inference derived from normative Bayesian computations⁷, which hypothesizes distinct sources of internal noise at sensory, inference and response selection stages⁴. Across two experiments, we fitted this model to human behaviour to determine whether and how tested participants adopted strategies to compensate for the costs of imprecise statistical inference identified above. To do so, we formulated and compared different strategies by which participants could mitigate the variability of beliefs formed through statistical inference. We obtained converging evidence that human observers control statistical inference in a cost-efficient fashion, by updating their internal beliefs based on incoming inputs only when the noisy sensory observations are deemed sufficiently reliable. This metacognitive strategy not only reduces the costs of internal sources of variability affecting inference, but also, counterintuitively, increases the accuracy of resulting decisions.

Results

Characterizing the suboptimality of statistical inference. We first assessed participants’ performance by comparing their reversal behaviour to the one predicted by the optimal Bayesian inference process⁷ (Methods, Eqs. 1 and 2). Importantly, this normative process accounts for sensory errors due to visual noise. As expected from previous work^{4,5}, optimal inference substantially outperformed participants (Fig. 1e,g; overall accuracy, optimal inference: 0.870; participants: 0.769 ± 0.011 , mean \pm s.e.m., paired two-sided t -test, difference: $t_{28} = 9.5$, $P < 0.001$, Cohen’s $d = 1.858$, 95% confidence

¹Laboratoire de Neurosciences Cognitives et Computationnelles, Institut National de la Santé et de la Recherche Médicale (Inserm), Paris, France.

²Département d’Études Cognitives, École Normale Supérieure, Université PSL, Paris, France. ³Department of Neurobiology, Harvard Medical School, Boston, MA, USA. ✉e-mail: julie.drevet@ens.fr; valentin.wyart@ens.fr

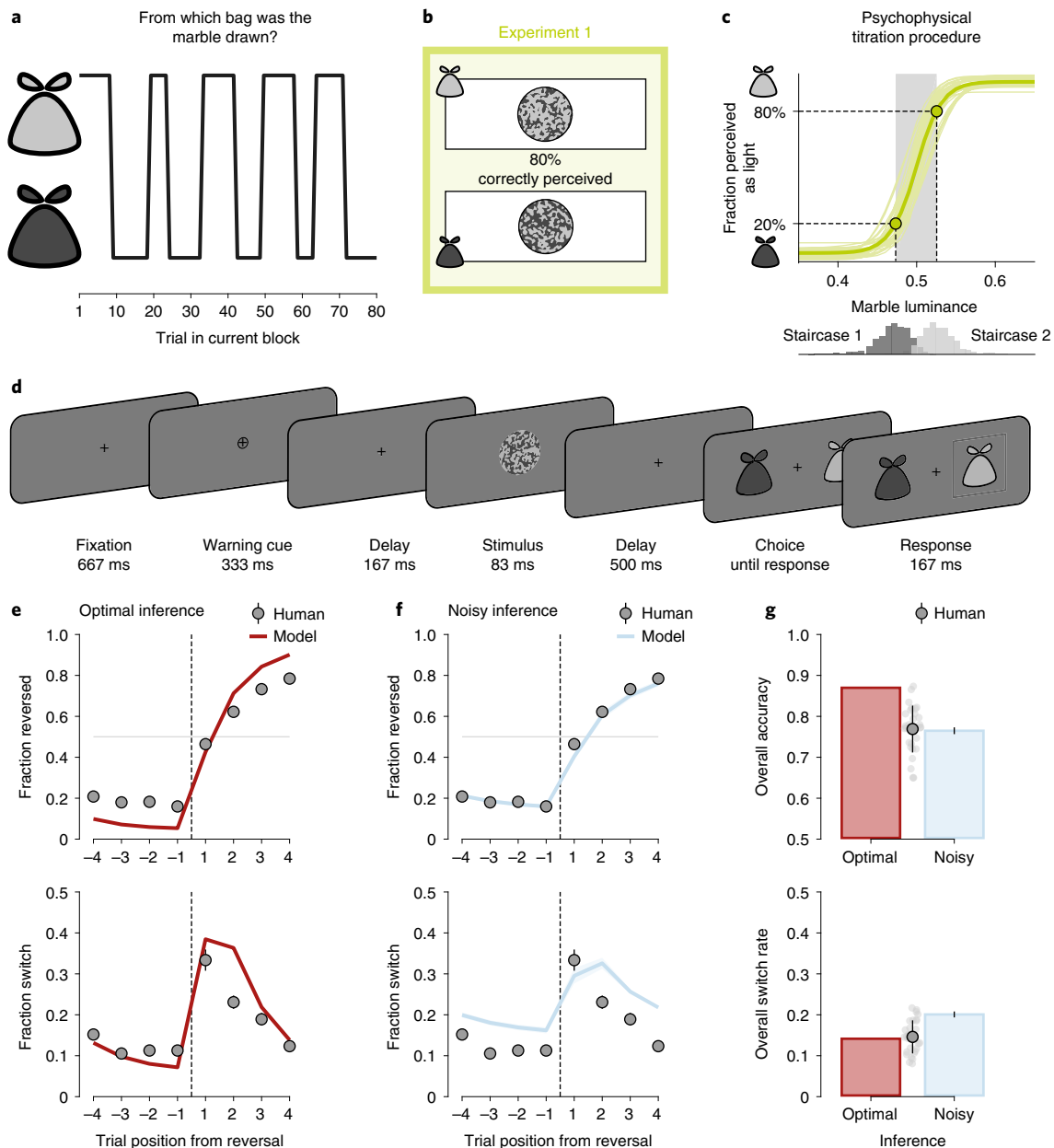


Fig. 1 | Description of experiment 1 ($n = 30$). **a**, Structure of the volatile decision-making task. Each block of trials features alternations between draws from the light and dark bags. **b**, Visual stimuli. The overall marble luminance is titrated such that 20% of presented marbles are miscategorized. **c**, Adaptive titration procedure. Top: psychometric curves estimated by the titration procedure (thin lines, participants; thick line, group-level average). Bottom: distributions of marble luminance presented by the staircases titrating the dark and light marbles. **d**, Trial description: 500 ms after a warning cue, a marble stimulus is presented for 83 ms, after which the participant indicates the bag from which marbles are currently drawn. **e**, Predictions of optimal inference. Top: response reversal curves indicating the fraction of correct responses towards the new drawn bag surrounding each reversal. Bottom: response switch curves indicating the fraction of trial-to-trial response switches surrounding the same reversals. The reversal is represented by the thin dotted line. Dots indicate human data (group-level average), whereas lines indicate predictions of optimal inference. Error bars correspond to s.e.m. **f**, Predictions of the noisy inference model. Top: response reversal curve predicted by the best-fitting noisy inference model (blue line). Bottom: response switch curve predicted by the best-fitting noisy inference model (blue line). Noisy inference captures well the accuracy of behaviour surrounding reversals (top) but overestimates the variability of the same behaviour (bottom). Same conventions as in **e**. Shaded areas correspond to s.e.m. **g**, Discrepancies between models and human behaviour. Top: overall accuracy of participants (grey dot with error bar indicates the mean and s.d. of all participants, light grey dots indicate single participants), optimal inference (red bar) and noisy inference (blue bar). Bottom: overall switch rate of participants, optimal inference and noisy inference. Despite their suboptimal accuracy, participants make response switches as often as optimal inference. Error bars on simulated noisy inference model (blue bars) correspond to s.e.m.

interval (CI) (0.079, 0.123)). To characterize this suboptimality, we started from the optimal model but corrupted the inference with imprecise computations by adding normally distributed noise on

the result of each update step (Eq. 4). This suboptimal Bayesian inference model is controlled by two free parameters: (1) the hazard rate h , that is, the subjective rate of reversals of the bag from

which marbles are drawn, and (2) the inference noise σ , that is, the standard deviation (s.d.) of these computation errors in the statistical inference process. On each trial, the model updates imprecisely its belief regarding the drawn bag by combining its prior belief in that trial with the imperfect information provided by the new visual stimulus—that is, its likely bag based on its perceived lightness (Methods).

We fitted this noisy inference model to two characteristic features of human reversal behaviour simultaneously, each evaluated as a function of the time step from a change in hidden state (that is, a reversal) on four trials preceding and four trials following each reversal (Fig. 1f and Methods): (1) the response reversal curve, corresponding to the fraction of correctly reversed responses regarding the new hidden state after reversal, and (2) the response switch curve, corresponding to the fraction of participants' response switches regarding their own response on the previous trial. These two behavioural effects of interest are not independent, but each accounts for a separate behavioural dimension: the response reversal curve measures the accuracy of participants' responses regarding the bag being drawn after a reversal, whereas the response switch curve measures the stability of participants' responses across successive trials⁹.

Simulations of the best-fitting noisy inference model provided a good fit to participants' response reversal curves, but a poor fit to the same participants' response switch curves when both features were fitted simultaneously. Indeed, the model showed the same overall accuracy as participants (Fig. 1g; 0.765 ± 0.008 , difference: $t_{28} = 0.8$, $P = 0.416$, Cohen's $d = 0.153$, 95% CI $(-0.015, 0.006)$, Bayes Factor in favor of a null effect $BF_{01} = 3.707$), but larger overall switch rate (Fig. 1g; 0.201 ± 0.007 ; participants: 0.146 ± 0.007 ; difference: $t_{28} = 10.0$, $P < 0.001$, Cohen's $d = 1.858$, 95% CI $(0.044, 0.066)$). Despite their suboptimal accuracy, participants showed a similar switch rate as optimal inference (Fig. 1g; 0.141 ; difference: $t_{28} = 0.6$, $P = 0.553$, Cohen's $d = 0.111$, 95% CI $(-0.020, 0.011)$, $BF_{01} = 4.294$). This discrepancy between the humans' and the noisy inference model's overall switch rates suggests that participants deployed an additional strategy to compensate for the variability of beliefs due to imprecise inference.

Before assessing the nature of this 'response stabilization' strategy, we validated that participants' suboptimal performance is due to imprecise inference. First, the hazard rate h inferred from the participants' choices (0.081 ± 0.010) does not differ statistically from the true hazard rate of drawn bags (0.081 ; difference: $t_{28} = 0.0$, $P = 0.999$, Cohen's $d = 0.0002$, 95% CI $(-0.020, 0.020)$, $BF_{01} = 5.066$). This suggests that participants do not have a biased perception of the volatility of the task. Second, we ruled out noise in response selection as an alternative source of behavioural variability (Supplementary Fig. 1a). Inference noise provides a better fit to behaviour than selection noise (exceedance $P > 0.999$), a result which we validated through model recovery (Methods). Third, we verified that participants' suboptimal behaviour was not best explained by a leaky integration process (Methods, Eq. 3) rather than by the normative one (Methods, Eq. 2). This linear approximation⁷—even with additive inference noise—fails to better predict participants' reversal behaviour, and the suboptimal Bayesian model provides a better fit to both metrics as revealed by Bayesian model selection (Supplementary Fig. 2, exceedance $P > 0.999$).

Comparing belief stabilization strategies for imprecise inference.

We sought to characterize the cognitive strategy used by participants to make their responses more stable across successive trials. For this purpose, we derived five candidate strategies that implement such 'stabilization' at different processing stages in the noisy inference model described above (Fig. 2) and compared them by simulating their specific effects surrounding reversals (Fig. 3). The first candidate strategy we considered is a perceptual bias (Fig. 2a), in which

the sensory representation of each noisy stimulus is shifted in direction of prior beliefs following Bayes' rule, resulting in a biased perception of the bag to which each stimulus belongs^{10–12} (Methods, Eq. 5). This perceptual bias, operating before inference, leads to less variable beliefs at the expense of slower response reversals (Fig. 3).

The next two strategies we considered operate directly at the inference stage (Fig. 2b). Instead of updating beliefs based on every noisy stimulus, we considered that participants may occasionally not update their beliefs, effectively ignoring the information provided by the presented stimulus and sticking to their prior beliefs (and their previous response) until the next trial. We contrasted two possible strategies resulting in such inference 'omission': (1) a fraction of inference lapses during which participants are distracted from the task (Methods, Eq. 6), or (2) a conditional inference strategy where participants update their beliefs only when the strength of the noisy sensory signal associated with the presented stimulus exceeds a certain threshold (Methods, Eqs. 7 and 8). Although both strategies lead to less variable beliefs, making statistical inference contingent on the reliability of sensory representations does not delay response reversals by only discarding the stimuli that are more likely to be misperceived (Fig. 3, middle column).

The last two considered strategies operate at the response selection stage, following the inference stage (Fig. 2c): (1) a repetition bias which shifts participants' response criterion in favour of the previous response (Methods, Eq. 9), and (2) a fraction of response lapses associated with 'blind' response repetitions (disconnected from current beliefs, see Methods, Eq. 10). Like the perceptual bias and the inference lapses described above, these last two candidate strategies decrease the trial-to-trial variability of responses at the expense of slower adaptation to reversals (Fig. 3).

Before identifying which of these candidate strategies best explains human behaviour surrounding reversals (corresponding to the response reversal and switch curves described above), we assessed the ability to arbitrate between them through model recovery (Methods). We generated synthetic data from each model and verified the ability to correctly recover the 'ground-truth' model through Bayesian model selection (Fig. 4a and Methods; all exceedance $P > 0.95$). We then compared the five response stabilization strategies in terms of their ability to fit response reversal and response switch curves: (1) quantitatively through Bayesian model selection (Fig. 4b and Methods) and (2) qualitatively by examining their best-fitting response reversal and switch curves (Fig. 4c). Bayesian model selection revealed that the conditional inference strategy stands out by explaining simultaneously the accuracy (response reversal curve) and stability (response switch curve) of participants' behaviour surrounding reversals (Fig. 4b; exceedance $P > 0.999$). Qualitatively speaking, while all five candidate strategies reproduce participants' response reversal curves, only the conditional inference strategy explains the large and transient increase in response switches on the first trial following each reversal (Fig. 4c). Besides, we confirmed that the alternative suboptimal models (selection noise and leaky accumulation), even when stabilized with conditional inference, could not explain participants' reversal behaviour better than the noisy Bayesian inference model with conditional inference (Supplementary Figs. 1b and 2b). These findings suggest that participants decrease the trial-to-trial variability of their beliefs by updating them only when the incoming sensory information is deemed sufficiently reliable—thereby discarding unreliable information that could otherwise trigger unwarranted changes-of-mind. In practice, the best-fitting reliability threshold δ (1.030 ± 0.122 , mean \pm s.e.m.) suggests that participants discard as many as 32.5% of presented stimuli as unreliable.

Validating specific predictions of conditional inference. To provide further evidence in favour of conditional inference, we designed and ran a second experiment ($n = 30$ new participants)

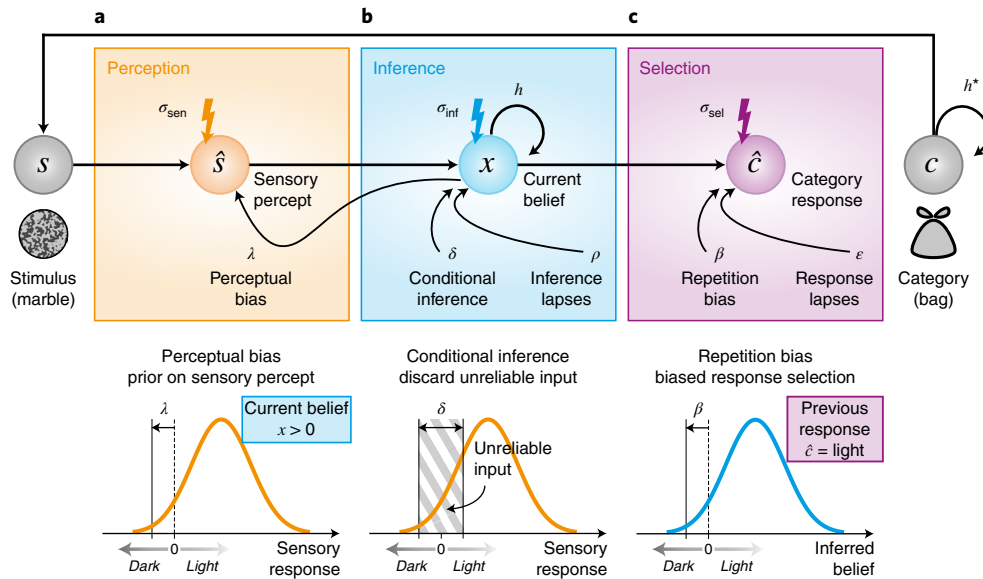


Fig. 2 | Candidate response stabilization strategies. A marble (stimulus s) is drawn from a given bag (category c , light or dark) at each trial. **a**, At the perception stage, the continuous sensory response to the stimulus, corrupted by sensory noise σ_{sen} is categorized into a binary sensory percept \hat{s} (light or dark). A perceptual bias, controlled by parameter λ , biases the perceptual categorization towards the current belief x at stimulus onset by shifting the categorization criterion. **b**, At the inference stage, the current belief x (expressed as the log-posterior odds ratio between the two bags) is updated as a function of the hazard rate h (prior term) and the incoming sensory percept \hat{s} (likelihood term). The inference process is corrupted by inference noise σ_{inf} . Belief stabilization is achieved either through conditional inference, by discarding stimuli whose associated sensory responses do not exceed a reliability threshold δ , or through a random fraction ρ of random inference lapses during which the current belief is not updated. **c**, At the selection stage, the current belief x is sampled with selection noise σ_{sel} to obtain a category response \hat{c} , corresponding to the bag perceived as being currently active (light or dark). Response stabilization is achieved either through a repetition bias towards the previous category response, controlled by parameter β , or through a random fraction ϵ of blind response repetitions (response lapses).

which allowed: (1) replicating the findings obtained in experiment 1 and (2) validating two specific predictions of conditional inference. First, we reasoned that if participants discard less reliable stimulus information, they should do so more often when the stimulus is difficult to categorize as light or dark—and hence triggers a weaker sensory signal. To test this first prediction, we had each bag contain marbles of three levels of difficulty (stimulus strength, corresponding to different proportions of light and dark areas) determined through the same titration procedure to achieve correct stimulus categorization of 70%, 80% and 90% (Fig. 5a). We predicted that participants should discard more difficult stimuli more often than easier stimuli.

Second, we hypothesized that the reliability threshold used by participants to discard sensory information should decrease with participants' confidence at categorizing isolated stimuli as light or dark. In other words, more confident participants should discard fewer stimuli as unreliable. For this purpose, we asked participants to provide confidence reports when categorizing stimuli during the titration procedure (Fig. 5b and Methods). These confidence reports showed classical signatures of decision evaluation (Fig. 5c): lower accuracy for decisions made with low confidence (repeated-measures ANOVA, $F_{1,29} = 84.9$, $P < 0.001$, effect size $\eta_p^2 = 0.745$), and a selective increase in confidence as a function of stimulus strength for correct decisions but not errors (correct: $F_{3,87} = 61.8$, $P < 0.001$, $\eta^2 = 0.681$; error: $F_{3,87} = 1.5$, $P = 0.241$, $\eta^2 = 0.048$).

As in experiment 1, participants showed suboptimal performance (Fig. 5d) and imprecise inference overestimated the trial-to-trial variability of participants' responses (Fig. 5e), which was again closer to the overall behavioural variability of optimal inference (Fig. 5f). In contrast to experiment 1, we could split response reversal and response switch curves as a function of stimulus strength on the

first trial following each reversal. Model simulations showed that conditional inference increases the degree of separation between these curves (Supplementary Fig. 3). This is because the growing fraction of discarded stimuli with categorization difficulty results in missed reversals when they are followed by a difficult stimulus. Model recovery confirmed that the conditional inference model is in principle identifiable from behaviour (Supplementary Fig. 4a; conditional inference M_3 : exceedance $P > 0.99$), and Bayesian model selection yielded strong evidence in favour of conditional inference (Supplementary Fig. 4b; exceedance $P > 0.999$). As in experiment 1, conditional inference is the only strategy that can explain the dynamics of participants' response switches surrounding reversals (Supplementary Figs. 3 and 4c).

To validate our first prediction that participants discard more difficult stimuli more often than easier ones, we fitted a separate reliability threshold for each stimulus strength. The reliability threshold is expressed in abstract units, and a more tangible metric of conditional inference corresponds to the associated 'discard rate': the overall fraction of discarded stimuli predicted by the model (Methods). The proposed conditional inference strategy predicts that the reliability threshold should not vary with stimulus strength (which varies from one trial to the next). We found that the reliability threshold remains indeed approximately constant (Fig. 6a; $F_{2,58} = 2.6$, $P = 0.095$, $\eta^2 = 0.081$), resulting in a decreasing discard rate with stimulus strength—from $33.3 \pm 3.8\%$ (mean \pm s.e.m.) for difficult stimuli categorized with 70% accuracy down to $13.6 \pm 1.2\%$ for easy stimuli categorized with 90% accuracy ($F_{2,58} = 14.6$, $P < 0.001$, $\eta^2 = 0.335$). This finding supports the notion that participants control inference by ignoring the sensory information below a fixed reliability threshold.

To validate our second prediction that the reliability threshold used to discard unreliable sensory information should decrease with

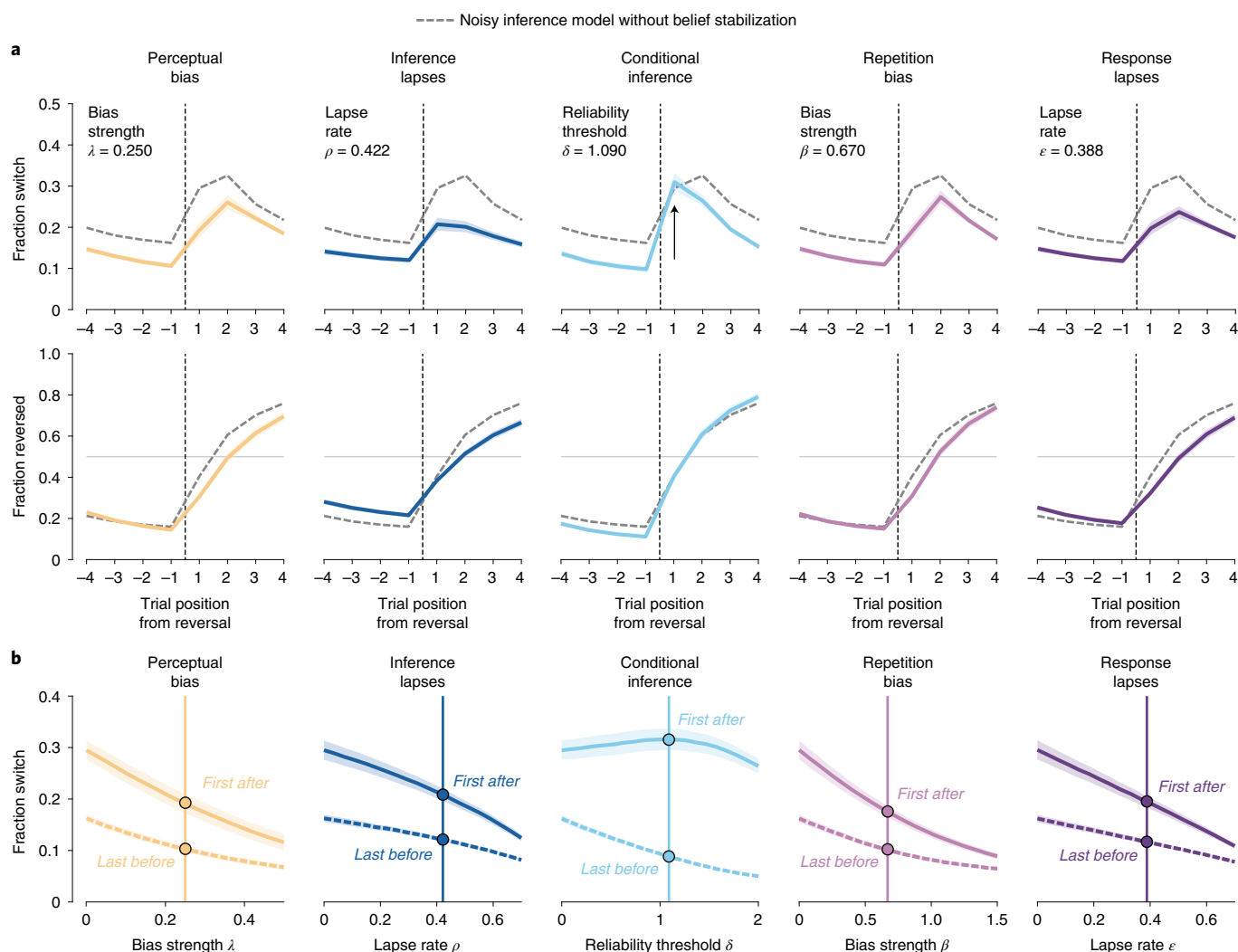


Fig. 3 | Predicted effects of response stabilization strategies. **a**, Simulated effects of the five candidate strategies on response switch curves (top row) and response reversal curves (bottom row). The parameters controlling each response stabilization strategy (λ , δ , ρ , β , ϵ) are set to match participants’ overall switch rate. The other parameters (h , σ_{sens} , σ_{intr} , σ_{sel}) are fixed to their best-fitting values for the noisy inference model without response stabilization. Solid coloured lines correspond to the reversal behaviour of each candidate model, whereas dotted grey lines correspond to the reversal behaviour of the noisy inference model without response stabilization (same for all panels). The conditional inference strategy stands out from other strategies on the first trial after reversal (arrow). **b**, Simulated effects of the five candidate strategies on response switches just before and after a reversal. Fraction of response switches on the last trial before each reversal (dotted lines) and the first trial after each reversal (solid lines) for each response stabilization strategy. Dots correspond to stabilization parameters set as in **a**. All candidate strategies reduce simultaneously response switches before and after reversals, except for conditional inference which reduces response switches only before reversals (that is, when they are not warranted). Shaded areas around curves in **a** and **b** correspond to s.e.m.

confidence across participants, we fitted the confidence threshold δ_C of each participant during the titration procedure (Supplementary Fig. 5 and Methods). We then regressed this confidence parameter—which decreases with confidence—against the parameters of the conditional inference model fitted to the same participants during the main task. As predicted, we found a positive relation between the confidence threshold and the reliability threshold across participants (Fig. 6b; linear correlation coefficient $r=0.392$, $P=0.033$, 95% CI (0.037, 0.659)). This positive relation remained significant when accounting for the apparent internal variability in confidence reports (Supplementary Fig. 5c; regression coefficient $b=0.250 \pm 0.089$, $P=0.009$, 95% CI (0.075, 0.425)). This second finding suggests that the reliability threshold used to discard unreliable sensory information depends on the confidence with which participants categorize individual stimuli.

Identifying individual differences in conditional inference.

To provide further support for conditional inference, we studied whether the reliability threshold used to discard stimuli shows specific individual differences across participants. We reasoned that, if individual differences in reliability threshold δ reflect genuine differences in conditional inference, then these differences should follow a behavioural ‘gradient’ that is distinct from the ones associated with the other two parameters in the model (the hazard rate h and the inference noise σ). More specifically, individual differences in δ should mainly affect the variability of participants’ responses surrounding reversals—that is, their response switch curves.

First, we observed that all three parameters show large individual differences (Fig. 7a). Importantly, these parameters are only moderately correlated with each other (Fig. 7b)—with at most 10% of shared pairwise variance—and each of them shows significant test–retest

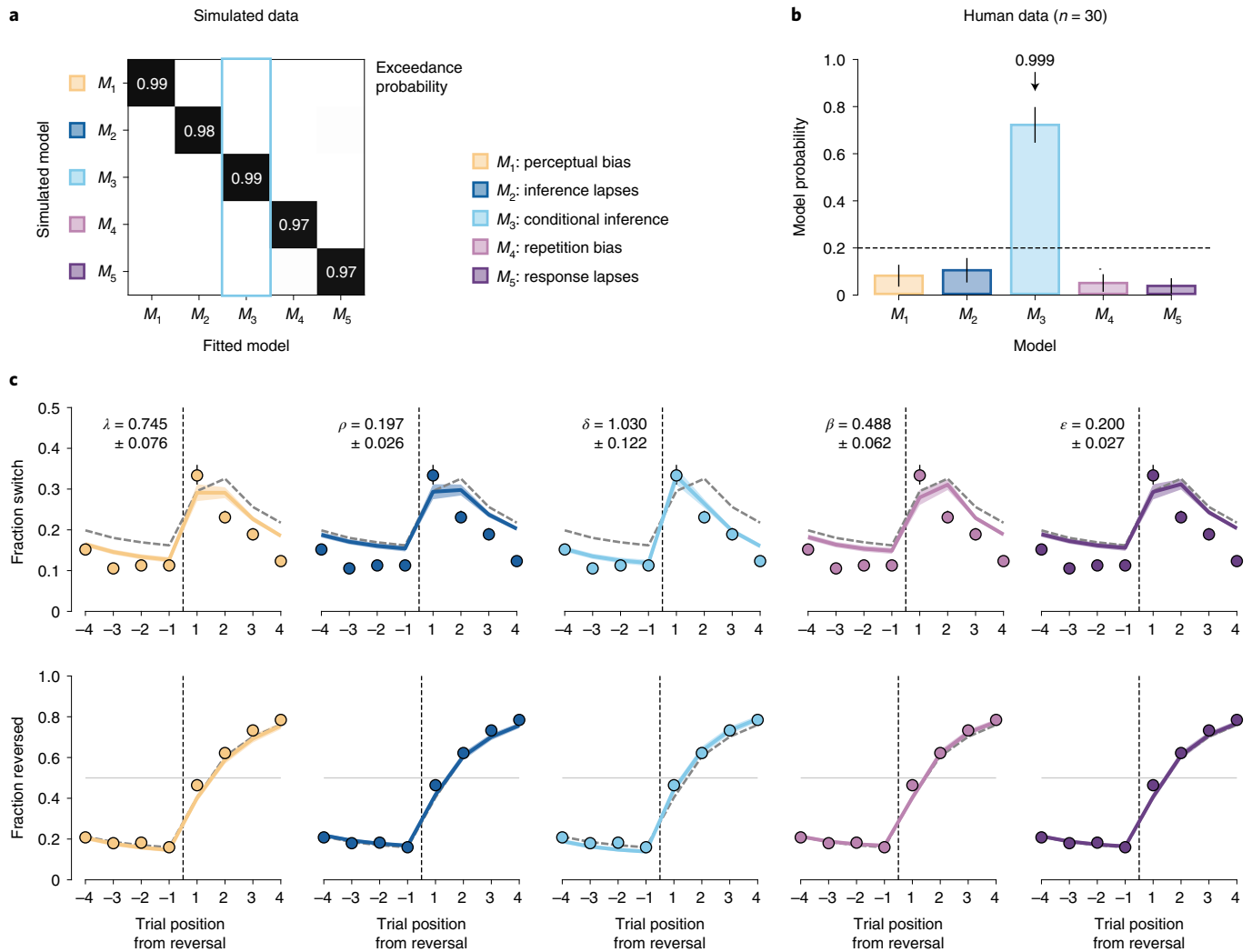


Fig. 4 | Belief stabilization through conditional inference (n = 30). **a**, Confusion matrix between response stabilization strategies depicting exceedance probabilities p_{exc} obtained from ex ante model recovery. The ‘ground-truth’ response stabilization strategy used to simulate synthetic behaviour was correctly recovered with $p_{\text{exc}} > 0.97$ for all five strategies, including conditional inference (the model that best describes participants’ behaviour) with $p_{\text{exc}} > 0.99$. **b**, Random-effects Bayesian model selection of the best-fitting strategy in participants’ data. Bars indicate the estimated model probabilities for the five candidate strategies. The conditional inference model (M_3) best describes the behaviour of more participants than other candidate strategies with $p_{\text{exc}} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution. The dashed line corresponds to the uniform distribution. **c**, Simulations of response stabilization strategies fitted to participants’ data. Simulated response switch curves (top row) and response reversal curves (bottom row) based on the best-fitting parameters of each model. The mean best-fitting parameter (mean \pm s.e.m.) is shown in the top-left corner for each subpanel. Solid coloured lines correspond to the reversal behaviour of each stabilized model, whereas dotted grey lines correspond to the reversal behaviour of the noisy inference model without response stabilization (same for all panels). Dots indicate human data (group-level average). Only the conditional inference model reproduces participants’ reversal behaviour. Shaded areas and error bars correspond to s.e.m.

reliability across the two experimental sessions performed by each participant on separate days (hazard rate $h:r=0.698$, $P < 0.001$, 95% CI (0.538, 0.809); inference noise $\sigma:r=0.308$, $P = 0.018$, 95% CI (0.056, 0.523); reliability threshold $\delta:r=0.574$, $P < 0.001$, 95% CI (0.373, 0.724); Supplementary Fig. 6 and Methods). To test whether each parameter is associated with a specific behavioural gradient, we split participants into two groups as a function of the best-fitting value of the parameter being considered and plotted the response reversal and switch curves for each group (Fig. 7c and Methods). We found that the behavioural gradient associated with the hazard rate h explains 48.7% of individual differences in response reversal curves, with a clear effect on the reversal time constant. The behavioural gradient associated with inference noise σ explains 22.5% of individual differences in response reversal curves, with a distinct effect on the overall accuracy of responses. By contrast, the

behavioural gradient associated with the reliability threshold δ explains only 4.2% of individual differences in response reversal curves, but 24.8% of individual differences in response switch curves. In line with the conditional inference model, participants with higher reliability thresholds show less variable responses but equally rapid adaptation to reversals (Fig. 7c).

To validate the existence of this behavioural gradient in a ‘bottom-up’ fashion, independently of fits of the conditional inference model, we performed a principal component analysis (PCA) of participants’ response reversal and switch curves (Supplementary Fig. 7 and Methods). Strikingly, the three first principal components (PC) obtained using this variance partitioning procedure map on the three parameters of the conditional inference model. The third PC (PC3) is associated with individual differences in response switch curves (16.4%), but little to no variability in response reversal

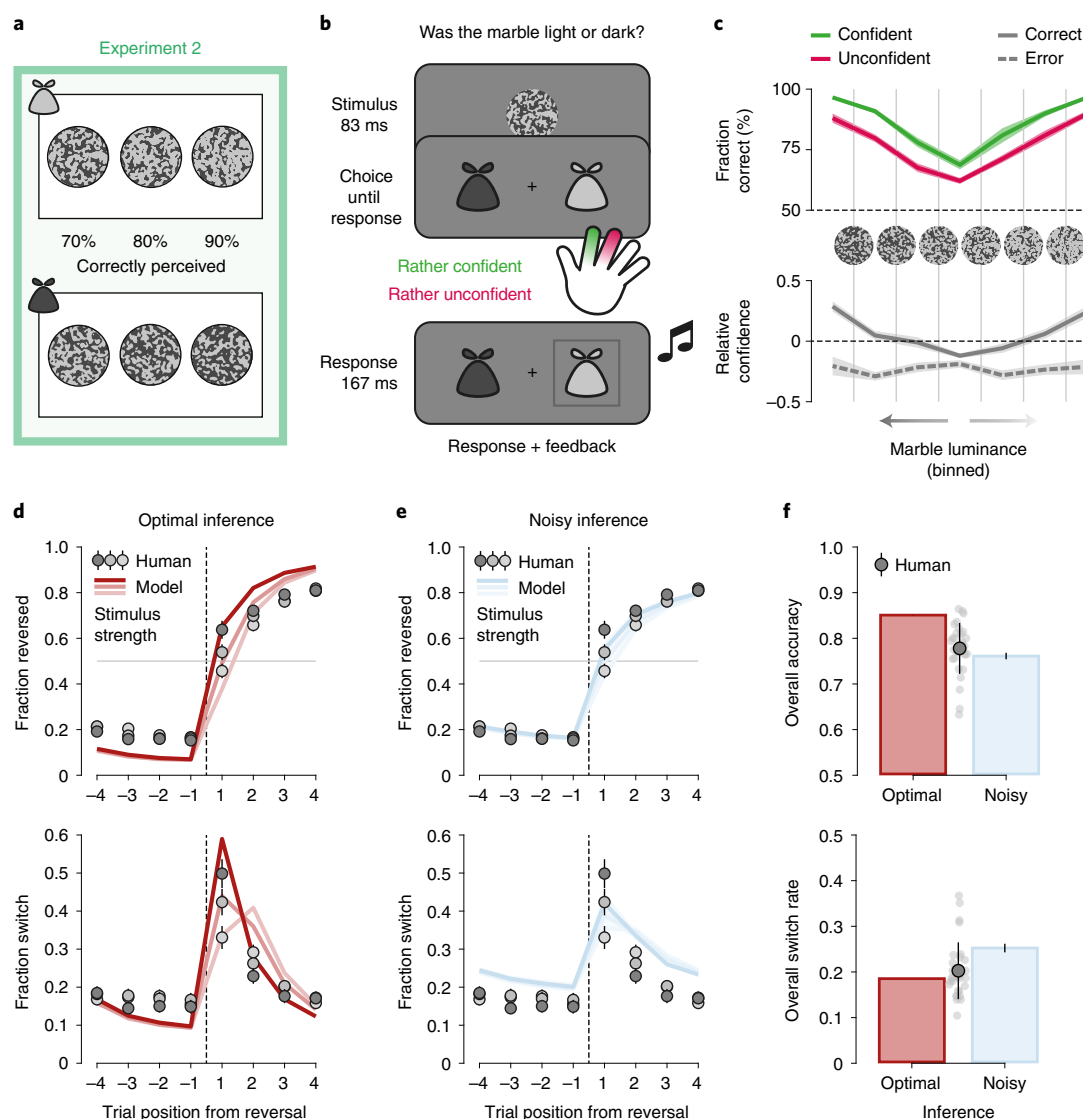


Fig. 5 | Description of experiment 2 ($n = 30$). **a**, Visual stimuli. Each bag is filled with three types of marbles, corresponding respectively to 70%, 80% and 90% of correctly categorized stimuli, using the same adaptive titration procedure as in experiment 1. **b**, Titration trials with confidence reports. Participants were asked to categorize the presented marble as light or dark, and report simultaneously their confidence level as rather high or rather low. Participants received auditory feedback after each titration trial. **c**, Relation between decision confidence and accuracy in titration trials. For each binned luminance, participants' fraction of correct decisions for confident (green) and unconfident (red) trials (top panel), and participants' fraction of confident decisions for correct (solid line) and error (thick dashed line) trials (bottom panel). Confidence is expressed relative to each participant's mean confidence (thin dashed line). Shaded areas correspond to s.e.m. **d**, Predictions of optimal inference as a function of stimulus strength on the first trial following each reversal. Top: response reversal curves indicating the fraction of responses towards the new drawn bag surrounding each reversal. Bottom: response switch curves indicating the fraction of trial-to-trial response switches surrounding the same reversals. The reversal is represented by the thin dotted line. Dots indicate human data (group-level average), whereas lines indicate predictions of optimal inference. Error bars correspond to s.e.m. **e**, Predictions of the noisy inference model as a function of stimulus strength on the first trial following each reversal. Noisy inference captures well the accuracy of behaviour surrounding reversals (top), but overestimates the variability of behaviour (bottom). Same conventions as in **d**. Shaded areas correspond to s.e.m. **f**, Discrepancies between models and human behaviour. Top: overall accuracy of participants (grey dot with error bar indicate the mean and s.d. of all participants, light grey dots indicate single participants), optimal inference (red bar) and noisy inference (blue bar). Bottom: overall switch rate of participants, optimal inference and noisy inference. Despite their suboptimal accuracy, participants make response switches as often as optimal inference. Error bars on simulated noisy inference model (blue bars) correspond to s.e.m.

curves (3.4%). Accordingly, participant-specific scores on this PC correlate selectively with reliability thresholds fitted with the conditional inference model ($r^2 = 0.348$, exceedance $P = 0.861$, 95% CI (0.168, 0.551)). Together, these top-down (model-based) and bottom-up (PCA-based) findings provide additional evidence for conditional inference by identifying a behavioural gradient characteristic of conditional belief updating across tested participants.

Measuring cognitive benefits of conditional inference. The reduced trial-to-trial variability of beliefs triggered by conditional inference mitigates the negative effects of external (unreliable sensory information) and internal (imprecise computations) sources of variability during statistical inference, by discarding the incoming stimulus information that does not exceed a minimum reliability level. We thus tested whether conditional inference may not only

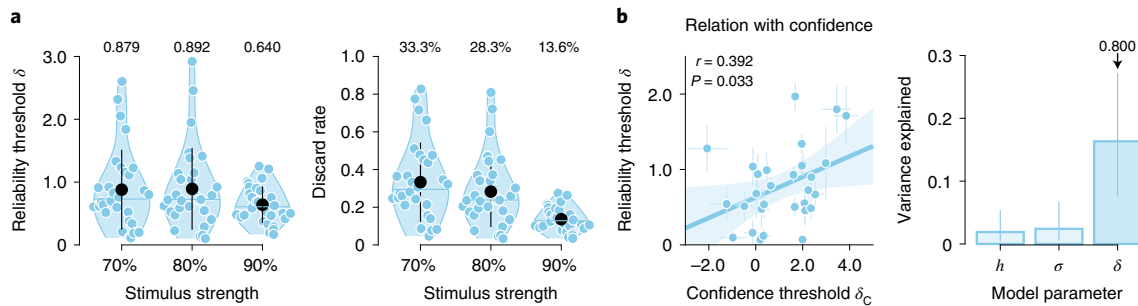


Fig. 6 | Validation of specific predictions of conditional inference ($n = 30$). **a**, Decrease in evidence discard rate with stimulus strength. The reliability threshold does not vary with stimulus strength, resulting in discard rates that decrease with stimulus strength (one blue dot per participant). Black dots represent group-level means and error bars represent their associated standard deviations. Horizontal blue lines represent group-level medians. **b**, Selective relation between reliability and confidence criteria. Left: correlation between confidence threshold and reliability threshold across participants. As predicted by conditional inference, the reliability threshold estimated in the volatile decision-making task correlates positively with the confidence threshold estimated when categorizing isolated stimuli. Parameters are presented as mean \pm s.d. of posterior distributions of each fit. Shaded area corresponds to the 95% confidence interval for the regression line. Right: variance of confidence threshold explained by each model parameter. The reliability threshold shares more variance with the confidence threshold than the other two model parameters. Bars correspond to r^2 and error bars to the interquartile ranges of each r^2 measure obtained through bootstrapping ($n = 10^4$).

decrease the variability of beliefs and resulting responses, but also, counterintuitively, increases accuracy. For this purpose, we compared the effects of the five tested belief stabilization strategies on accuracy across both experiments through simulations (Fig. 8a). We found that all strategies decrease behavioural variability at the expense of lower accuracy, except for conditional inference which shows an inverted U-shaped relation with accuracy. The accuracy-optimizing value δ^* of the reliability threshold ($\delta^* = 1.281$, associated with 5.5% more accurate responses) is associated with as much as 41.3% of discarded stimuli categorized with an overall 80% accuracy (Supplementary Fig. 8). This performance benefit of conditional inference is due to the combination of less variable beliefs (and responses) at the end of ‘episodes’—that is, successive draws from the same bag—and equally fast adaptation to reversals (Supplementary Fig. 8). By contrast, other belief stabilization strategies trade less variable responses against slower adaptation to reversals. A ‘soft’ version of the conditional inference model (Methods) that downweights the evidence that falls below the reliability threshold rather than discarding it also predicts an inverted U-shaped relation with accuracy. However, it fails at predicting a decreased behavioural variability and explains human behaviour less accurately than the conditional inference model (exceedance $P < 0.001$; Supplementary Figs. 9 and 10).

We measured this performance benefit of conditional inference in participants’ data by fitting a multiple regression model of accuracy as a function of each parameter: the hazard rate h (entered as a quadratic effect as it is expected to show an inverted U-shaped relation with accuracy, maximal accuracy expected around the true value of the hazard rate, see Methods), the inference noise σ_{inf} (entered as a linear effect as it is expected to show a negative relation with accuracy) and the reliability threshold δ (entered as a quadratic effect as it is also expected to show an inverted U-shaped relation with accuracy; Fig. 8a). This multiple regression model provides an accurate fit to participants’ data (Fig. 8b; $r^2 = 0.859$, d.f. = 53, $P < 0.001$), and confirms the inverted U-shaped relation between reliability threshold and accuracy (quadratic coefficient $b_5 = -0.028 \pm 0.009$, mean \pm s.e.m., $t_{53} = -3.2$, $P = 0.002$, 95% CI $(-0.046, -0.011)$). The accuracy-optimizing value of the reliability threshold estimated from the best-fitting model ($\delta^* = 1.042 \pm 0.174$, mean \pm s.e.m.) is consistent with simulations (Fig. 8a). We confirmed that the quadratic relation between reliability threshold observed in participants’ data is predicted by simulations of the best-fitting conditional inference model, by replacing participants’ accuracy with the

accuracy of model simulations (Fig. 8c; $b_5 = -0.025 \pm 0.006$, $t_{53} = -3.9$, $P < 0.001$, 95% CI $(-0.0383, -0.0123)$). Implementing conditional inference with the optimal reliability threshold δ^* is associated with comparable performance improvements in participants and simulations (participants: +3.2%; simulations: +5.5%). These findings confirm that conditional inference not only mitigates the variability of beliefs arising from imprecise inference, but also improves the accuracy of these beliefs by trading the costs of inference against its expected benefits based on the reliability of incoming sensory information.

Discussion

Human observers combine multiple pieces of sensory evidence to form or revise beliefs about the current uncertain state of their environment. These beliefs remain nonetheless imprecise, not only because of the expected uncertainty associated with each piece of evidence or the unexpected uncertainty associated with transitions between states¹³, but also due to the imprecision of the inference process itself^{1–6}. Variability in beliefs might therefore help adapt to new contingencies, but might also be mistaken for genuine changes in the environment and trigger unwarranted changes-of-mind. Here, we investigated whether and how human observers mitigate the trial-to-trial variability of their imprecise beliefs, by studying human behaviour in a volatile decision-making task. Using computational modelling, we found that humans stabilize their imprecise beliefs using an efficient ‘conditional’ strategy—whereby they discard pieces of sensory evidence that are not deemed as sufficiently reliable. We show that this conditional belief updating strategy is associated with large individual differences linked to participants’ perceptual confidence, and improves the accuracy of resulting decisions.

It is well-known that human decisions show a suboptimal trial-to-trial variability under uncertainty^{5,6}. This behavioural variability can arise from internal ‘noise’ located at different stages of information processing. Classical cognitive models consider behavioural variability as a consequence of noise corrupting either sensory processing or the decision policy, in this later case by introducing a softmax at the action selection stage^{7,14,15}. However, recent findings have shown that the large fraction of the observed behavioural variability arises from computation noise in inference^{4,6,16}. Both policy noise and inference noise generate behavioural variability, but these two sources of noise differ regarding the trial-to-trial variability of underlying beliefs. Indeed, policy noise does not generate any trial-to-trial variability in beliefs but only variability in behaviour.

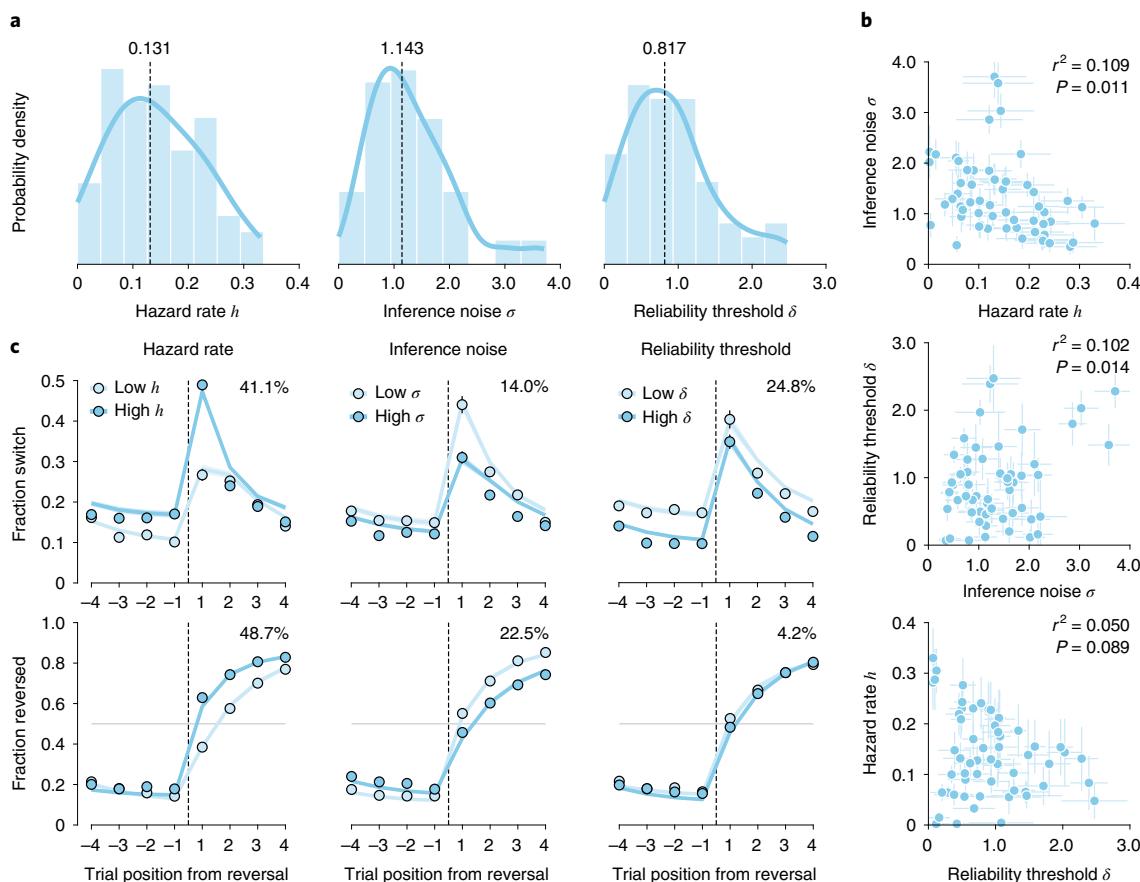


Fig. 7 | Interindividual variability in conditional inference ($n = 60$). **a**, Distribution of conditional inference model parameter values across participants from both experiments. Best-fitting parameters and their median value (dotted lines) with density approximation (solid blue lines). **b**, Pairwise relations between the parameters of the conditional inference model. Parameters are only mildly correlated with one another. Parameters are presented as mean \pm s.d. of posterior distributions of each fit. **c**, Distinct gradients of behavioural variability associated with each parameter of the conditional inference model (left: hazard rate; middle: inference noise; right: reliability threshold). Response switch curves (top) and response reversal curves (bottom) obtained by sorting participants in two median-split groups as a function of the best-fitting value of each parameter. Lines correspond to simulations of the best-fitting conditional inference model, whereas dots correspond to participants' behaviour. Percentages indicate the variance in response switch curves (top row) and response reversal curves (bottom row) explained by each parameter. Shaded areas and error bars correspond to s.e.m.

By contrast, inference noise generates trial-to-trial variability in both beliefs and behaviour. Recent neurophysiological findings support the idea that the large moment-to-moment variability of neural activity in cortical areas which reflects evidence accumulation corresponds to genuine variability in beliefs¹⁷.

Large moment-to-moment fluctuations in beliefs are likely to be costly, both psychologically and neurally. At the psychological level, changes-of-mind require changes in behaviour, which entail increased cognitive control and slower response latencies. At the neural level, previous studies have shown that changes-of-mind are associated with strong increases in neural activity in the prefrontal cortex in humans¹⁸ and rodents¹⁹. We thus reasoned that humans may rely on a form of 'belief stabilization' strategy to mitigate the impact of inference noise on belief variability and prevent unwarranted changes-of-mind. We found that human behaviour is less variable than what is predicted from the amount of inference noise that matches behavioural accuracy. To characterize how humans reduce the variability of their behaviour, we compared several stabilization strategies at different processing stages. Some of them mitigate the variability of behaviour but not beliefs, whereas others stabilize the underlying beliefs^{11,12}. We found that human observers use a 'conditional' strategy to stabilize their beliefs. This means that humans discard (ignore) a fraction of the available sensory

evidence judged as too weak and unreliable. Interestingly, this strategy challenges the idea that more information is always better in the presence of inference noise when performing a volatile decision-making task.

Previously described strategies typically reduce the variability of beliefs through some form of 'confirmation bias', either by shifting the perception of sensory evidence in the direction of prior beliefs as in the 'perceptual bias' model that we have considered among alternative stabilization strategies, or by underweighting (or even discarding) the sensory evidence that is inconsistent with prior beliefs. In particular, recent work^{20,21} has identified such a 'stimulus-consistency' bias mechanism during sequential cue combination. This bias, which resembles a similar bias described during reinforcement learning²², was observed in stable environments. In these conditions where the latent cause does not change, it was shown to increase accuracy by reducing decision sensitivity to misleading sensory observations. By contrast, in volatile environments, the same bias typically decreases accuracy because inconsistent observations can signal a change in their latent cause, and should therefore not be underweighted²³. Future work should further investigate the effect of task parameters on the stabilization strategies deployed by humans to mitigate the negative effects of inference noise.

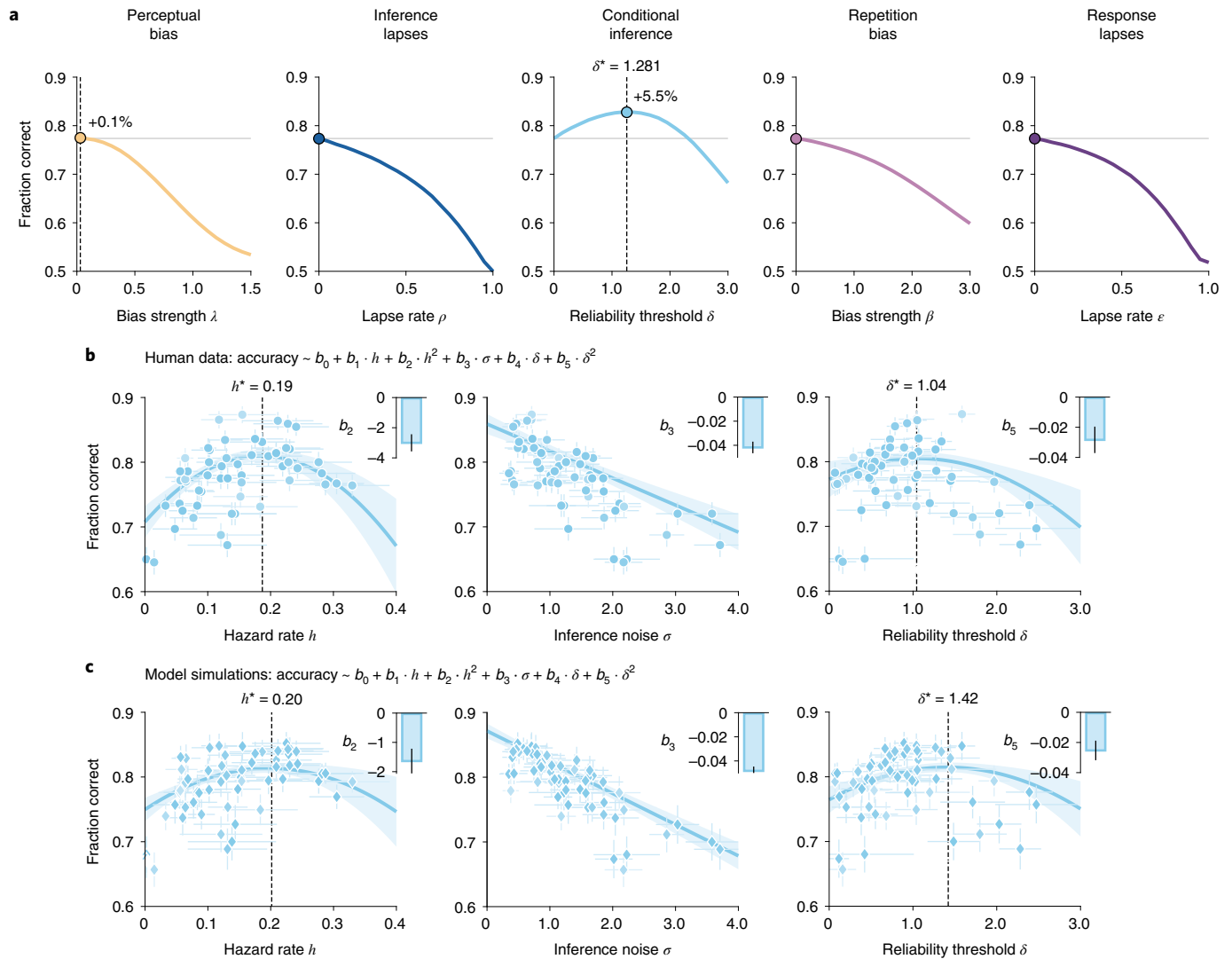


Fig. 8 | Increased decision accuracy through conditional inference. a, Simulated effects of response stabilization strategies on decision accuracy. Fraction of overall correct responses when increasing each stabilization parameter drops for all models except for the conditional inference model, for which it increases the accuracy by +5.5% for $\delta^* = 1.28$. Dots correspond to the best simulated accuracy. **b**, Observed effects of best-fitting model parameters on human decision accuracy. For each best-fitting conditional inference model parameter, corresponding human accuracy (dots) and robust regression (solid blue lines, $r^2 = 0.859$, $P < 0.001$). Inverted U-shaped relation between reliability threshold and accuracy (negative quadratic coefficient b_5 , $P = 0.002$). Shaded area corresponds to the 95% confidence interval for predicted values. Parameters are presented as mean \pm s.d. of posterior distributions and human accuracy as mean \pm s.d. given the number of trials provided. Top right insets correspond to estimated regression coefficients presented as mean \pm s.e.m. **c**, Predicted effects of best-fitting model parameters on modelled decision accuracy. For each best-fitting conditional inference model parameter, corresponding conditional inference model predicted accuracy (diamonds) and robust regression (solid blue lines). Inverted U-shaped relation between reliability threshold and accuracy (negative quadratic coefficient b_5 , $P < 0.001$). Shaded areas correspond to the 95% confidence interval for predicted values. Parameters are presented as mean \pm s.d. of posterior distributions and vertical error bars correspond to the accuracy s.d. of best-fitting model simulations. Top right insets correspond to estimated regression coefficients presented as mean \pm s.e.m.

In our task, tested participants discard about a third of presented stimuli, a surprisingly large fraction for stimuli categorized in isolation with 80% overall accuracy in a forced-choice task. Importantly, the ‘lapse’ rate measured in the condition where participants categorize isolated stimuli is widely smaller than the ‘conditional discard’ rate measured in the main task (lapse rate: 9.3%; conditional discard rate: 28.9%). This means that the conditional inference strategy observed in our task does not reflect lapses in attention. Consistently with the idea that apparent lapses in behaviour not only reflect inattention, recent work in mice has shown that lapses measured during perceptual decisions can reflect exploration²⁴ or alternations between discrete behavioural strategies²⁵. In our task,

we propose that the discarding of unreliable stimuli is controlled by a general metacognitive process linked to perceptual confidence. Indeed, the reliability threshold used for discarding stimuli correlates with the confidence threshold measured in the same participants during an independent task.

In addition, tethering inference to the reliability of incoming sensory evidence shows positive effects as it reduces the cognitive costs of imprecise inference, in terms of unwarranted changes-of-mind as already mentioned, but also in terms of mental effort by waiving inference on a substantial fraction of trials. The fact that mental effort can be aversive enough for people to accept non-negligible levels of physical pain²⁶ can thus partially explain why humans rely

on an efficient metacognitive strategy to reduce mental effort. We further show that conditional inference simultaneously increases the accuracy of decisions while decreasing mental effort, making it a win–win strategy in our standard reversal learning task.

Like most decision parameters²⁷, conditional inference shows large individual differences across tested participants. We observed that the reliability threshold used to discard unreliable stimuli is associated with a specific behavioural gradient across participants. The fact that the position of individual participants on this gradient remains relatively stable across both experimental sessions indicates a genuine trait-like parameter. One important question concerns the relation between this trait-like parameter and other cognitive traits. We have shown that the reliability threshold used to discard stimuli correlates with the confidence threshold adopted by the same participants in a different task—another trait-like parameter which shows high test–retest reliability across experimental sessions. This means that the reliability threshold could be related to transdiagnostic psychiatric symptom dimensions²⁸ known to correlate with confidence threshold in various tasks²⁹.

In our task, we introduced expected uncertainty by degrading stimuli using sensory noise, but this is not the only way by which expected uncertainty can be introduced in decision tasks¹². Indeed, existing tasks often introduce uncertainty by reducing the information conveyed by a single stimulus about its generative category—a form of uncertainty referred to as category ambiguity. The reliability-contingent control of inference observed in our task has not been identified in previous tasks where, unlike the present task, the sensory reliability of presented stimuli is large and it is category ambiguity that is low⁴. Future work could investigate more explicitly the different strategies deployed for mitigating the impact of imprecise beliefs in conditions of high sensory noise versus high category ambiguity. Another important question concerns the neural processes that may underlie this conditional inference strategy. Our model is purely conceptual and suggests a form of hierarchical inference, in line with other proposed belief stabilization strategies^{11,12}. In our model, this means that beliefs are updated downstream (and separately from) the categorization and evaluation of individual stimuli. Future work will be needed to investigate the possible neural mechanism underlying conditional inference, which may be a form of ‘gating’—an all-or-none, non-linear filtering mechanism—that requires sufficiently large sensory responses to trigger changes in the downstream decision (belief updating) circuit.

Taken together, our results show that humans condition the imprecise update of their beliefs to the reliability of incoming sensory information—an efficient metacognitive control process which not only prevents unwarranted changes-of-mind, but also reduces mental effort by waiving belief updating. Avoiding the mental effort spent on cognitive computations that are less likely to pay off (in this case, based on unreliable evidence) proves a surprisingly effective strategy that simultaneously increases the objective accuracy of resulting decisions—thereby bypassing the general cognitive trade-offs between the benefits and costs of cognitive operations described in the literature³⁰.

Methods

Participants. Sixty healthy adult participants (30 females, age: 25.8 ± 3.9 years, all right-handed, experiment 1: $n = 30$, 15 females, age: 24.7 ± 3.9 years; experiment 2: $n = 30$, 15 females, age: 27.16 ± 3.7 years) took part in the study. One participant was excluded from analyses due to poor performance (overall decision accuracy more than 3.5 s.d. below the group-level mean). Participants were screened for the absence of any history of neurological and psychiatric disease or any current psychiatric medication. All participants provided written informed consent, and received 45 euros as compensation for their participation after completing the second experimental session. The study followed guidelines from the Declaration of Helsinki and specific procedures which received ethical approval from relevant authorities (Comité de Protection des Personnes Ile-de-France VI, ID RCB: 2007-A01125-48, 2017-A01778-45).

Task and procedure. Both experiments were based on the same reversal learning task. In repeated trials, participants were presented with visual cues described as ‘marbles’, which could be drawn either from a bag containing light marbles or a bag containing dark marbles. Participants had to infer the bag the marble was drawn from (light or dark). Participants were instructed that marbles were not drawn randomly and independently across trials, but from the same bag for a certain number of successive trials (that is, an episode) before being drawn from the other bag (that is, a generative category reversal). Participants were herewith instructed about the presence of reversal but did not know the number of trials (or episode length) before a reversal. The length of these episodes was drawn from a bounded exponential distribution, resulting in an approximately flat hazard rate—a procedure equivalent to a Markov process with fixed transition probability. Each marble corresponded to a two-tone disc with light and dark shades of grey, in different proportions and spatially scrambled over the surface of the disc. The disc was generated using Gaussian noise filtered through a two-dimensional Gaussian smoothing filter and eventually binarized accordingly to the expected proportions between the light and dark shades of grey. All stimuli were presented on a uniform medium grey background. In experiment 1, the relative proportions of light and dark shades of grey were adjusted using a titration procedure to reach 80% of marbles perceived as light for the light bag, and 80% of marbles perceived as dark for the dark bag. In experiment 2, three levels of difficulty were titrated for each bag to achieve correct stimulus identification of 70%, 80% and 90%, respectively. In this second experiment, participants were not explicitly instructed of the three difficulty levels.

The titration procedure preceded each reversal learning block, and corresponded to two adaptive psychophysical staircases³¹ adjusting the relative proportions of light and dark shades of grey in presented marbles. We ran two separate staircases because no symmetry was assumed: participants could have a biased perception towards the light or the dark category. The trials of the two staircases were randomly interleaved during the procedure. In experiment 1, each staircase converged towards a light/dark proportion, one proportion was used to generate the marbles populating the light and the other for marbles populating the dark bag, each corresponding to 80% accuracy (light bag: $52.3 \pm 1.1\%$; dark bag: $47.6 \pm 1.0\%$). In experiment 2, an additional logistic regression was performed to extrapolate the light/dark proportions corresponding to 70% and 90% accuracy levels from the light/dark proportions corresponding to 80% accuracy (light bag: $51.6 \pm 1.4\%$ for 70% accuracy, $52.6 \pm 1.5\%$ for 80% accuracy, $54.1 \pm 1.7\%$ for 90% accuracy; dark bag: $48.40 \pm 1.3\%$ for 70% accuracy, $47.4 \pm 1.4\%$ for 80% accuracy, $45.9 \pm 1.5\%$ for 90% accuracy). The titration blocks were presented to participants as a distinct game (‘game 1’): participants were instructed to sort individual marbles, and informed that marbles were presented in random order (unlike the main task). Auditory feedback was provided at the end of each trial in titration blocks.

The reversal learning blocks were presented as ‘game 2’ with instructions to report whether the computer is drawing marbles from the light or the dark bag after the presentation of each marble. Unlike the titration blocks, no feedback was provided during the reversal learning blocks. The experiment consisted of 16 blocks (eight blocks from the task condition described above, and eight blocks of another task condition not relevant for the current study), divided into two experimental sessions of eight blocks, each session lasting approximately 90 min and taking place on different days. Each reversal learning block consisted of 80 trials. Experiment 1 consisted of two types of blocks: ‘low volatility’ blocks containing five reversals (that is, hazard rate $h = 1/16$) and ‘high volatility’ blocks containing 10 reversals (that is, hazard rate $h = 1/8$). Experiment 2 consisted of ‘high volatility’ blocks only. The first generative category of each block was counterbalanced pseudo-randomly across blocks and participants.

Participants were instructed that each marble consisted of both light and dark shades of grey spatially scrambled. They were also instructed that all marbles contained in the light bag were predominantly light, whereas all marbles contained in the dark bag were predominantly dark (Supplementary Fig. 12). Before the experiment, they performed a short training for both ‘games’ (titration and reversal learning blocks) to get familiarized with the stimuli and the difficulty of identifying this predominant dark/light colour due to the spatial scrambling and the short stimulus presentation time (83 ms).

The experiment was coded in MATLAB and run using Psychtoolbox-3 (ref. ³²). Participants performed the experiment in a soundproof booth with their head positioned on a chin rest at 75 cm from a 24-inch LCD screen operating at 60 Hz with a resolution of 1920×1080 pixels.

Bayes-optimal inference. For each trial t , the decision-maker tries to find the associated hidden state corresponding to one of the two alternative categories c , $c_t = +1/-1$ (light or dark bag). At each trial, the belief about the current value of the hidden state is updated by combining the prior belief with the new incoming evidence according to Bayes rule. The normalized model of decisions between two alternatives is formalized in ref. ⁷, with the belief at trial t , L_t , defined as the logarithm of the posterior odds of the alternative categories accumulated until this trial. The sign of the log-odds belief indicates which category is more likely, whereas the magnitude of the log-odds belief indicates the strength of the belief in favour of the more likely hidden state. The update rule combining prior belief L_{t-1} and new incoming evidence LLR_t is defined as follows:

$$L_t = \psi(L_{t-1}, h) + \text{LLR}_t, \quad (1)$$

with ψ the time-varying prior expectation defined by

$$\psi(L_{t-1}, h) = L_{t-1} + \log \left[\frac{1-h}{h} + \exp(-L_{t-1}) \right] - \log \left[\frac{1-h}{h} + \exp(L_{t-1}) \right], \quad (2)$$

and h the hazard rate, the expected probability at each trial t that the category will switch, which corresponds to a non-linear integration leak. We also considered a linear approximation of the prior expectation function⁷ (Eq. 2) turning the model into a leaky evidence accumulator (Supplementary Fig. 2):

$$\psi(L_{t-1}, h) \approx (1 - 2h) \times L_{t-1}. \quad (3)$$

Suboptimal Bayesian inference. Each update of the log-odds belief with new stochastic incoming evidence LLR_t is corrupted by internal Gaussian noise— inference noise—with s.d. σ_{inf} turning the deterministic belief update equation (Eq. 1) into a stochastic draw from a normally distributed random variable:^{4,33}

$$L_t \sim N(\psi(L_{t-1}, h) + \text{LLR}_t, \sigma_{\text{inf}}^2). \quad (4)$$

Model-based analysis of behaviour and stabilization strategies. We modelled human behaviour in the experimental condition using a hierarchical suboptimal Bayesian model with two free parameters h and σ_{inf} and with an additional model-specific stabilization parameter. We consider the sensory response to the stimulus presented at trial t as a normally distributed random variable generated by $\mathcal{N}(\pm 1, \sigma_{\text{sensory}}^2)$, depending on the true underlying category of the presented stimulus (+1 for light, -1 for dark). The s.d. of sensory noise σ_{sensory} is set to correspond to 80% of correct categorization—achieved through the online titration procedure.

In the first (‘categorized’) version of the model, this sensory response is then categorized as light ($c_t = +1$) or dark ($c_t = -1$) and expressed as a log-odds ratio. By design of the task, with $p(\hat{s}_t > 0 | c_t = +1)$ being the conditional probability of the sensory response \hat{s}_t being positive given that the true category is light ($c_t = +1$), and $p(\hat{s}_t > 0 | c_t = -1)$ being the conditional probability of the same sensory response \hat{s}_t being positive given that the true category is dark ($c_t = -1$), the log-likelihood ratio is ideally $\text{LLR}_t = \log \frac{p(\hat{s}_t > 0 | c_t = +1)}{p(\hat{s}_t > 0 | c_t = -1)} = \log \frac{0.8}{0.2}$ for evidence towards $c_t = +1$ and, respectively, $\log \frac{0.2}{0.8}$ for evidence towards $c_t = -1$. In the second (‘non-categorized’) version of the model, the log-odds LLR_t is computed using the continuous normally distributed stimulus sensory response without any explicit categorization (Supplementary Fig. 9). The log-odds ratio between two symmetrical Gaussian distributions $\mathcal{N}(+1, \sigma_{\text{sensory}}^2)$ and $\mathcal{N}(-1, \sigma_{\text{sensory}}^2)$ or vice versa, simplifies to a scaled version of the sensory response by a factor $2/\sigma_{\text{sensory}}^2$ thus normally distributed with mean equal to $\pm 2/\sigma_{\text{sensory}}^2$ and s.d. equal to $2/\sigma_{\text{sensory}}^2$. The combination of the evidence provided by the stimulus (LLR_t) and the prior belief (L_{t-1}), also expressed as log-odds ratio and updated according to hazard rate h is then corrupted by Gaussian inference noise of s.d. σ_{inf} following Eq. 4. Inference noise captures both internal variability in the perceived log-odds ratio LLR_t and variability in the update of the belief L_t .

The response selection policy is based on the sign of the decision variable: the newly formed belief L_t . We considered a normative, ‘greedy’ selection process and a noisy selection process (Supplementary Fig. 1) modelled by sampling responses from the sign of a normally distributed decision variable with mean L_t and variance σ_{sel}^2 .

We introduced stabilizing parameters to the suboptimal Bayesian inference process described above, operating at different stages of information processing.

1. The perceptual bias model, controlled by a prior bias strength λ , shifts the sensory representation of the stimulus already expressed as log-odds

$$\hat{s} \sim \mathcal{N}(\pm 2/\sigma_{\text{sensory}}^2, (2/\sigma_{\text{sensory}}^2)^2) \text{ towards the prior belief } L_{t-1};$$

$$\hat{s} \sim N(\pm 2/\sigma_{\text{sensory}}^2, (2/\sigma_{\text{sensory}}^2)^2) + \lambda \times L_{t-1}. \quad (5)$$

In this case, the log-likelihood ratio differs from ideal as the integration range is shifted by $\lambda \times L_{t-1}$.

2. The inference lapse model, controlled by a fraction of lapses ρ_{lapse} , ignores a fraction of the stimuli whose evidence is not used to update the current belief:

$$L_{t_{\text{lapse}}} = \psi(L_{t_{\text{lapse}}-1}, h). \quad (6)$$

3. The conditional inference model, controlled by a reliability threshold δ , uses only the stimuli whose sensory responses exceed a fixed threshold, $|\hat{s}| > \delta$, to update the current belief:

$$L_{t_{|\hat{s}|>\delta}} = \psi(L_{t_{|\hat{s}|>\delta}-1}, h) + \text{LLR}_t. \quad (7)$$

In this case, the log-likelihood ratio LLR_t accounts for the presence of the reliability threshold by integrating conditional likelihoods above the reliability threshold (instead of zero for the standard model). If the sensory response does not exceed the reliability threshold, the stimulus is ignored, leading to:

$$L_{t_{|\hat{s}|>\delta}} = \psi(L_{t_{|\hat{s}|>\delta}-1}, h). \quad (8)$$

A discard rate—corresponding to the fraction of unreliable evidence not used for belief update—can be derived by integrating the normal distribution function $\mathcal{N}(\pm 2/\sigma_{\text{sensory}}^2, (2/\sigma_{\text{sensory}}^2)^2)$ on the remaining range $[-\delta; +\delta]$.

4. The confirmation bias model, controlled by a bias strength β , operates at the selection stage by systematically shifting the decision criterion in favour of the previous response:

$$r_t = \text{sign}(L_t - \beta \times r_{t-1}). \quad (9)$$

5. The response lapse model, controlled by a lapse rate ε , makes a fraction of ‘blind’ repetitions of the previous response irrespective of the current belief:

$$r_{t_{\text{lapse}}} = r_{t_{\text{lapse}}-1}. \quad (10)$$

We considered a ‘soft’ version of the conditional inference model (Supplementary Figs. 9 and 10), controlled by the same reliability threshold δ , not discarding but downweighting the stimuli whose sensory responses do not exceed the threshold. In this model, inference is performed at every trial but with two distinctly weighted log-likelihood ratio LLR_t , the first integrating likelihoods below and the second integrating likelihoods above the reliability threshold.

Fitting procedure. Due to the presence of internal noise during sensory processing and hidden-state inference, which propagates across trials, no solution for model likelihoods can be derived analytically. We therefore used simulation-based methods to obtain noisy likelihoods for tested sets of parameter values, which were fed to specific fitting algorithms capable of handling such noisy likelihood functions.

Instead of fitting the candidate models to all choices using simulation-based methods, we focused on the two behavioural signatures characteristic of reversal learning behaviour: the response reversal curve and the response switch curve. This focused fitting procedure aimed at testing which of the tested candidate models can best explain specifically our behavioural effects of interest. Note that fits to all choices yield the same results (Supplementary Fig. 11). We fitted the parameter values (h , σ_{inf} and the optional stabilizing parameter) which best explained these two psychometric curves, both evaluated on four trials preceding and four trials following a reversal, for a total of eight trials per reversal for each metric. We conducted model recovery analyses, detailed below, to ensure the reliability of the simulation-based fitting method outlined above.

In practice, we simulated model responses ($n = 1,000$) to the sequences of trials presented to participants, we then derived the corresponding mean and s.d. of reversal and switch curves (modelled as normally distributed variables at each trial) and estimated the log-likelihood of observed (participant) reversal and switch curves given model simulations. We combined the log-likelihood estimate with prior distributions on parameter values to get a log-posterior estimate, whose prior distributions over parameters were defined as truncated β and γ distributions (h : β distribution with shape parameters $\alpha = 2$ and $\beta = 18$, range $[0, 0.5]$; σ_{inf} : γ distribution with shape parameter $k = 4$ and scale parameter $\theta = 0.25$, range $[0, 5]$; λ and δ : γ distribution with shape parameter $k = 1$ and scale parameter $\theta = 0.75$, range $[0, 5]$; ρ_{lapse} and ε : β distribution with shape parameters $\alpha = 1$ and $\beta = 9$, range $[0, 1]$; β : γ distribution with shape parameter $k = 1$ and scale parameter $\theta = 0.25$, range $[0, 5]$). In the first experiment, we verified for the winning model that using looser prior distributions did not change the best-fitting parameter values (Spearman correlations: h : $\rho = 0.779$, $P < 0.001$, 95% CI (0.650, 0.784); σ_{inf} : $\rho = 0.854$, $P < 0.001$, 95% CI (0.763, 0.862); δ : $\rho = 0.834$, $P < 0.001$, 95% CI (0.670, 0.843); differences: h : $t_{28} = 0.508$, $P = 0.615$, Cohen’s $d = 0.094$, 95% CI $(-0.022, 0.013)$, $BF_{01} = 4.497$; σ_{inf} : $t_{28} = 0.863$, $P = 0.395$, Cohen’s $d = 0.16$, 95% CI $(-0.349, 0.142)$, $BF_{01} = 3.603$; δ : $t_{28} = 1.152$, $P = 0.259$, Cohen’s $d = 0.214$, 95% CI $(-0.326, 0.091)$, $BF_{01} = 2.778$).

The unnormalized log-posterior is given as argument for the Variational Bayesian Monte Carlo (VBMC) algorithm³⁴ (version 1.0; <https://github.com/lacerbi/vbmc>) returning a variational approximation of the full posterior and a lower bound on the log-marginal likelihood. The VBMC algorithm supports stochastic estimates of the log-posterior and we provide estimates of its s.d. by bootstrapping the s.d. of the estimated log-likelihood. We finally take the posterior mean to obtain best-fitting parameter values.

For the second experiment, we computed three distinct reversal curves and three distinct switch curves depending on the strength of the stimulus presented on the first trial following a reversal, for both model simulations and human data. We summed the log-likelihood across these six metrics.

We used the lower bound of the marginal likelihood as model evidence for Bayesian Model Selection (BMS) analysis. BMS was conducted using a random-effects approaches assuming that different participants may rely

on different models, and consists of estimating the distribution over models that participants draw from. We used the Dirichlet parameterization of the random-effects approach implemented in SPM12 (refs. ^{35,36}) with default Dirichlet prior set to one—corresponding to a uniform distribution (<http://www.fil.ion.ucl.ac.uk/spm>).

Our approach focusses on behaviour around reversals through the two psychometric curves described above. Alternatively, we considered fitting human data using per-trial choice log-likelihoods using a particle filter and the same VBMC algorithm. In this case, each trial—and not only trials around reversals—equally contributes to the goodness of fit, rather than focusing on our behavioural effects of interest around reversals. Yet, the more responses considered, the more likely a response can reflect a lapse and be a blind repetition of the previous response. To prevent the fitting procedure to try to fit those responses with the same importance as the other responses—especially the ones around reversal—ergo to prevent those lapse responses to corrupt the goodness of fit, we included a lapse rate parameter ϵ as additional parameter to each of the first four stabilizing models (the fifth previously described candidate model being fully formalized by the lapse rate parameter ϵ). The prior distributions on parameter values remain the same as previously reported, except for the additional parameter ϵ meant to be kept low (β distribution with shape parameters $\alpha = 1$ and $\beta = 39$, range $[0,1]$). The analysis of model fits to all choices using the same BMS yielded the same results as the ones obtained from model fits to behavioural effects of interest using our simulation-based approach (Supplementary Fig. 11).

Model recovery procedure. Model simulations were done for each participant's true stimulus sequences using parameters corresponding to the mean of their respective prior distributions. We used prior distribution to have an a priori assessment that the constructed models are well-posed. Each simulation was fitted by a given model and we used the Dirichlet parameterization as described above to estimate the distribution over recovered models. We then sampled 10^6 times from the estimated Dirichlet distributions and computed column-wise (for each recovered model) the exceedance probability of each simulated model.

Confidence fitting procedure. We modelled human behaviour in titration trials using three free parameters, the sensory noise s.d. σ_{sensory} , the confidence noise s.d. $\sigma_{\text{confidence}}$ and the confidence threshold $\delta_{\text{confidence}}$. Each stimulus sensory response at trial t is modelled as a normally distributed random variable generated by $\mathcal{N}(+1 \text{ or } -1, \sigma_{\text{sensory}}^2)$ depending on the staircase it originates from. Based on this sensory response, the noisy confidence is reported with respect to a confidence threshold $\delta_{\text{confidence}}$. We fitted this model using a particle filter ($n = 1,000$) to the factorized responses to titration trials (light and unconfident; light and confident; dark and unconfident; dark and confident) using VBMC to maximize the log-likelihood of factorized response probabilities.

PCA. PCA was performed on participants' response reversal and response switch curves using MATLAB in-built function which centres the data and uses the singular value decomposition algorithm. Based on the first three obtained PC scores, we estimated the explained variance in reversal and switch curves through bootstrapping ($n = 10,000$).

Multiple regression model. We predicted human and conditional inference model accuracies as a function of each parameter using a multiple regression model. We chose to include the perceived hazard rate h and the reliability threshold δ up to the second-order in the multiple regression, but the inference noise σ_{inf} appears only up to the first-order as it only corrupts accumulation of evidence. Second-order dependencies, if any, allow exhibition of inverted U-shaped relations with accuracy—ergo a beneficial effect on accuracy. As we investigate the benefits of the conditional inference strategy, the reliability threshold δ was included up to the second-order, and showed indeed an inverted U-shaped relation with accuracy, no discarding of evidence corresponding to a classical suboptimal inference model and a large δ implying omission of almost all evidence. The multiple regression model was fitted using the MATLAB in-built function `fitnlm.m` with robust fitting (weighted least-squares).

Statistical testing and reproducibility. Unless noted otherwise, statistical analyses of differences between scalar metrics in both experiments relied on two-tailed parametric tests (classical and Bayesian paired t -tests, repeated-measures ANOVA) between tested participants and model predictions in a paired manner using MATLAB functions, `simple_mixed_anova.m`³⁷ and JASP³⁸. For repeated-measures ANOVA, reported P values are Greenhouse–Geisser corrected. Given our sample sizes, these statistical tests were applied outside the small-sample regime. Data were not explicitly tested for normality. Reported statistics are not corrected for multiple comparisons. Given the absence of prior effect sizes, we chose a sample size ($n = 30$ for each experiment) that exceeded the average sample size used in human psychophysical studies with similar trial number per participant ($n = 640$ across two sessions).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated during and analysed during the current study are freely available online on figshare: https://figshare.com/projects/Efficient_stabilization_of_imprecise_statistical_inference_through_conditional_belief Updating/140170. Source data are provided with this paper.

Code availability

The analysis code supporting the reported findings is freely available online on github: <https://github.com/juliedrevet/CONDINF>.

Received: 19 July 2021; Accepted: 11 August 2022;

Published online: 22 September 2022

References

- Wald, A. & Wolfowitz, J. Optimum character of the sequential probability ratio test. *Ann. Math. Stat.* **19**, 326–339 (1948).
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765 (2006).
- Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (John Wiley, 1966).
- Drugowitsch, J., Wyart, V., Devauchelle, A.-D. & Koehlin, E. Computational precision of mental inference as critical source of human choice suboptimality. *Neuron* **92**, 1398–1411 (2016).
- Wyart, V. & Koehlin, E. Choice variability and suboptimality in uncertain environments. *Curr. Opin. Behav. Sci.* **11**, 109–115 (2016).
- Findling, C. & Wyart, V. Computation noise in human learning and decision-making: origin, impact, function. *Curr. Opin. Behav. Sci.* **38**, 124–132 (2021).
- Glaze, C. M., Kable, J. W. & Gold, J. I. Normative evidence accumulation in unpredictable environments. *eLife* **4**, 1–27 (2015).
- Murphy, P. R., Wilming, N., Hernandez-Bocanegra, D. C., Prat-Ortega, G. & Donner, T. H. Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments. *Nat. Neurosci.* **24**, 987–997 (2021).
- Palminteri, S., Wyart, V. & Koehlin, E. The importance of falsification in computational cognitive modeling. *Trends Cogn. Sci.* **21**, 425–433 (2017).
- Stocker, A. A. & Simoncelli, E. P. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* **9**, 578–585 (2006).
- Luu, L. & Stocker, A. A. Post-decision biases reveal a self-consistency principle in perceptual inference. *eLife* **7**, e33334 (2018).
- Lange, R. D., Chatteraj, A., Beck, J. M., Yates, J. L. & Haefner, R. M. A confirmation bias in perceptual decision-making due to hierarchical approximate inference. *PLoS Comput. Biol.* **17**, e1009517 (2021).
- Soltani, A. & Izquierdo, A. Adaptive learning under expected and unexpected uncertainty. *Nat. Rev. Neurosci.* **20**, 635–644 (2019).
- Griffiths, T. L. & Tenenbaum, J. B. Optimal predictions in everyday cognition. *Psychol. Sci.* **17**, 767–773 (2006).
- Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
- Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S. & Wyart, V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat. Neurosci.* **22**, 2066–2077 (2019).
- Peixoto, D. et al. Decoding and perturbing decision states in real time. *Nature* **591**, 604–609 (2021).
- Donoso, M., Collins, A. G. E. & Koehlin, E. Foundations of human reasoning in the prefrontal cortex. *Science* **344**, 1481–1486 (2014).
- Karlsson, M. P., Tervo, D. G. R. & Karpova, A. Y. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* **338**, 135–139 (2012).
- Glickman, M., Moran, R. & Usher, M. Evidence integration and decision confidence are modulated by stimulus consistency. *Nat. Hum. Behav.* **6**, 988–999 (2022).
- Salvador, A. et al. Premature commitment to uncertain decisions during human NMDA receptor hypofunction. *Nat. Commun.* **13**, 338 (2022).
- Lefebvre, G., Lebreton, M., Meyniel, E., Bourgeois-Gironde, S. & Palminteri, S. Behavioural and neural characterization of optimistic reinforcement learning. *Nat. Hum. Behav.* **1**, 1–9 (2017).
- Lefebvre, G., Summerfield, C. & Bogacz, R. A normative account of confirmation bias during reinforcement learning. *Neural Comput.* **34**, 307–337 (2022).
- Pisupati, S., Chartarifska-Lynn, L., Khanal, A. & Churchland, A. K. Lapses in perceptual decisions reflect exploration. *eLife* **10**, e55490 (2021).
- Ashwood, Z. C. et al. Mice alternate between discrete strategies during perceptual decision-making. *Nat. Neurosci.* **25**, 201–212 (2022).
- Vogel, T. A., Savelson, Z. M., Otto, A. R. & Roy, M. Forced choices reveal a trade-off between cognitive effort and physical pain. *eLife* **9**, e59410 (2020).

27. Moutoussis, M. et al. Decision-making ability, psychopathology, and brain connectivity. *Neuron* **109**, 2025–2040.e7 (2021).
28. Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A. & Daw, N. D. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* **5**, e11305 (2016).
29. Rouault, M., Seow, T., Gillan, C. M. & Fleming, S. M. Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biol. Psychiatry* **84**, 443–451 (2018).
30. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).
31. Kaernbach, C. Simple adaptive testing with the weighted up-down method. *Percept. Psychophys.* **49**, 227–229 (1991).
32. Kleiner, M. et al. What's new in Psychtoolbox-3. *Perception* **36**, 1–16 (2007).
33. Weiss, A., Chambon, V., Lee, J. K., Drugowitsch, J. & Wyart, V. Interacting with volatile environments stabilizes hidden-state inference and its brain signatures. *Nat. Commun.* **12**, 2228 (2021).
34. Acerbi, L. in *Advances in Neural Information Processing Systems* 33 (eds Larochelle, H. et al.) 8211–8222 (Curran Associates, 2020).
35. Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J. & Friston, K. J. Bayesian model selection for group studies. *NeuroImage* **46**, 1004–1017 (2009).
36. Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group studies - revisited. *NeuroImage* **84**, 971–985 (2014).
37. Caplette, L. Simple RM/Mixed ANOVA for any design. *MATLAB Central File Exchange* <https://www.mathworks.com/matlabcentral/fileexchange/64980-simple-rm-mixed-anova-for-any-design> (2022).
38. JASP Team. JASP v.0.16 <https://jasp-stats.org/> (2021).

Acknowledgements

We thank B. De Martino and M. Usher for their insightful comments and suggestions during peer review. This work was supported by the European Research Council (starting grant No. ERC-StG-759341 to V.W.), the National Institute of Mental Health (US–France collaborative research grant No. 1R01MH115554-01 to J. Drugowitsch and V.W.) and the Agence Nationale de la Recherche (grant No. ANR-17-NEUC-0001-02 to

J. Drugowitsch and V.W., and a department-wide grant No. ANR-17-EURE-0017). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

J. Drevet contributed to conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft, writing—review and editing, and visualization. J. Drugowitsch contributed to conceptualization, methodology, writing—review and editing, supervision, project administration and funding acquisition. V.W. contributed to conceptualization, methodology, software, validation, formal analysis, resources, writing—original draft, writing—review and editing, visualization, supervision, project administration and funding acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41562-022-01445-0>.

Correspondence and requests for materials should be addressed to Julie Drevet or Valentin Wyart.

Peer review information *Nature Human Behaviour* thanks Benedetto De Martino and Marius Usher for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection The experiment was coded in MATLAB R2018b and run using Psychtoolbox-3 (Kleiner et al., 2007). Participants performed the experiment in a soundproof booth with their head positioned on a chin rest at 75 cm from a 24-inch LCD screen operating at 60 Hz with a resolution of 1920 × 1080 pixels.

Data analysis The model parameters were fitted using the Variational Bayesian Monte Carlo (VBMC) algorithm (version 1.0; Acerbi, 2020; <https://github.com/lacerbi/vbmc>). The random-effects Bayesian model selection procedure was implemented in SPM12 (Wellcome Center for Human Neuroimaging; <http://www.fil.ion.ucl.ac.uk/spm>). The custom scripts used to analyze the behavioral data were written using MATLAB R2018b-R2020a and are freely available on github: <https://github.com/juliedrevet/CONDINF>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The datasets generated during and analyzed during the current study is freely available online on figshare:
https://figshare.com/projects/Efficient_stabilization_of_imprecise_statistical_inference_through_conditional_belief Updating/140170

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

Sixty healthy adult participants (30 females, age: 25.8 ± 3.9 years, all right-handed) took part in the study:
 Experiment 1: N = 30, 15 females (self-reported gender), age: 24.7 ± 3.9 years.
 Experiment 2: N = 30, 15 females (self-reported gender), age: 27.16 ± 3.7 years.
 Participants were screened for the absence of any history of neurological and psychiatric disease or any current psychiatric medication. All subjects had normal or corrected-to-normal vision.
 No sex- and gender-based analyses have been performed because we compare conditions in a within-subject design.

Population characteristics

See above.

Recruitment

Participants were recruited through online posting on the public mailing list from the Relais d'Information en Sciences Cognitives (RISC) in France. Subscribers to the public mailing list are typically young adults (university students in particular) in the age range that was targeted for our experiment (18-35 years). We do not expect any self-selection biases that could impact results.

Ethics oversight

All participants provided written informed consent. The study followed guidelines from the Declaration of Helsinki and specific procedures which received ethical approval from relevant authorities (Comité de Protection des Personnes Ile-de-France VI, ID RCB: 2017-A01778-45).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Across two independent datasets, participants were asked to play a sequential categorization reversal learning task based on carefully titrated noisy visual stimuli. The staircase procedure took place on separate blocks. To investigate how participants mitigate the significant costs of imprecise cognitive computations, we used computational modelling based on a noisy Bayesian inference process and compared participants' reversal learning behavior both qualitatively and quantitatively to the one predicted by five candidate models.

Research sample

Sixty healthy adult participants (30 females, age: 25.8 ± 3.9 years, all right-handed, Experiment 1: N = 30; Experiment 2: N = 30) took part in the study. One participant was excluded from analyses due to poor performance (overall decision accuracy more than 3.5 s.d. below the group-level mean). Participants were screened for the absence of any history of neurological and psychiatric disease or any current psychiatric medication. All subjects had normal or corrected-to-normal vision. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar or larger to those reported in previous publications (Drugowitsch, J., Wyart, V et al., 2016; Urai, A.E. et al. 2017).

Sampling strategy

No statistical methods were used to pre-determine sample sizes but our sample size (N = 30 for each experiments) matches the commonly accepted good practices in this field. In particular, the chosen sample size was determined a priori for both experiments. The eight blocks of the main task consisted of 640 trials and were interleaved with eight blocks for the titration procedure (400 trials on the first session block and 60 trials for the remaining blocks). The first generative category was counter-balanced pseudorandomly across blocks and participants. Given the absence of prior effect sizes, we chose a sample size (N = 30 for each experiment) which

	exceeded the average sample size used in human psychophysical studies with similar trial number per participant (N = 640 across two sessions).
Data collection	The experiment was coded in MATLAB R2018b and run using Psychtoolbox-3 (Kleiner et al., 2007). Participants performed the experiment in a soundproof booth with their head positioned on a chin rest at 75 cm from a 24-inch LCD screen operating at 60 Hz with a resolution of 1920 × 1080 pixels. Pupillometric signals were collected using an eye-tracker EyeLink-1000 (SR Research). Pupillometric signals were not analyzed in this paper. No one was present besides the experimenter and the participant during data collection. The experimenter was not blinded to experimental condition or study hypothesis.
Timing	The data for experiment 1 (N = 30 participants) was collected between November 2018 and January 2019. The data for experiment 2 (N = 30 participants) was collected between August 2019 and September 2019.
Data exclusions	One participant was excluded from analyses due to poor performance (overall decision accuracy more than 3.5 s.d. below the group level mean).
Non-participation	N=3 participants dropped out after the first session of the experiment because of very low-quality pupillometric signals (unable to reduce blinks, pupillometry was not analyzed in the current paper). They were compensated for their participation.
Randomization	Participants were not allocated into distinct experimental groups. We used a within-subject design, the main task blocks and the titration blocks were distinct in the instructions provided to the participants, and they were thus by definition not blind to participants.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Supplementary information

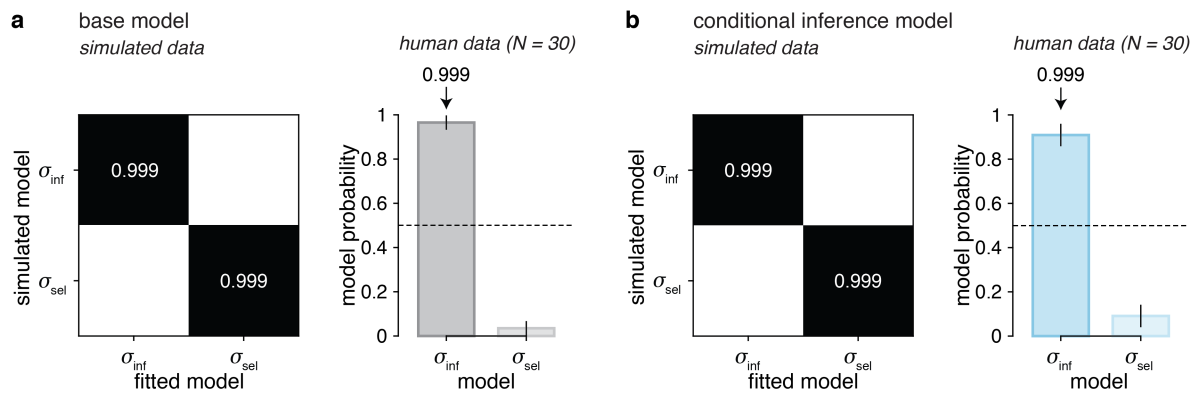
**Efficient stabilization of imprecise
statistical inference through conditional
belief updating**

In the format provided by the
authors and unedited

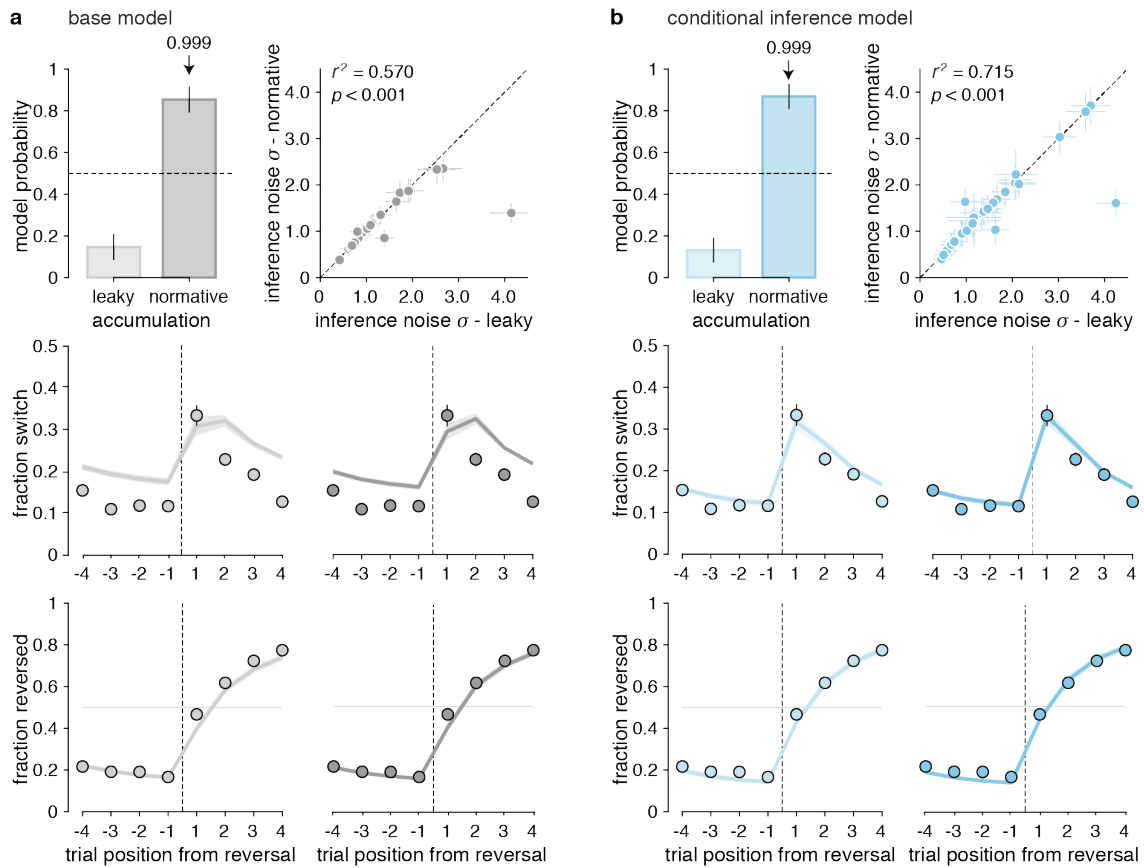
Supplementary Information

Efficient stabilization of imprecise statistical inference through conditional belief updating

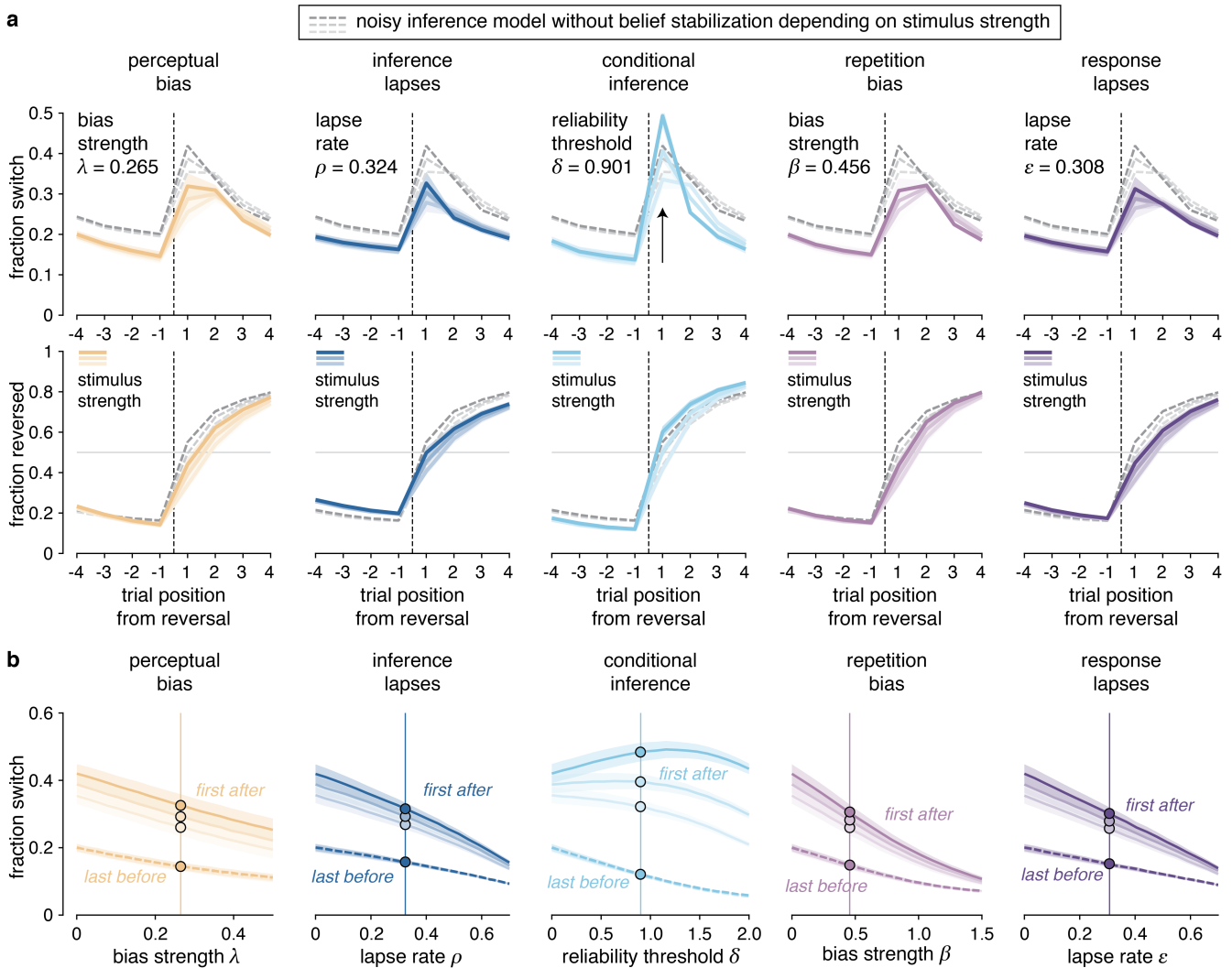
Julie Drevet, Jan Drugowitsch and Valentin Wyart



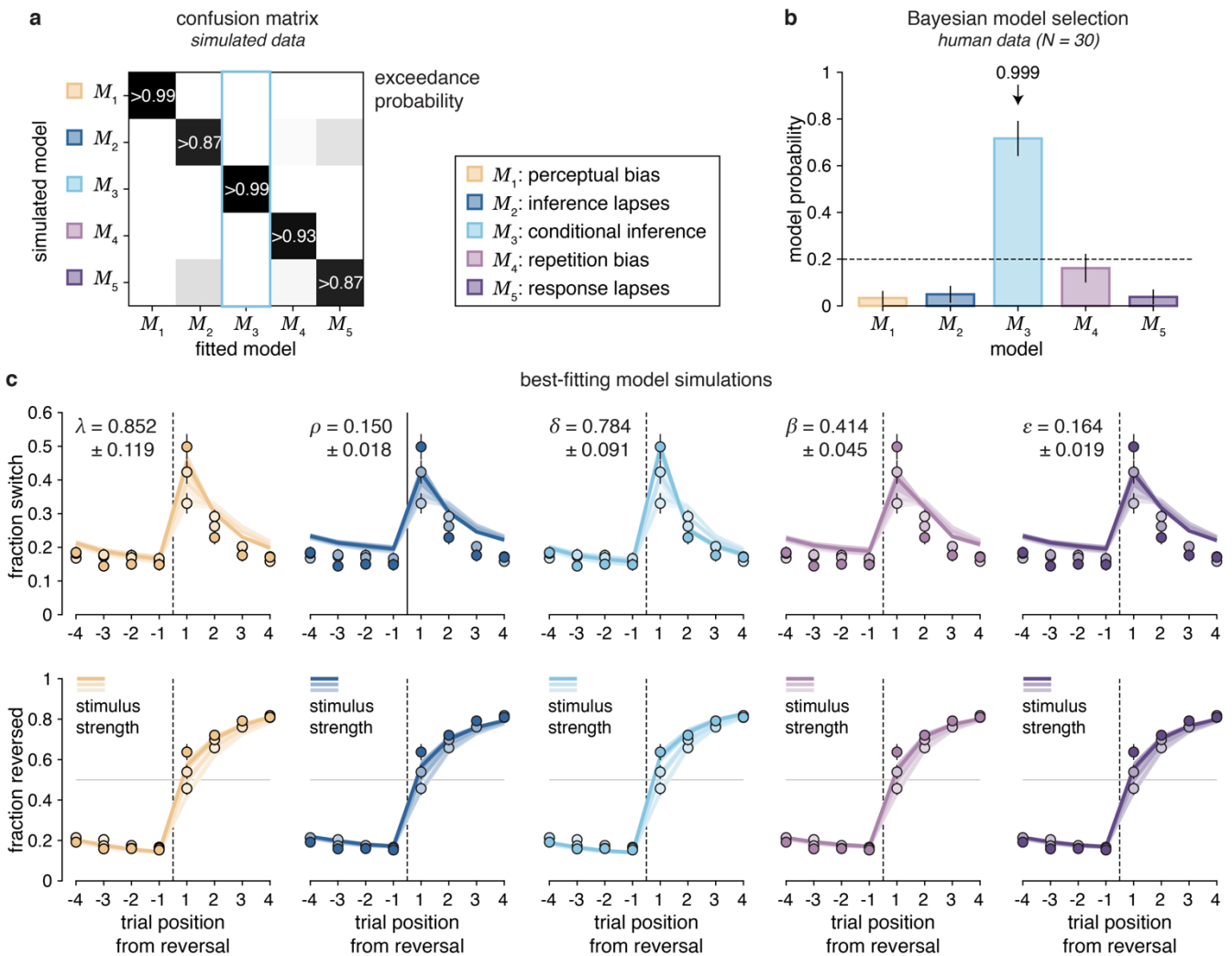
Supplementary Figure 1 | Comparison between inference and selection noise. (a) Bayesian inference model corrupted by internal noise (either inference noise or selection noise). Left: *ex-ante* confusion matrix from model recovery on simulated data, indicating that for each recovering noise source, the model used to simulate the data was the same with exceedance $p_{exc} > 0.999$. Right: estimated model probabilities for inference and selection noise when fitted to human data, Bayesian model selection shows a clear preference for inference noise over selection noise with exceedance probability $p_{exc} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution. (b) Conditional inference model. Left: *ex-ante* confusion matrix from model recovery on simulated data, indicating that for each recovering noise source, the model used to simulate the data was the same with exceedance $p_{exc} > 0.999$. Right: estimated model probabilities for inference and selection noise when fitted to human data, Bayesian model selection shows a clear preference for inference noise over selection noise with exceedance probability $p_{exc} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution.



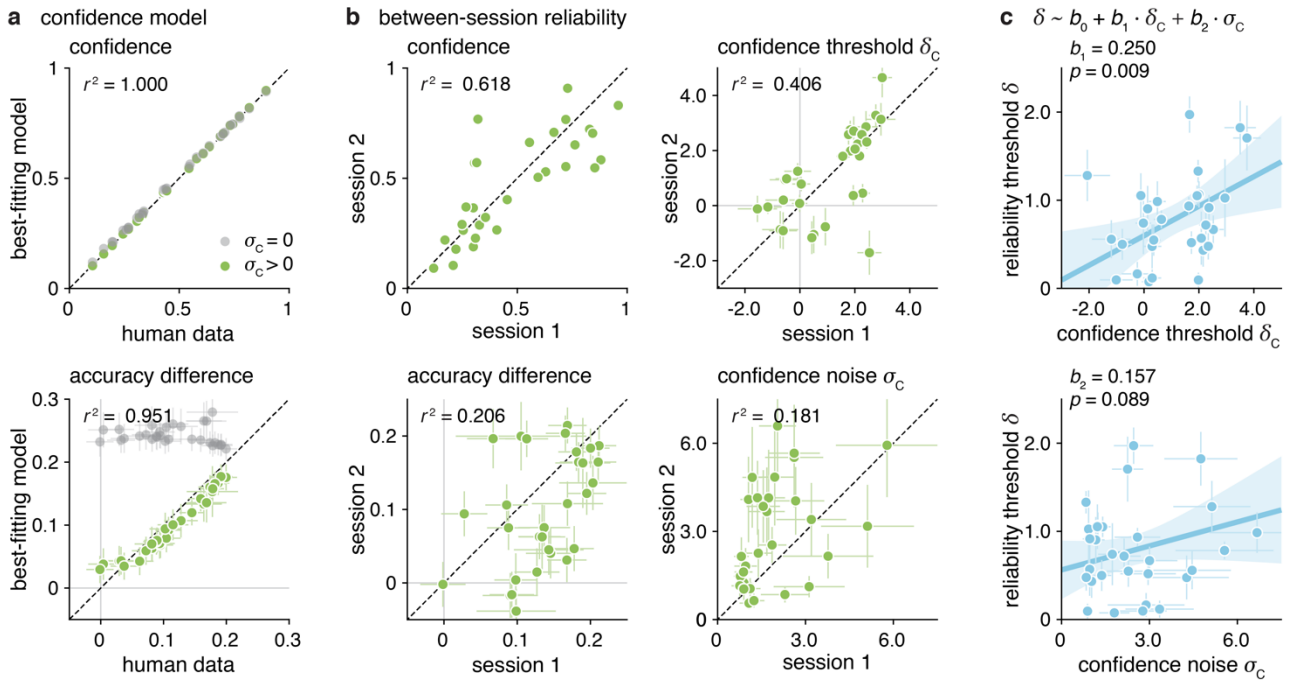
Supplementary Figure 2 | Comparison between leaky and normative evidence accumulation ($N = 30$). (a) Inference noise added to a leaky accumulation model or to the normative accumulation model. Top left: estimated model probabilities for leaky and normative accumulation when fitted to human data, Bayesian model selection shows a clear preference for normative over leaky accumulation with exceedance probability $p_{exc} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution. Top right: positive correlation between best-fitting inference noise parameter added to the leaky accumulation model and best-fitting inference noise parameter added to the normative accumulation model. Error bars correspond to posterior s.d. Bottom: Response reversal curves (top row) and response switch curves (bottom row) predicted by the best-fitting leaky accumulation model (light grey line) and best-fitting normative accumulation model (dark grey line). Both leaky and normative accumulation models capture correctly the accuracy of behavior surrounding reversals but overestimates the variability of the same behavior, especially the leaky accumulation model. The reversal is represented by the thin dotted line. Dots indicate human data (group-level average). Error bars and shaded areas correspond to s.e.m. (b) Same panels as in (a) but with additional conditional inference stabilization strategy (either leaky accumulation or normative accumulation). Similar amount of inference noise is found in both accumulation models.



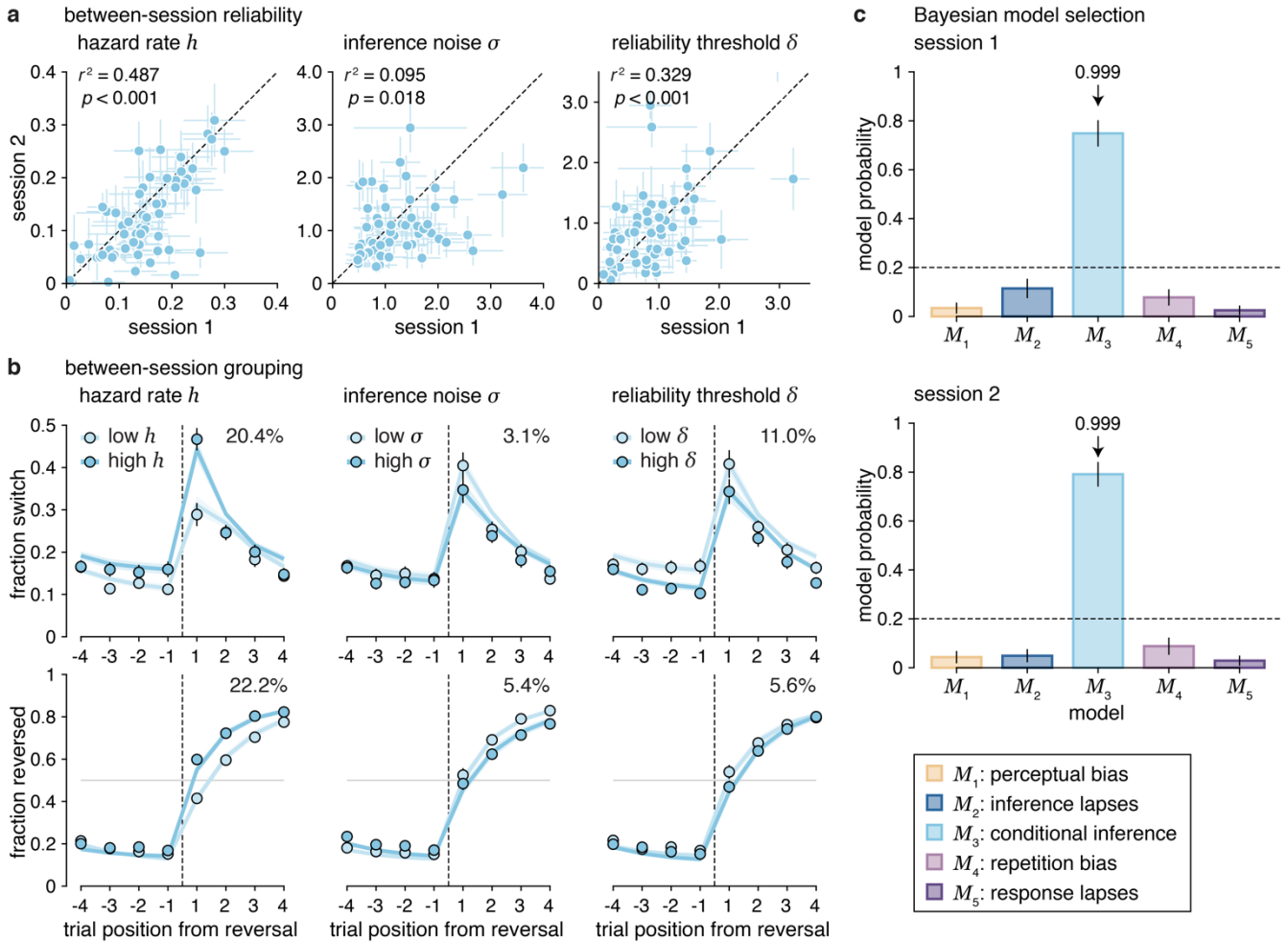
Supplementary Figure 3 | Predicted effects of response stabilization strategies. (a) Simulated effects of the five candidate strategies on response switch curves (top row) and response reversal curves (bottom row). The parameters controlling each response stabilization strategy (λ , δ , ρ , β , ε) is set to match participants' overall switch rate. The other parameters (h , σ_{sen} , σ_{inf} , σ_{sel}) are fixed to their best-fitting values for the noisy inference model without response stabilization. Solid colored lines correspond to the reversal behavior of each candidate model depending on the stimulus strength in the trial following a reversal. Dotted gray lines correspond to the reversal behavior of the noisy inference model without response stabilization depending on the stimulus strength in the trial following a reversal (same for all panels). Stronger color corresponds to stronger stimuli. The conditional inference strategy stands out from other strategies on the first trial after reversal (arrow). Shaded areas around curves correspond to s.e.m. (b) Simulated effects of the five candidate strategies on response switches just before and after a reversal. Fraction of response switches on the last trial before each reversal (dotted lines) and the first trial after each reversal (solid lines) for each belief stabilization strategy. Stronger colors correspond to a stronger (more informative) stimulus on the first trial following a reversal. All candidate strategies reduce simultaneously response switches before and after reversals, except for conditional inference which reduces response switches only before reversals (i.e., when they are not warranted). Dots correspond to stabilization parameters set as in (a). Shaded areas correspond to s.e.m.



Supplementary Figure 4 | Belief stabilization through conditional inference – Experiment 2 ($N = 30$). (a) Confusion matrix for *ex-ante* model recovery. For each recovering model, the model used to simulate the data was the same with exceedance $p_{\text{exc}} > 0.87$ and especially for the conditional inference model the exceedance reached is $p_{\text{exc}} > 0.999$. (b) Bayesian model selection of response stabilization strategy. Estimated model probabilities for the five candidate models. The conditional inference model outperforms other models with exceedance $p_{\text{exc}} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution. The dashed line corresponds to the uniform distribution. (c) Simulations of response stabilization strategies fitted to the human data. Simulations of three switch curves (top row) and three reversal curves (bottom row) are based on best fitting parameters for each stabilizing model. Dots indicate human data (group-level average), whereas lines indicate model predictions. The mean best-fitting parameter of each model (mean \pm s.e.m.) is indicated in the top-left corner of each subpanel. Stronger colors correspond to a stronger (more informative) stimulus on the first trial following a reversal. Only the conditional inference model qualitatively reproduces humans stabilized behavior around reversal and the large transient increase of switches on the trial just after reversal. Shaded areas and error bars correspond to s.e.m.

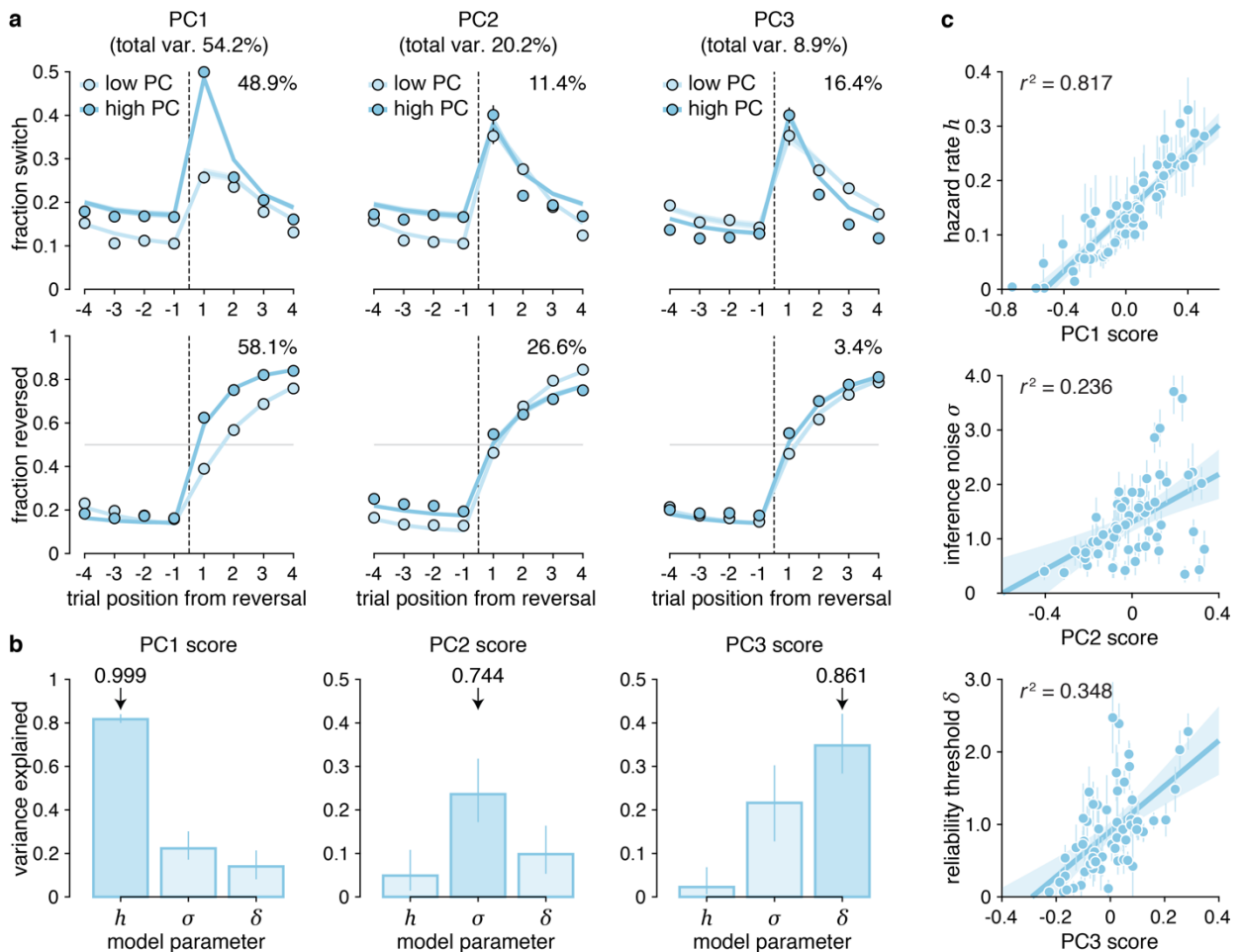


Supplementary Figure 5 | Decision confidence model ($N = 30$). (a) Reliability of confidence model predictions. Correlation between observed data and predicted confidence (top) and predicted accuracy difference between confident and unconfident trials (bottom) of a confidence model with (green) or without (gray) confidence noise σ_c . All data are presented as mean values and error bars correspond to s.d. of the accuracy difference of best-fitting model simulations and estimated s.d. of human accuracy difference given the number of trials provided. (b) Between-session reliability in confidence reports and parameters. Reported confidence and best-fitting parameters found for the first experimental session correlate positively with reported confidence and best-fitting parameters found for the second session. Data are presented as mean values. Error bars on best-fitting parameters correspond to s.d. of posterior distribution and error bars on accuracy difference correspond to estimated s.d. of human accuracy difference given the number of trials provided. (c) Reliability of positive relation between reliability threshold and confidence threshold. Linear Regression between reliability threshold and confidence threshold accounting for internal variability in confidence reports. Parameters are presented as mean \pm s.d. of posterior distributions of each fit. Solid blue line correspond to the best-fitting regression line (in the least-squares sense) and shaded area corresponds to the 95% confidence interval for the regression line.

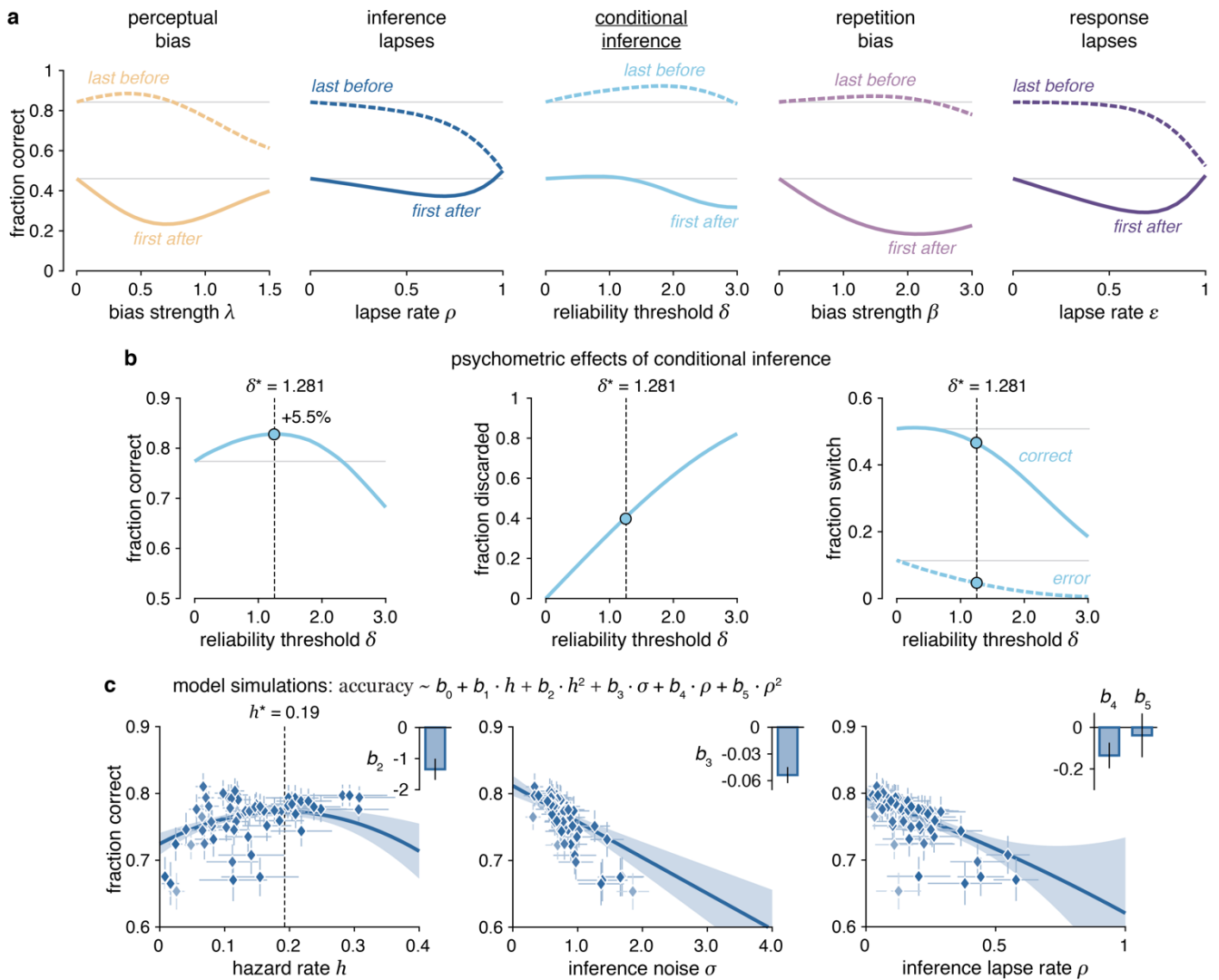


Supplementary Figure 6 | Between-session reliability in conditional inference model parameters ($N = 60$).

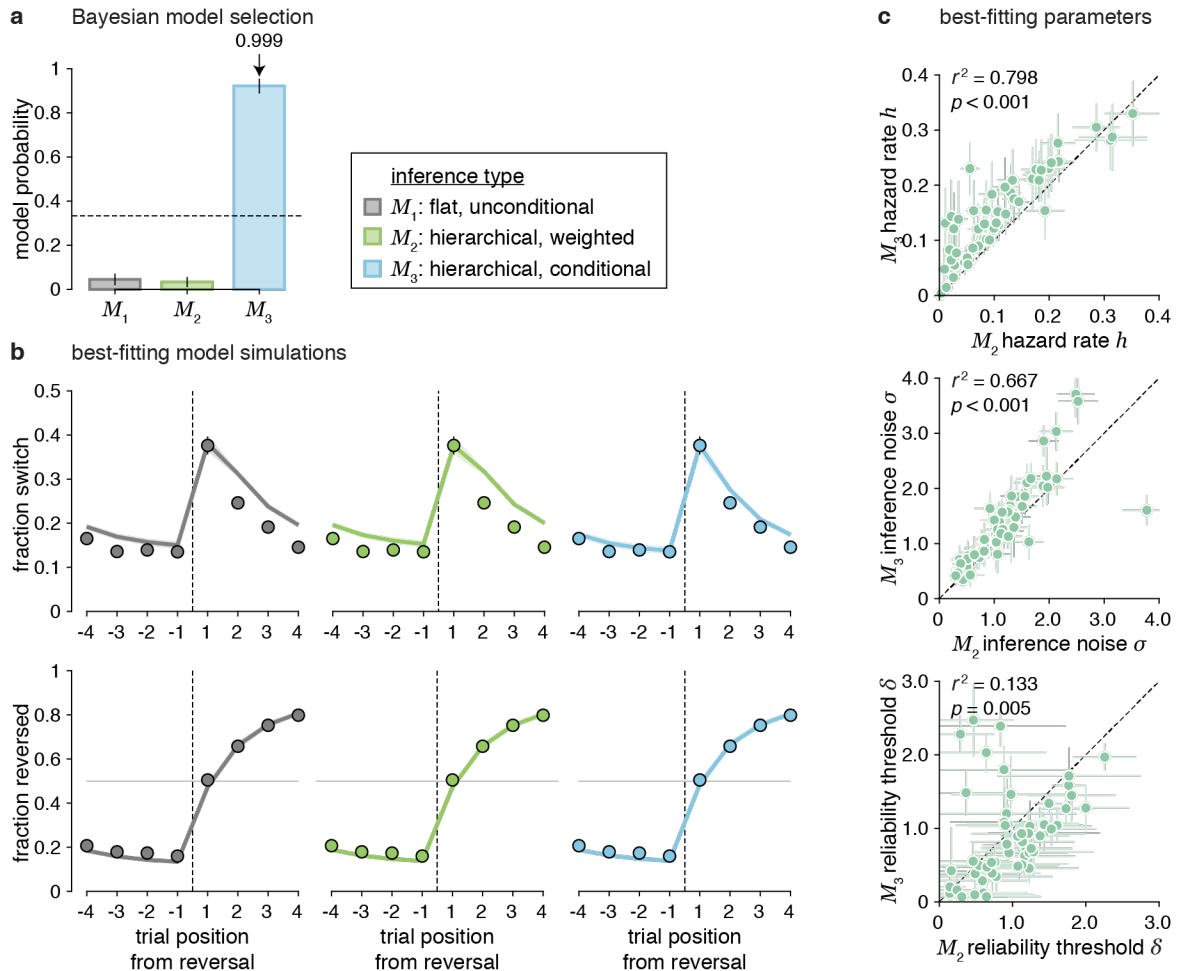
(a) Between-session reliability in model parameter values. Best-fitting parameters found for the first experimental session correlate positively with best-fitting parameters found for the second session. Parameters are presented as mean \pm s.d. of posterior distributions of each fit. Dotted lines correspond to the identity line. **(b)** Between-session (cross-validated) effect on response switch and reversal curves. Median-split on each best-fitting parameter for conditional inference model simulations (lines) and observed data (dots) and corresponding variance explained by each parameter. Error bars and shaded areas correspond to s.e.m. **(c)** Session-wise Bayesian model selection. The conditional inference model outperforms other models for both sessions with exceedance $p_{exc} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution.



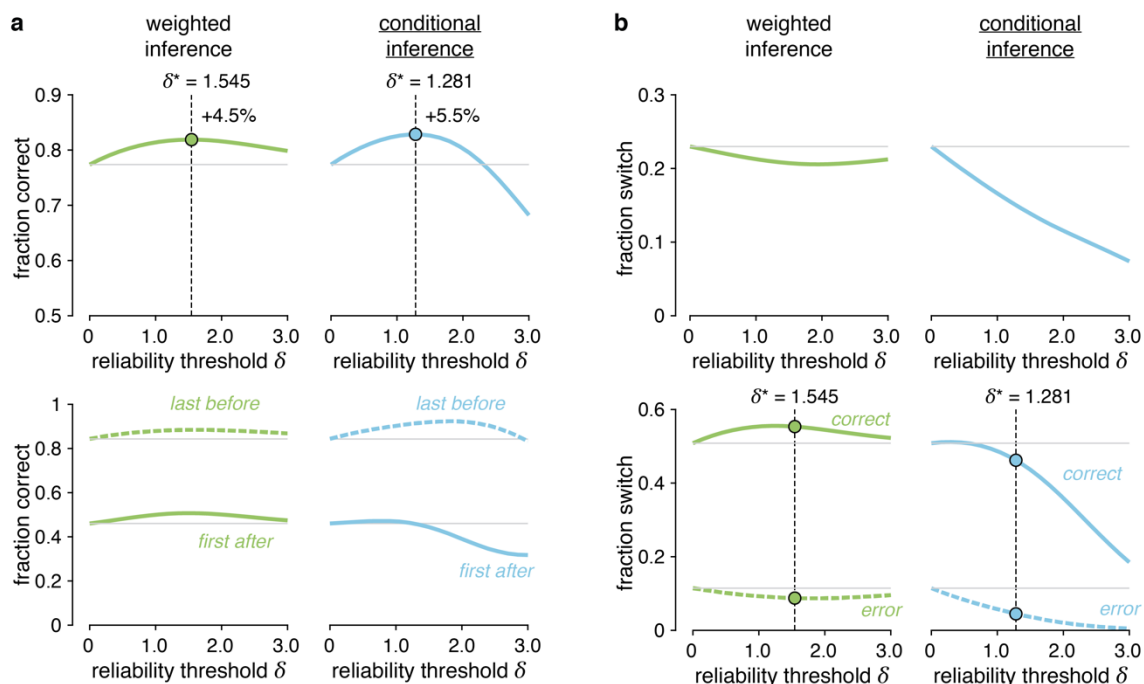
Supplementary Figure 7 | Principal Component Analysis of interindividual variability ($N = 60$). **(a)** Behavioral gradients associated with the first three PCs. Median-split on each principal component score for conditional inference model simulations (lines) and observed data (dots) and corresponding variance explained by each principal component. Error bars and shaded areas correspond to s.e.m. **(b)** Relations between each PC and model parameters. Each conditional inference best-fitting model parameter explains best the variance of one principal component score. The perceived hazard rate best explains the first PC score with exceedance $p_{\text{exc}} > 0.999$, the inference noise best explains the second PC score with exceedance $p_{\text{exc}} > 0.744$ and the reliability threshold best explains the third PC score with exceedance $p_{\text{exc}} > 0.861$. Bars correspond to r^2 and error bars to the interquartile ranges of each r^2 measure obtained through bootstrapping ($N = 10^4$). **(c)** Relation between PC scores and their associated parameters. Each principal component score significantly correlates with one conditional inference model parameter. Solid blue line corresponds to the best-fitting regression line (in the least-squares sense) and shaded area corresponds to the 95% confidence interval for predicted values. Parameters are presented as mean \pm s.d. of posterior distributions.



Supplementary Figure 8 | Conditional inference model effects. (a) Simulated effects on response accuracy before and after a reversal. Accuracy computed on the trial just before (dotted lines) and just after (solid lines) a reversal depending on each stabilization parameter. All stabilization parameters reduce accuracy on the trial following a reversal except for the conditional inference model which maintains the accuracy stable. The asymptotic accuracy reached before reversal is increased by three stabilization parameters: bias strength, reliability threshold and bias strength with a stronger and systematic improvement provided by the conditional inference model. (b) Psychometric effects of conditional inference. Left: Overall accuracy is increased when introducing a conditional threshold by +5.5% for $\delta^* = 1.281$ (see also main Figure 9a). Middle: fraction of discarded evidence depending on the reliability threshold (41.3% for $\delta^* = 1.281$). Right: Fraction of switch leading to a correct response (solid line) and leading to an error (dotted line). (c) Predicted effects of best-fitting inference lapses modeled decision accuracy (robust multiple regression: r -squared = 0.888, $p < 0.001$). For each best-fitting inference lapse model parameter, corresponding inference lapse model predicted accuracy (diamonds) and robust regression (solid blue lines). The lapse rate quadratic regressor b_5 is not significant ($p = 0.72$), therefore no inverted U-shaped relation to accuracy emerges. Solid blue line corresponds to the best-fitting robust regression line (in the weighted least-squares sense). Shaded areas correspond to the 95% confidence interval for predicted values. Parameters are presented as mean \pm s.d. of posterior distributions and vertical error bars correspond to the accuracy s.d. of best-fitting model simulations. Top right insets correspond to estimated regression coefficients presented as mean \pm s.e.m.

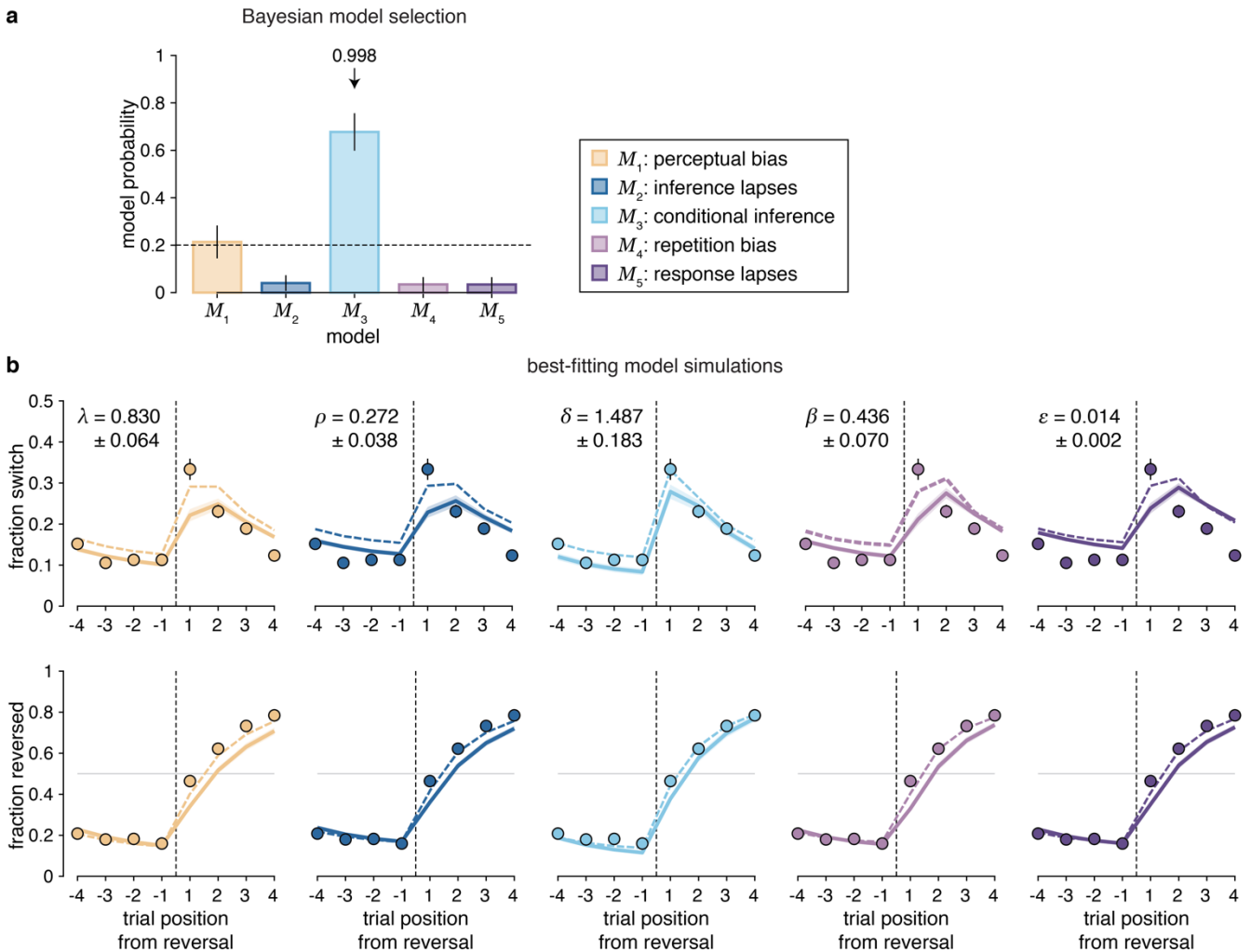


Supplementary Figure 9 | Comparison between flat unconditional, hierarchical weighted and hierarchical conditional inference models ($N = 60$). (a) Estimated model probabilities for flat, unconditional inference, for hierarchical, weighted, and for hierarchical, conditional inference when fitted to human data from both experiments pooled, random-effects Bayesian model selection provides clear support for the hierarchical conditional inference model with exceedance probability $p_{exc} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution. Dashed lines correspond to the uniform distribution. (b) Simulations of flat unconditional, of hierarchical weighted and of hierarchical conditional inference model fitted to participants' (experiments pooled). Simulated response switch curves (top row) and response reversal curves (bottom row) based on the best-fitting parameters of each model. Solid colored lines correspond to the reversal behavior of each stabilized model whereas dots indicate human data (group-level average). If the flat unconditional model reproduces well participants' reversal behavior for both metrics, especially on the trial just following a reversal for the response switch curve – unlike most other stabilization models considered in this paper – it fails at reproducing participants' stabilized behavior on the following trials as well as the hierarchical conditional model. The hierarchical weighted inference model reproduces well participants' fraction reverse but fails at reproducing their stabilized behavior. Shaded areas and error bars correspond to s.e.m. (c) Best-fitting parameter correlations. Weighted and conditional hierarchical inference model have the same free parameters and their best-fitting values correlate positively. Parameters are presented as mean \pm s.d. of posterior distributions. Dotted lines correspond to the identity line.



Supplementary Figure 10 | Comparison between weighted inference and conditional inference model effects.

(a) Simulated effects of weighted and conditional inference models on decision accuracy. Top: Fraction of overall correct responses when increasing the reliability threshold. The accuracy is increased by +4.5% for $\delta^* = 1.55$ for the weighted inference model, a little less than for the conditional inference model increasing the accuracy by +5.5% for $\delta^* = 1.28$. Dots correspond to the best simulated accuracy. Down: Simulated effects on response accuracy before and after a reversal. Accuracy computed on the trial just before (dotted lines) and just after (solid lines) a reversal depending on the reliability threshold for weighted and conditional inference models. Both models increase the asymptotic accuracy reached before a reversal and the weighted inference model maintain the accuracy even more stable on the trial after a reversal than the conditional inference model. **(b)** Simulated effects of weighted and conditional inference models on switching behavior. Top: weighted inference barely reduces overall behavioral variability compared to the conditional inference model. Infinitely large values of the reliability threshold predict the same effect on the fraction switch as no reliability threshold at all. Bottom: Fraction of switch leading to a correct response (solid line) and leading to an error (dotted line) for both models.



Supplementary Figure 11 | Belief stabilization through conditional inference – fit to each choice ($N = 30$).

(a) Random-effects Bayesian model selection of the best-fitting strategy in each participants' choice – instead of focusing only on the two reversal metrics. Bars indicate the estimated model probabilities for the five candidate strategies for experiment 1. Again, the conditional inference model (M_3) outperforms other candidate strategies with exceedance $p_{exc} > 0.999$. Model probabilities are presented as mean and s.d. of the estimated Dirichlet distribution. Dashed line corresponds to chance probability. **(b)** Simulations of response stabilization strategies fitted to each participants' choice. Simulated response switch curves (top row) and response reversal curves (bottom row) based on the best-fitting parameters of each model. The mean best-fitting stabilizing parameter \pm s.e.m. is indicated in the top-left corner for each model. Solid colored lines correspond to the reversal behavior of each stabilized model with parameters best-fitting each participant's choice, whereas dashed colored lines correspond to the reversal behavior of each stabilized model with parameters best-fitting these characteristic two metrics around reversal (same as main Figure 4). Dots indicate human data (group-level average). Again, the conditional inference model reproduces participants' reversal behavior better than any other models. Both curves are qualitatively less accurate when maximizing choice probability log-likelihood instead of these both metrics log-likelihoods. Shaded areas and error bars correspond to s.e.m.

INSTRUCTIONS

In this experiment, we consider 2 bags of marbles :
The bag with light marbles and the bag with dark marbles



There will be **two games**:

1. either you will have to **determine the color** of the marble to store it in the correct bag
2. or you will have to **determine from which bag** the computer draws the marbles

1

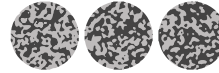
Difficulty

the marbles in the bags are all **bicolor**, yet always

either predominantly light in the light bag



or predominantly dark in the dark bag



To determine the color of the marble will not always be obvious, hence the difficulty of the two games!

2

Task course

Game #1 « I store the marble in the bag »

a marble appears quickly on the screen
if it is mostly light, you will have to store it in the light bag,
if it is mostly dark, you will have to store it in the dark bag.



Game #2 « From which bag was the marble drawn »

the computer draws a marble in one of the two bags,
the drawn marble appears quickly on the screen, you will have
to determine from which bag the marble was drawn.



3


Game #1

« I store the marble in the bag »

at each trial, a marble to be stored is presented,
then the two bags appear on the screen



you will have to give you answer by choosing one of the
bag depending on the side on which it appears on the screen.
[experiment 2: indicating how sure you are]

for each sorted marble, you will hear a sound  that
will indicate if you stored the marble correctly

4

Game #2

« From which bag was the marble drawn »

at each trial, a **drawn marble** is presented,
then the two bags appear on the screen



you will have to give you answer by choosing one of the
bag depending on the side on which it appears on the screen.

there will be no sound to indicate if the chosen bag was
the correct one!

5

Special feature of Game #2

« From which bag was the marble drawn »

The computer will **not draw randomly!**


It will start drawing marbles in one bag for a given number of trials,
then it will switch to the other bag in which it will draw marbles for
(another) given number of trials, and so on.

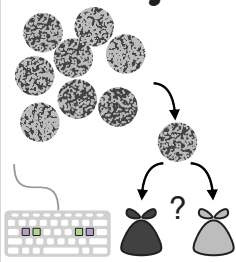
Be careful! The question is not « what is the color of the presented
marble » but « from which bag was the marble drawn », to be
among the best players, you will have to consider the marble
that was just presented, **but also the previous trials!**

6

Game #1


« I store the marble in the bag »

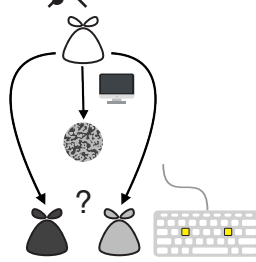
- guess **the color** of the marble
- marbles are shown **randomly**
- after each trial: 



Game #2

« From which bag was the marble drawn »

- guess **from which bag** comes the marble
- the computer draws **serials of marbles**
- after each trial: 



- After each block, your performance will be compared to other players that already played at this task.

- Stars will indicate if you are among the best players for this block:

- ★ ★ ★ worse than average
- ★ ★ ★ better than average
- ★ ★ ★ better than 75% of the players
- ★ ★ ★ better than 95% of the players !

8

- each game 1 is followed by a game 2
- there will be 8 blocks of (Game #1 + Game #2) per session
- try to memorize the instructions to be sure not to not mix up Game #1 and Game #2!
- the side of the bags will be assigned randomly, in does not have any influence on the task
- try to give the better and the prompter answer you can
- you will start with a short training
- the session lasts for 1h30min, there will be breaks between blocks.

Don't hesitate if you have any question!

Good luck!

9