Discrete confidence levels revealed by sequential decisions

Matteo Lisi[a,*]
Gianluigi Mongillo[b,c]
Georgia Milne[d]
Tessa Dekker[d,e]
Andrei Gorea[f,2]

[a] Department of Psychology, University of Essex, Colchester, UK
[b] Institut de la Vision, Sorbonne Université, Paris, France
[c] Centre National de la Recherche Scientifique, Paris, France
[d] Institute of Ophtalmology, University College London, London, UK
[e] Division of Psychology and Language Sciences, University College London, UK
[f] Université Paris Descartes, Sorbonne Paris Cité, Paris, France

[*] Corresponding author: m.lisi@essex.ac.uk; ORCID 0000-0003-3554-385X

[2] Deceased January 7th, 2019.

Abstract

Humans can meaningfully express their confidence about uncertain events. Normatively, these beliefs should correspond to Bayesian probabilities. However, it is unclear whether the normative theory provides an accurate description of the human sense of confidence, partly because the self-report measures used in most studies hinder quantitative comparison with normative predictions. To measure confidence objectively, we developed a dual-decision task in which the correctness of a first decision determines the correct answer of a second decision, thus mimicking real life situations in which confidence guides future choices. Whilst participants were able to use confidence to improve performance, they fell short of the ideal Bayesian strategy. Instead, behavior was better explained by a model with a few discrete confidence levels. These findings question the descriptive validity of normative accounts, and suggest that confidence judgments might be based on point-estimates of the relevant variables, rather than on their full probability distributions.

Introduction

Decision making is faced with uncertainty for every organism. A growing amount of evidence indicates that humans use estimates about the uncertainty of sensory data to guide their behaviour, for example when combining information from multiple sources[1]. Moreover, other studies have shown that humans can judge the accuracy of their perceptual decisions (an ability that belongs to the domain of metacognition) and report degrees of confidence that correlates with their objective performance[2]. This sense of confidence is an essential component of decision making (e.g. the decision to undertake a risky operation depends on the confidence in a correct diagnosis), but the underlying computations are unclear. According to one popular view confidence is Bayesian: it should coincide with the posterior probability that a choice was correct, computed following the rules of Bayesian inference. Although human confidence judgements tend to be qualitatively consistent with statistical principles[3], such qualitative agreement is neither necessary nor sufficient to conclude that the underlying computations are indeed Bayesian[4] - in the sense that they explicitly use probability distributions to quantify uncertainty. Indeed, other studies have shown deviations from Bayesian optimality[5], which suggest that confidence judgements could instead be based on simpler or approximated strategies.

One major issue for testing the Bayesian account of human confidence is methodological, and lies in the difficulty of measuring subjective probabilities, i.e. the degrees of belief that a human observer assigns to events. Subjective probability can be formally defined as the marginal rate of substitution[6,7]: our subjective probability about the event E occurring or having occurred would be $p(E)$ if we are indifferent to gaining one unit of utility contingent on E against gaining $p(E)$ units of utility for sure. Methods that measure subjective probabilities in perceptual decisions using opt-out or post-decision wagering techniques are straightforward operationalizations of this definition. A well known problem with these methods, however, is that they rely on unverifiable assumptions about the utility function of the participant[7] - such methods cannot disentangle subjective probability from factors such as opportunity costs. More frequently used methods that requires explicit confidence ratings suffer from miscalibrations and response biases[2,8,9] that hinder a rigorous, quantitative comparison of human behavior with the predictions of the Bayesian account. Yet another approach involves confidence forced-choice paradigms[10], in which observers engaging in perceptual discrimination can improve their performance by choosing the less uncertain among pair of stimuli. While this technique allows quantitatively accurate measurement of the ability to discriminate different levels of uncertainty across stimuli and tasks, it cannot detect global biases in confidence: observers may make accurate comparative judgements while overestimating or underestimating by a large amount their probability of being correct.

Here, we present an approach for studying subjective probabilities that allows us to detect confidence biases and directly compare human behaviour with the ideal Bayesian strategy and alternative models. In our experiments, human observers were presented with two consecutive signals and asked to decide whether they were above or below a reference value. The key innovation was that the statistics of the second signal were made contingent upon the decision-maker making a correct decision about the first signal. More specifically, correct/incorrect first decisions resulted in signals above/below the reference value for the second decision. An implementation of this experimental protocol is shown in Fig. 1A: participants first indicate whether the orientation of the random array of lines is tilted left or right from vertical (first decision). After this response, they are presented with a dynamic random-dot motion stimulus, and asked to report its direction (second decision). Crucially, the direction of the random-dot stimulus depends deterministically on the correctness of the first decision: if the first decision was correct, then the motion will be to the right in the second decision, otherwise it

will be to the left. Participants were informed about this rule, and were requested to use it to maximize their accuracy. It was impossible to know with certainty whether the first decision was correct and what the second response should be, as the difficulty of both stimuli was varied randomly from trial to trial, independently of whether the correct response was left or right. However, participants could assess their confidence in being correct on the first decision and use it to update prior expectations in the second decision and improve their performance. This paradigm mimics realistic situations in which confidence in past decisions, in absence of immediate feedback, is used to guide behaviour and inform future decisions, rather than being reported onto a rating scale without real consequences.
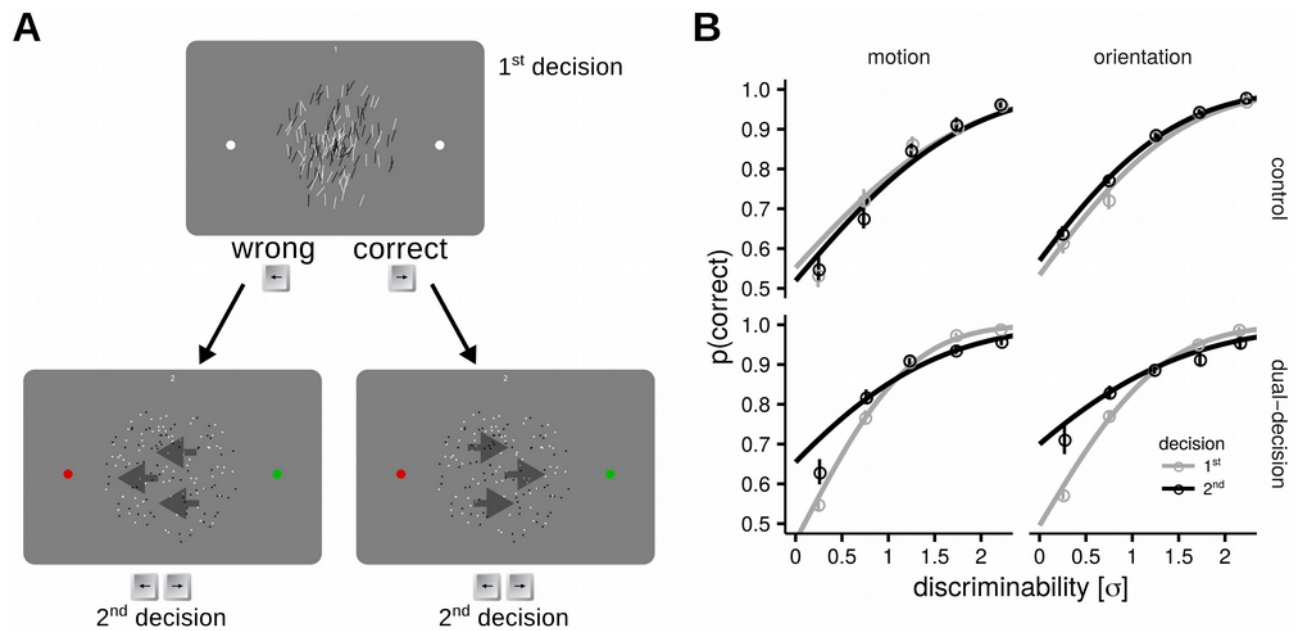
Figure 1. Dual-decision paradigm and results. A. Dual-decision paradigm: in this example the first decision is about the average tilt of an array of lines, and is followed by a second decision about the predominant direction of a random dot motion pattern. The central area of the screen where the stimulus was presented was sided by two dots, which were white for the first decision but turned red and green for the second one, to remind which directions were to be expected in case of incorrect and correct first response, respectively. The difficulty of both decisions (the opposite of the evidence discriminability, e.g. mean tilt from vertical or motion coherence, quantified as standard deviation of internal noise, σ) is drawn from the same uniform distribution (and therefore is, on average, equal), however the direction of the signal in the second decision (left vs right) was made dependent on the accuracy of first response. Each participants ran the task in both ways (i.e. orientation-motion and motion-orientation) in both the dual decision task and in a control condition where the direction of the signal was randomly drawn regardless of their responses. **B.** Performance plotted as a function of discriminability in the control conditions (top facets) and in the dual-decision conditions (lower facets). In the dual-decision task performance at either of the two visual tasks is higher in the second decision, especially for trials where the discriminability is small and, in the absence of prior information, performance near chance. Note that the data was grouped according to the perceptual task (motion, orientation) to highlight the effect of decision order within a single perceptual task. Thus, the panels displaying performance in motion and orientation judgments contain data from two different sessions with alternated task order. Error bars represents standard errors of the mean across participants; the curves show probit fits to the aggregated data.

4

Results

Confidence as a prior in sequential decisions: the dual-decision task

The paradigm used in the experiments is illustrated in Fig. 1A with an example of the orientation-motion discrimination task (experiment 1, see Methods for details). Participants made a first decision about the mean tilt of the array of lines, followed by a decision about the direction of a random-dot motion stimulus. Correct/incorrect first decisions determined whether the direction of the random dot motion pattern was rightward/leftward, respectively, and this rule was known to participants. All participants also ran a version of the experiment with the two discrimination tasks presented in the reverse order (first motion then orientation), see Methods for details. To control for sequential effects on responses between first and second decisions we included a control condition in which the second stimulus was independent of the first decision. We collected 2 independent datasets (with dataset 2 being a replication run on a different cohort of participants; see Methods). For completeness here we present statistical tests conducted both on the aggregated data as well as on the two datasets separately. Overall,we found a significant effect of decision order in the dual-decision condition (Fig. 1B): the mean odds-ratio of participants responding correctly in the second decision relative to the first one was greater than one: 1.51, 95% CI [1.28, 1.61], $t(28)=6.03$, $p=1.66\times10^{-6}$. The effect was significant when tested separately in dataset 1, mean odds-ratio 1.48, 95% CI [1.24, 1.64], $t(13)=4.78$, $p=7.19\times10^{-4}$; as well as in dataset 2, mean odds-ratio 1.55, 95% CI [1.21, 1.74], $t(14)=3.88$, $p=3.33\times10^{-3}$ (Bonferroni-corrected p-values). There was no statistically significant effect of decision order in the control condition: 1.07, 95% CI [0.96, 1.19], $t(28)=1.21$, $p=0.23$; the estimated Bayes factor (null/alternative) indicated that the data were 3.46 times more likely under the null hypothesis, indicating moderate support for the null hypothesis. We obtained similar results when analyzing the control condition separately for each dataset: dataset 1, mean odds-ratio 0.97, 95% CI [0.84, 1.15], $t(13)=0.33$, $p=1.00$; dataset 2, mean odds-ratio 1.17, 95% CI [1.00, 1.33], $t(14)=2.39$, $p=0.06$ (Bonferroni-corrected p-values). In sum, these analyses show that participants were able to exploit the structure of the task and use their confidence to improve their performance in the second decision. Subsequent analyses focused on the dual-decision condition and were conducted on the aggregated data; importantly the size of the effect of decision order in the dual-decision condition did not differ across the two datasets: mean difference in odds-ratio 0.07, 95% CI [-0.28, 0.41], $t(27)=0.18$, $p=0.85$, the Bayes factor (null/alternative) was 3.70, indicating moderate support for the null hypothesis.
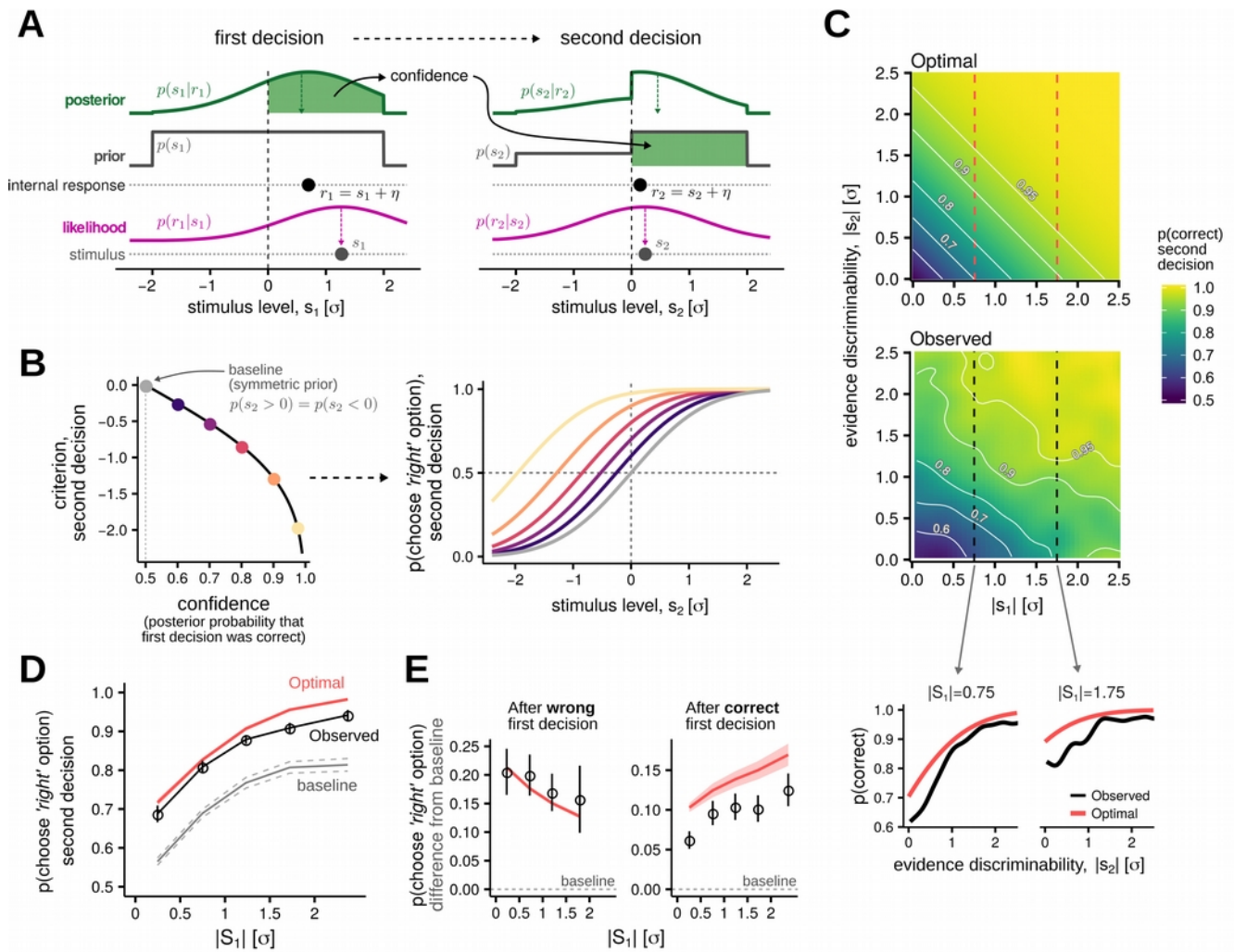
Ideal Bayesian observer

We considered how a Bayesian observer would solve this task (Fig. 2A). The observer's task is to decide whether the stimulus value (e.g., the mean tilt from vertical in the orientation task) is greater or smaller than a reference level, call it zero. However, the observer has access only to a noisy estimate of the stimulus value. To decide whether the stimulus is greater or smaller than zero, the Bayesian observer would integrate the posterior probability of the stimulus over the corresponding semi-interval and choose the most likely alternative. The probability of the chosen alternative also corresponds to their confidence in the decision (the posterior probability that the choice was correct). The confidence is then used to update prior expectations about the second signal: the Bayesian observer assigns a prior probability equal to their confidence to the possibility that the second signal will be positive (e.g. corresponding to a 'right' tilt), and the complement of this probability to the alternative possibility. Updating the prior amounts to shifting the criterion used for the second decision by an amount that

increases with confidence (Fig. 2B; see Supplementary Information for details). Trials with higher confidence in the first decision then result in a larger prior probability for 'positive' stimuli and, hence, a stronger bias for responding 'right' in the second decision (Fig. 2B). This bias can be quantified by examining how choice probabilities in the second decision vary as a function of the discriminability of the first stimulus. Note that, because the signal sign in the second decision depends on the accuracy of the first decision, an increase in the discriminability of the first stimulus will result in an increased rate of responding 'right' to the second stimulus regardless of observer's use of confidence information (Fig. 2D). Thus, we examined differences in choice probabilities relative to baseline probabilities that would be obtained given completely independent responses to the two stimuli. When averaged over all trials, the observed pattern of choices displays a trend similar to that predicted by the Bayesian model (Fig. 2D), which is remarkable considering that this model is fully constrained by the data. To provide a more stringent comparison between human behaviour and the Bayesian model, we next split data according to the accuracy of the first decision: confidence should decrease with greater signal discriminability after an incorrect decision, and increase after a correct one[3]. However, while human participants closely match the ideal after incorrect first decisions (Fig. 2E, left sub-panel), they systematically chose 'right' less frequently than the Bayesian observer after correct first decisions (Fig. 2E, right sub-panel). This pattern can be interpreted as an under-confidence bias and suggests deviations from the Bayesian strategy. Moreover, the finding of an under-confidence bias is inconsistent with sampling-based approximations of the Bayesian strategy, as these would predict an opposite over-confidence pattern (see Supplementary Information).

Human performance is not explained by a fixed bias

One possible explanation for the observed sub-optimality of behaviour is that participants did not monitor their confidence after each decision, but rather adopted a fixed bias towards responding 'right' in the second decision. Indeed, since they performed above chance in the first decision (mean proportion of correct first decision: 0.80), a fixed bias would improve their performance in the second. We estimated the best-fitting fixed bias for each participant and compared the predictions with the observed behaviour. As shown in Fig. 3A, the fixed bias model fails to explain how choice probabilities in the second decision vary with the discriminability of the first stimulus. As a further test, we fit logistic regressions for each participant to measure the influence of the accuracy of the first decision on the probability of choosing 'right' on the second decision. The results of this analysis indicate that, even after controlling for stimulus values, participants were significantly more likely to choose 'right' in the second decision after a correct first decision than after an incorrect one: odds-ratio 2.61, 95% CI [1.84, 3.55], $t(28)=5.71$, $p=3.99\times10^{-6}$, indicating that they monitored their confidence on a trial-by-trial basis rather than used a fixed strategy. Moreover, when adding the discriminability of the first stimulus as a predictor, we found that it modulates the probability of choosing 'right' in the second decision: for each unitary increase in discriminability, the odds-ratio increases multiplicatively by a factor of 1.95, 95% CI [1.42, 3.36], $t(28)=2.95$, $p=0.006$, after a correct response and decreases by a factor of 0.68, 95% CI [0.49, 0.95], $t(28)=2.18$, $p=0.037$, after an incorrect response. Taken together these results allow us to reject the hypothesis that the bias toward responding 'right' in the second decision was constant across trials and instead conclude that participants adjusted their criterion on a trial-by-trial basis

Figure 2. Ideal Bayesian model and comparison with human behaviour. A. Representation of probability density function and belief updating in the Bayesian observer, first and second decision represented from left to right. The physical stimulus evokes an internal estimate or response corrupted by Gaussian noise, $\eta \sim N(0, \sigma^2)$. The ideal observer knows about the variability of its perceptual system and computes a posterior distribution over possible stimulus values. After having made the first decision, the ideal observer integrate the posterior distribution over the chosen semi-axis, to evaluate its confidence, that is the subjective probability that the decision was correct, given the evidence. The confidence is then used to update its prior for the second decision: a prior probability equal to the confidence is assigned to the possibility that the stimulus values will be drawn from the positive (rightward) semi-axis. **B.** It can be shown (see Supplementary Information) that updating the prior amounts to shifting the decision criterion: here we show on the left the optimal criterion shift as a function of the confidence level, and on the right the psychometric functions which describe choice probabilities for several confidence levels, ranging from 0.5 (grey) to 0.975 (yellow). **C.** Optimal and observed performance in the second decision, plotted as a function of both first and second stimuli discriminabilities (the 2D surface showing the observed performance was computed by smoothing the accuracy in individual trials, pooled from all observers, with a Gaussian kernel with standard deviation of 0.2). Although participants were able to use the structure of the task to improve their performance in the second decision (see also Fig.1B), their performance did not reach the ideal benchmark provided by the ideal Bayesian observer, as shown also by the two example cross-section at the bottom. **D.** Choice probability in the second decision was influenced by the discriminability of the first stimulus. This pattern is better appreciated when data are plotted as a difference from the baseline, that is a psychometric function that describes what would be the choice probabilities in the absence of prior information (corresponding to a confidence level of 0.5, illustrated with the grey dot and lines in panel B). All error bars indicate bootstrapped standard errors.

A biased-Bayesian model does not fully account for sub-optimal choice behaviour

An alternative explanation for the under-confidence pattern (Fig. 2D) is that participants might systematically overestimate their sensory noise (see Supplementary Information for details). To test for this possibility we computed for each participant how much the estimate of the noise should have deviated from the true value to explain the under-confidence pattern. We found that on average participants should have over-estimated their sensory noise by a factor of 2.94, 95% CI [1.99, 4.03]. This biased Bayesian model can account well for the under-confidence following correct first decisions (and therefore improve the overall fit compared to the unbiased Bayesian model, since correct first decisions were more frequent). However, over-estimating the internal noise reduce the bias toward responding 'right' both after correct and incorrect first decisions, thus degrading the model's ability to predict performance after incorrect decisions (Fig. 3B). This indicates that participants not only were biased toward under-confidence (relative to the ideal) but also that confidence bias varied depending on the discriminability of the stimulus.

Non Bayesian models with discrete confidence levels provide the best description of human behaviour

The systematic deviations from Bayesian confidence exposed above could result from observers that are under-confident at large and small discriminability, but over-confident at intermediate discriminability. In other words, the function linking evidence to confidence must have a slope much larger than the one of the Bayesian observer in a small range of discriminabilities (going from small to intermediate). The simplest function with these characteristics is a step function (Fig. 3C). We note that a step function would result naturally if observers make decisions using a point estimate of the stimulus. In this case, confidence assessments must be obtained by comparing the estimate to fixed, possibly learned, criteria. If the observer has only one confidence criterion, then the resulting evidence-confidence function would have two confidence levels (low and high). In this case the observer would shift their decision criterion to increase the probability of responding 'right' only when in the high-confidence state (see Supplementary Information for details). We estimated the parameters of the discrete-confidence model and compared it against the alternative models described above. We compared the goodness of fit of models using the Akaike Information Criterion[11] (AIC), which takes into account model complexity. An analysis of simulated data confirmed that the models were distinguishable and correctly recovered (see Extended Data Figure 3). The smallest AIC, indicating better fit, was attained by the discrete-confidence model. The AIC difference between the discrete-confidence model and other models (summed over subjects) revealed little or no support[35] for the remaining models: the difference was 151 for the biased-Bayesian model, 95% CI [67, 253]; 343 for the fixed-bias model, 95% CI [163, 597]; and 1591 for the ideal Bayesian model, 95% CI [1020, 2240]. Additionally, we compared the models using a random-effects approach[12,13], which assumes that different participants could have used different strategies, corresponding to different models, to generate their behavior (see Supplementary Information for details). The protected exceedance probability of the discrete model (the second-order probability that participants were more likely to use this model to generate behavior) was >0.99 and the Bayesian omnibus risk (the posterior probability that all models have the same frequencies) was 1.70×10^{-4}. The discrete model thus provides a better fit than the alternatives and can explain better the pattern of under-confidence following correct first decisions combined with the relatively high confidence following incorrect decisions (Fig. 3C). These high-confidence errors would occur whenever the internal signal not only has the opposite sign to the

physical stimulus (due to random noise fluctuations), but is also large enough to exceed the confidence criterion.
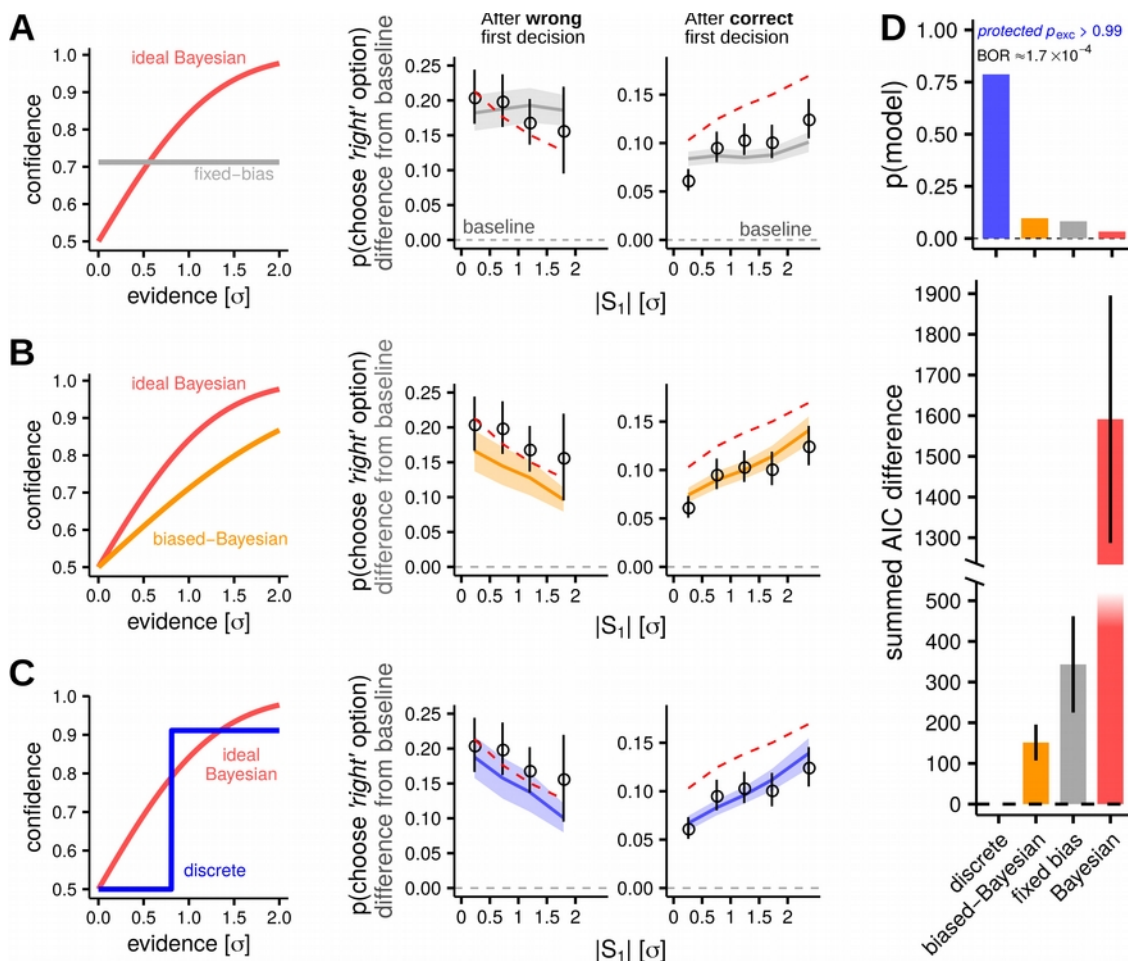
Figure 3. Alternative sub-optimal models. A. Fixed-bias model. **B.** Biased-Bayesian model. **C.** Discrete confidence model. In both A, B and C the left panel shows the functions linking confidence to evidence in each of the models (for comparison each also panel includes the function of the ideal Bayesian observer); these functions were generated using the median of the estimated parameters across observers. Right panels show data and model predictions. Note in particular that the biased-Bayesian model, whilst providing an overall good fit to the data, tends to under-estimate the probability of choosing 'right' after a wrong decision more than the discrete model. This pattern is evident also in individual data (see Extended Data Figure 1). **D.** Model comparison: the top panel represents model frequencies estimates using a random-effects approach that allows heterogeneity in the models that participants used to generate behaviour[12]; the annotation reports the (protected) exceedance probability of the winning discrete-confidence model, together with the Bayesian omnibus risk (BOR), i.e. the probability that observed differences in model frequencies are due to chance. The bottom panel represent differences in the Akaike Information Criterion, relative to the winning model, summed over subjects (see Extended Data Figure 2 for a plot of individual AIC values). All error bars and bands indicate bootstrapped standard errors.

Discussion

We presented an approach to study confidence in perceptual decisions objectively. Our protocol involves a sequential decision task in which participants may improve their performance in a second decision by using their confidence in the first. This experimental protocol not only allows for investigation of metacognitive monitoring, but also of metacognitive control of behaviour[14]. It can therefore be regarded as a laboratory proxy of real-life situations where current decisions might depend on the pending outcomes of previous choices. One advantage of our experimental protocol is that it permits quantitative testing of hypotheses about the human sense of confidence without requiring arbitrary assumptions, for example about how participants map their subjective confidence onto an external rating scale in the case of self-report measures. Additionally, the current approach has the advantage that the decision strategy depends only on internal sources of uncertainty, as opposed to external ones, such as stochasticities in the environment, that typically elicit conservative and sub-optimal placements of decision criteria[15].

Taken together our results revealed that participants performing the dual-decision task can meaningfully use confidence in a prior choice to inform subsequent decisions. However participants also exhibited specific and systematic deviations from the predictions of the ideal Bayesian model. Our analysis revealed that the relationship between the internal signals (the sensory evidence) and the confidence does not have the shape that would be expected if subjective confidence were computed as if by a Bayesian observer, by integrating a posterior probability distribution over the values of the stimulus that are consistent with the choice made. The pattern found in our data could instead be accounted for by an alternative, non-Bayesian model with only two discrete levels of confidence. This model posits that confidence is discretised in few distinct levels, as also suggested by previous work[16]. More specifically, in our discrete model there would be two distinct confidence states: confident - i.e more likely that the response was correct - vs. non-confident - i.e., equally likely that the response was correct or incorrect (see Fig. 3C). This one confidence criterion variant of our model provided the best and most parsimonious description of the empirical data.

In the discrete model, sensory evidence is transformed to confidence via a discrete step function. This suggests that human observers do not assign numerically precise subjective probabilities to perceptual interpretations and implies that confidence in a perceptual decision may be based on only a single point estimate of the most likely stimulus, rather than on a full probability distribution over all possible stimuli[5]. Whilst this algorithmic description of the mechanisms underlying subjective confidence valuation may seem simplistic, it could be the natural consequence of a limit in how much information can be accessed and used for meta-perceptual judgments. In fact the discrete model effectively compresses (in the information-theoretical sense) the amount of information that is encoded from the first decision and remembered for later use in the second decision (it compresses this information to only 1 bit). One intriguing possibility therefore, is that the critical constraint preventing subjects from accurately assigning probabilities and achieving optimality is a capacity limit in the information that is encoded for later use. This constraint could be formulated as a maximum limit in the mutual information between the internal evidence used in the decision and the internal variable representing the degree of confidence. This idea fits well with "second order" models of meta-cognition, which postulate a separation (in space and/or time) of the internal states supporting decision and confidence[17] and assume that the samples of sensory evidence guiding decisions and confidence are distinct, although coupled. Our results suggest an extension of this framework where the nature of this

coupling emerges from the loss of information during confidence valuation. According to this idea the internal states supporting decisions and confidence could be viewed as the input and output of an information channel with limited capacity (i.e. with an upper bound on the rate at which information can be transmitted reliably). Our results thus provide additional support for the idea that any rational account of confidence needs to incorporate implementational constraints and resource limitations[18] to be descriptively accurate.

It is interesting to note that this pattern, in which subjective probabilities both overestimate and underestimate the objective probability in distinct sub-domains of the same task, resembles the distortions of subjective probability that have been found also in non-perceptual tasks, such as those involving decisions from experience[19], suggesting that there may be common limitations in the internal representation of subjective probabilities across different tasks. Additionally, the finding that confidence can be used to improve performance in our dual-decision task across different perceptual tasks (here motion and orientation) is, in principle, consistent with previous claims that confidence may be represented on a common scale and act as a 'common currency' between different perceptual judgments[20,21]. Indeed, in an additional 'within-task' experiment that involved the same type of perceptual judgment in both decisions, we found a pattern of deviation from optimality that was qualitatively similar to that observed in our main experiment. However, quantitative modelling of the within-task experiment failed to provide decisive evidence in favor of either the discrete or biased-Bayesian models, leaving open the possibility that participants might be able to make slightly more efficient use of information when both decisions involve the same perceptual judgment (see Supplementary Results for details and a discussion of caveats in the interpretation of within-task experiments).

Finally, our results have implications for the current investigation of the neural substrate of confidence and metacognition, because many of these studies assume, more or less explicitly, that subjective confidence corresponds to the Bayesian posterior probability[22,23]. However, the evidence supporting the hypothesis that Bayesian inference is algorithmically implemented in the brain is debated[5,24,25] leaving open the possibility that the algorithms implemented by the brain may involve heuristic shortcuts[25]. Distinguishing these possibilities requires novel experimental and theoretical approaches to test the validity of Bayesian inference as a *process* model[26] for the human sense of confidence. In the present study we have not only found further evidence that brain processes that subtend confidence judgments do not conform to the ideal benchmark represented by the Bayesian models, but also provided a new experimental protocol that we believe will be a valuable tool for future investigations of confidence and metacognition and for translating advances in this field into concrete, theory-driven applications[27].

Methods

Participants

All experiments were performed in accordance with the requirements of the Declaration of Helsinki. The experimental protocols were approved by the Research Ethics Committee of the School of Health Sciences of City, University of London and by the UCL Research Ethics Committee. A total of 31 subjects participated in the experiments (mean age 31, standard deviation 11; 12 females). The first 16 participants (dataset 1) ran the experiment at City University (London, UK). The remaining 15 participants (datset 2) were testedt in a replication of the experiment ran at the University College London, Division of Psychology and Language Science (London, UK), using the same procedure and an equivalent experimental setup. Two participants displayed a very large response bias in one of the two tasks (see Analysis) and were excluded from the analysis. No statistical methods were used to pre-determine sample size; sample size was chosen to be similar to those reported in previous studies[28,29]. All participants had normal or corrected-to-normal vision and gave their informed consent in written form prior to participation. All participants were naïve to the specific purpose of the experiment. Naïve participants were compensated with £10 pounds for each hour of experiment.

Behavioural tasks

Behavioural protocols required participants to make a sequence of two perceptual decisions each trial. There was no maximum time limit to give the response and in all tasks participants were explicitly asked to prioritize accuracy over speed (nevertheless response times covaried with confidence, see Supplementary Results, Analysis of response times). Experimental protocols were implemented using MATLAB (The MathWorks Inc., Natick, Massachusetts, USA) and the Psychophysics Toolbox[30,31].

Participants were asked in each decision to either discriminate the mean orientation (left or right from vertical) of an array of lines, or the direction of motion (left or right) of dynamic, random-dot motion stimulus. All trials always included both discrimination tasks (e.g., first orientation then motion, as in Fig. 1, or vice-versa) but the order differed across sessions. The lines were 150 black and white segments (0.7 degrees of visual angle or dva long, anti-aliased), presented on a grey background and randomly placed using a uniform distribution over a circular area (diameter 8 dva) placed at the center of the screen. The orientation of the lines were sampled from a Gaussian distribution with varying mean and standard deviation fixed at 15 degrees. The random-dot stimulus was made of 200 black and white dots (diameter 0.1 dva), presented within the same circular area, moving at a speed of 4 dva/sec, with a lifetime of 4 refresh cycles (≈67 msec). Both lines and dots were presented for 300 msec. A number (1 or 2) was present on top of the display to remind the participants whether the current decision was the first or the second of each trial, together with two dots on the left and right side of the stimulus, which were both white for the first decision but red and green for the second, indicating the expected alternative in case of incorrect and correct first decision, respectively. Stimuli were presented on a LCD monitor, 51.5cm wide, placed at 77cm of viewing distance (the distance was kept constant by means of a chin rest), in a quiet, dimly lit room. The monitor resolution was 1920×1200.

The experiment was split into three sessions run on separate days; participants were compensated with 10£ for each hour of experiment. In the first session, participants completed 300 trials of the control task in which the first and second decisions were independent. Each trial comprised two perceptual decisions. In one block (150 trials) the first decision was about the orientation of the lines

while the second was about motion direction. This order was reversed in the second block. An auditory beep after the second decision indicated whether subjects responded incorrectly in either of the two decisions.

In each of the following two sessions participants completed 300 trials of the dual-decision task, where the correctness of the first response was predictive of the direction of the second stimulus. The order of the two discrimination tasks varied across sessions with order counter-balanced across participants. Before each session of the dual-decision task, participants were reminded of the conditional rule and ran a short practice session. The practice continued until participants confirmed they understood the structure of the experiment and were comfortable to continue to the experiment. In the dual-decision trials, an auditory beep after each trial indicated if participants responded correctly to the second decision. Stimulus discriminability (mean tilt of the lines or motion coherence) were sampled randomly from a uniform distribution, spanning from 0 to a maximum tilt of 12 degrees or 80% coherence in the initial control task. The maximum values in the subsequent two sessions were set to 2 standard deviations of internal noise (σ, as estimated from control trials; see Analysis for details).

Analysis

Choices in the second decision (coded either as left vs right or correct vs incorrect) were analyzed by means of logistic regressions fit at the individual level (for details of logistic regression models see the Supplementary Information: Model-agnostic analyses). Effects were tested at the group level either by computing bootstrapped confidence interval on the estimated odds-ratios, or by means of t-tests on the coefficients of the logistic regression (which corresponds to the logarithm of the estimated odds-ratios and whose distribution could be assumed normal, as indicated by Shapiro-Wilk tests). Bootstrapped confidence intervals were computed using the bias-corrected and accelerated method[32]. All tests were two-tailed and were run with an alpha level of 0.05. In the case of non significant t-tests, evidence for the null hypothesis was estimated by means of Bayes factors (null/alternative), computed using the JZS prior[33] (which consists of a Cauchy prior on the standardized effect size and uninformative Jeffreys prior on variances). All analyses were performed in the open-source software R[34].

Estimation of internal noise. For each participant we estimated the standard deviation of the internal noise (σ) by considering only first decisions and control trials, in which the prior probability of the stimulus was identical between left and right alternatives. To account for the possibility that participants might have made attention lapses (e.g. stimulus-independent errors) or be biased, we fit 4 different psychometric functions, ranging from the simplest function with only one free parameter (that is the noise parameter, σ), to a most complex function that had 3 free parameters, including a lapse rate parameter and a bias term (a non-zero location parameter). In all the 4 psychometric functions the parametric form was a cumulative Gaussian (see Supplementary Information for details), so each function included the same noise parameter (σ). Each psychometric function provided a slightly different estimate of the internal noise σ. For analyses, we combined the 4 estimates by weighting them according to the Akaike weight[35] of each psychometric function. This results in a more robust estimate of the noise that is not conditioned on any of the specific psychometric functions considered but instead permits all scenarios whilst reducing the chance of overfitting. The model-averaged estimates were used to transform the values of the stimulus from raw units (degrees or coherence) to units of internal noise (σ). Two participants were excluded from subsequent analyses because their first decisions in one

of the two task displayed a large systematic bias (more than 2 standard deviations of the average bias of all participants).

Computational modelling.

To model performance in our task we considered that observers can have access only to a noisy estimates of the stimulus levels, corrupted by Gaussian noise (i.e. the internal response as defined in signal detection theory). We then developed alternative computational models implementing different strategies and compared their ability to account for performance in the second decision. We fit models to individual data by numerically finding the values of the free parameters that maximized the (log) likelihood of choices in the second decision. Only second decisions were used in the model comparison since all the models make similar prediction for the first decision, but differ in how they use information from the first choice to inform the second decision. We performed two types of model comparisons, both based on the Akaike Information Criterion[11] (AIC) as a measure of fit quality. First we compared the AIC values of each model, summed across subjects. However, this analysis assumes that all participants used the same strategy (corresponding to a given model) to generate their behavior. Therefore we also performed a more conservative analysis, which takes into account the possibility that different subjects may have used different models to generate their behavior and aims at inferring the probability of each model[12] (see Supplementary Information for details). Additionally, to ensure that the models are distinguishable, we generated synthetic data and verified that the generative models could be correctly recovered based on their AIC values (see Extended Data Figure 3).

In the following we summarise the main features of the computational models tested; the mathematical details are provided in the Supplementary Information.

Ideal Bayesian observer. The ideal Bayesian observer has full knowledge of the statistics of the internal noise and can use this knowledge to accurately compute the posterior probability of having responded correctly (i.e. its confidence). The confidence is then used to update prior expectations for the second stimulus; in the Supplementary Information we show that updating prior expectations amounts to shifting the decision criterion and we provide a derivation of the optimal shift in decision criterion, given the internal response and the level of internal noise. The only parameter of the ideal model is the internal noise standard deviation (σ), however since this is estimated only from first decisions, this model can be regarded as having no free parameters.

Bayesian sampler. This class of models represents a sampling-based approximation of the ideal Bayesian model. According to this model the observer does not have access to the full probability distribution of their internal response, but can approximate it based on a finite number of samples. The performance of this model will approach the ideal Bayesian model as the number of samples grow to infinity, however it is expected to display systematic biases for a small number of samples. In the Supplementary Information we show that these biases are characterised by over-confidence, that is in the opposite direction with respect to what we found for most of our subjects. These sampling-based approximations of the ideal Bayesian model are thus inconsistent with our behavioural results.

Biased-Bayesian. This model performs a Bayesian computation, but is sub-optimal because it uses an incorrect estimates of its own internal noise. Specifically, while the variability of its internal response is described by a Gaussian distribution with standard deviation σ, to assess its confidence the biased-Bayesian observer uses $m\sigma$ (with $m>0$). The only free parameter is the factor m characterising the extent by which the internal noise is mis-estimated. If $m=1$ the biased-Bayesian model is equivalent to the ideal Bayesian, whereas it will display systematic global biases for values different from 1.

Values of $m>1$ indicate over-estimation of the internal variability, and will lead to global under-confidence biases, whereas values $0<m<1$ indicate under-estimation of internal variability and will lead to over-confidence.

Discrete confidence. This model does not assume any knowledge about the nature of the internal stochastic processes linking the stimulus to the internal response. Therefore the only information available to monitor the uncertainty of the decision is the distance of the internal response (which can be considered a point-estimate) to the decision criterion; specifically, it assess confidence by comparing the internal response to an additional confidence criterion. This model's confidence can therefore be described as a set of discrete 'states': a 'confident' state that is obtained when the internal response exceed the confidence criterion; and a complementary 'uncertain' state. Only when in the 'confident' state will it shift the criterion used for the second decision. It has two free parameters: one first parameter determining the location of the confidence criterion, and a second parameter indicating by how much it will shift the decision criterion (for the second decision) when 'confident' about the first.

Fixed-bias. This is the model of an observer that does not monitor uncertainty on a trial-by-trial basis, and instead adopts a simple fixed-bias that increases the probability of choosing the 'right' option in the second decision. It has one free parameter, which corresponds to the decision criterion used in the second decision.

Data availability

The data that support the findings of this study are available at https://osf.io/w74cn/.

Code availability

The code for models and analyses that support the findings of this study is available at https://osf.io/w74cn/.

References

1. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
2. Mamassian, P. Visual Confidence. *Annual Review of Vision Science* **2**, annurev-vision-111815-114630 (2016).
3. Sanders, J. I., Hangya, B. & Kepecs, A. Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron* **90**, 499–506 (2016).
4. Adler, W. & Ma, W. J. Limitations of proposed signatures of Bayesian confidence. *bioRxiv* (2018). doi:10.1101/218222
5. Rahnev, D. & Denison, R. N. Suboptimality in Perceptual Decision Making. *Behavioral and Brain Sciences* **41**, 1-66 (2018).
6. De Finetti, B. La prévision: ses lois logiques, ses sources subjectives. in *Annales de l'institut Henri Poincaré* **7**, 1–68 (1937).
7. Savage, L. J. Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association* **66**, 783–801 (1971).

8. Drugowitsch, J., Moreno-Bote, R. R. & Pouget, A. Relation between Belief and Performance in Perceptual Decision Making. *PLoS ONE* **9**, e96511 (2014).

9. Fleming, S. M. & Lau, H. C. How to measure metacognition. *Frontiers in human neuroscience* **8**, 1–9 (2014).

10. Barthelmé, S. & Mamassian, P. Evaluation of Objective Uncertainty in the Visual System. *PLoS Computational Biology* **5**, e1000504 (2009).

11. Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723 (1974).

12. 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).

13. Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group studies - Revisited. *NeuroImage* **84**, 971–985 (2014).

14. Nelson, T. O. Metamemory: A Theoretical Framework and New Findings. in *Psychology of Learning and Motivation* **26**, 125–173 (Elsevier, 1990).

15. Ulehla, Z. J. Optimality of perceptual decision criteria. *Journal of Experimental Psychology* **71**, 564–569 (1966).

16. Zhang, H., Daw, N. D. & Maloney, L. T. Human representation of visuo-motor uncertainty as mixtures of orthogonal basis distributions. *Nature Neuroscience* **18**, 1152–1158 (2015).

17. Fleming, S. M. & Daw, N. D. Self-evaluation of decision performance: A general Bayesian framework for metacognitive computation. *Psychological review* **124**, 1–59 (2016).

18. Adler, W. T. & Ma, W. J. Comparing Bayesian and non-Bayesian accounts of human confidence reports. *PLOS Computational Biology* **14**, e1006572 (2018).

19. Zhang, H. & Maloney, L. T. Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience* **6**, 1–14 (2012).

20. De Gardelle, V., Le Corre, F. & Mamassian, P. Confidence as a common currency between vision and audition. *PLoS ONE* **11**, (2016).

21. de Gardelle, V. & Mamassian, P. Does Confidence Use a Common Currency Across Two Visual Tasks? *Psychological Science* **25**, 1286–1288 (2014).

22. Meyniel, F. & Dehaene, S. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences* **114**, E3859–E3868 (2017).

23. Bang, D. & Fleming, S. M. Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences* **115**, 6082-6087 (2018).

24. Laquitaine, S. & Gardner, J. L. A Switching Observer for Human Perceptual Estimation. *Neuron* **97**, 462-474.e6 (2018).

25. Gardner, J. L. Optimality and heuristics in perceptual neuroscience. *Nature Neuroscience* **22**, 514-523 (2019).

26. Maloney, L. T. & Mamassian, P. Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. *Visual Neuroscience* **26**, 147 (2009).

27. Huys, Q. J. M., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience* **19**, 404–413 (2016).

28. Meyniel, F., Schlunegger, D. & Dehaene, S. The Sense of Confidence during Probabilistic Learning: A Normative Account. *PLoS Computational Biology* **11**, 1–25 (2015).

29. Aitchison, L., Bang, D., Bahrami, B. & Latham, P. E. Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. *PLoS Computational Biology* **11**, 1–23 (2015).

30. Kleiner, M., Brainard, D. & Pelli, D. What's new in psychtoolbox-3. Perception 36 ECVP Abstract Supplement (2007).

31. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial vision* **10**, 437–442 (1997).

32. Efron, B. Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association,* **82**, 171-185 (1987).

33. Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., & Iverson, G. Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review* **16**, 225-237 (2009).

34. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2020).

35. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. (Springer New York, 2002).

Author contributions

M.L., A.G. and G.Mo. conceived the study. M.L., and G.Mo. developed the computational models. M.L. programmed the experiments. M.L. and G.Mi. collected the data. M.L analysed the data. M.L, G.Mo. and T.D. interpreted the results. M.L. wrote the paper with feedback from all authors.

Competing interest

The authors declare no competing interests.

Supplementary Information

for ***"Discrete confidence levels revealed by sequential decisions"***, by

Matteo Lisi[1,*], Gianluigi Mongillo[2,3], Georgia Milne[4], Tessa Dekker[4,5] & Andrei Gorea[6].

[1]Department of Psychology, University of Essex, Colchester, UK.
[2]Centre National de la Recherche Scientifique, Paris, France.
[3]Institut de la Vision, Sorbonne Universite, Paris, France.
[4]Institute of Opthalmology, University College London, London, UK.
[5]Division of Psychology and Language Sciences, University College London, UK.
[6]Université Paris Descartes, Sorbonne Paris Cité, Paris, France.

[*]m.lisi@essex.ac.uk

Contents

1 Supplementary methods

1.1 Computational models

1.1.1 Ideal Bayesian observer

To model performance in the dual-decision task, we assumed that the observer has only access to r_i, a corrupted version of the true stimulus value s_i, defined as $r_i = s_i + \eta_i$, where the subscript $i = 1, 2$ indicate the decision at hand (first or second) and η_i is Gaussian noise with variance σ_i^2. Note that if the two decisions involve different perceptual judgments, the variance of the noise can be different depending on the decision (first or second). The perceptual discrimination task amounts to deciding whether s_i was above or below a reference level, call it zero. The ideal Bayesian observer has full knowledge of the statistics of the internal noise, and to make a decision computes the posterior probability that the variable s_i was positive or negative (i.e. above or below the reference level), given the noisy internal response r_i. Since the internal noise is Gaussian, the likelihood function giving the probability of observing r_i given s_i is a normal distribution with mean s_i and variance σ_i^2.

In the first decision, the prior probability of s_1 is uniform within the range $(-R, R)$, that is $p(s_1) = 1/(2R)$ if $-R \leq s \leq R$, and 0 otherwise. Combining the prior and the likelihood function, the unconditioned probability of observing r_1 can be expressed as:

$$
\begin{aligned}
p(r_1) &= \int_{-R}^{R} p(r_1|s_1)\, p(s_1)\, ds_1 \\
&= \frac{1}{4R}\left[\mathrm{erf}\left(\frac{R + r_1}{\sigma_1\sqrt{2}}\right) + \mathrm{erf}\left(\frac{R - r_1}{\sigma_1\sqrt{2}}\right)\right].
\end{aligned}
$$

The posterior probability of s_1 after having observed r_1 is obtained applying Bayes rule:

$$
p(s_1 \mid r_1) = \frac{p(r_1 \mid s_1)\, p(s_1)}{p(r_1)}.
$$

Finally, the decision variable $c_1^+(r_1)$, corresponding to the probability that the s_1 was greater than 0 is given by

$$
c_1^+(r_1) = \int_0^R p(s_1|r_1)\, ds_1 = \frac{\mathrm{erf}\left(\frac{r_1}{\sigma_1\sqrt{2}}\right) + \mathrm{erf}\left(\frac{R - r_1}{\sigma_1\sqrt{2}}\right)}{\mathrm{erf}\left(\frac{R + r_1}{\sigma_1\sqrt{2}}\right) + \mathrm{erf}\left(\frac{R - r_1}{\sigma_1\sqrt{2}}\right)}. \tag{1}
$$

When $c_1^+(r_1) \geq 1/2$ the observer chooses 'right' (i.e. they decide to report that s_1 was positive, an outcome hereafter indicated with the notation d_1^+) otherwise they choose 'left' (i.e. they report that s_1 was negative, hereafter notated as d_1^-). The posterior probability of being correct in the first decision, which corresponds to the confidence of the ideal observer, is given by $c_1 = \max\left[c_1^+(r_1),\, 1 - c_1^+(r_1)\right]$. The ideal observer uses the probability c_1 to adjust prior expectations for the second decision, specifically by assigning a prior probability equal to c_1 to the possibility that s_2 will be drawn from the positive interval:

$$
p(s_2) = \begin{cases} \frac{c_1}{R}, & \text{if } 0 < s_2 \leq R \\ \frac{1 - c_1}{R}, & \text{if } -R \leq s_2 < 0 \\ 0, & \text{otherwise.} \end{cases}
$$

By applying the same calculation as above with the updated prior $p(s_2)$, one obtains the decision variable $c_2^+(r_2)$ for the second decision:

$$c_2^+(r_2) = \frac{c_1\left[\mathrm{erf}\left(\frac{r_2}{\sigma_2\sqrt{2}}\right) + \mathrm{erf}\left(\frac{R-r_2}{\sigma_2\sqrt{2}}\right)\right]}{(1-c_1)\,\mathrm{erf}\left(\frac{R+r_2}{\sigma_2\sqrt{2}}\right) + c_1\mathrm{erf}\left(\frac{R-r_2}{\sigma_2\sqrt{2}}\right) + (2c_1-1)\,\mathrm{erf}\left(\frac{r_2}{\sigma_2\sqrt{2}}\right)}. \tag{2}$$

This equation reduces to the one for the first decision variable (apart from the difference in the variance of the noise) when $c_1 = 1/2$, as it should.

The range R on which s_1 and s_2 takes values is immaterial. It is possible to simplify Eqs (1) and (2) by taking the limit $R \to \infty$. In this limit one obtains:

$$c_1^+(r_1) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{r_1}{\sigma_1\sqrt{2}}\right)\right] \tag{3}$$

$$c_2^+(r_2) = \frac{c_1\left[1 + \mathrm{erf}\left(\frac{r_2}{\sigma_2\sqrt{2}}\right)\right]}{1 + (2c_1-1)\,\mathrm{erf}\left(\frac{r_2}{\sigma_2\sqrt{2}}\right)}.$$

The decision rules described by these equations amount to comparing the internal signal r_i to a criterion, and decide accordingly (i.e., if $r_i \geq \theta_2$, choose d_2^+, otherwise choose d_2^-). The criterion for the first decision, θ_1, can be expressed as:

$$c_1^+(\theta_1) = \frac{1}{2} \quad \Rightarrow \quad \mathrm{erf}\left(\frac{\theta_1}{\sigma_1\sqrt{2}}\right) = 0$$

which is satisfied for $\theta_1 = 0$. Similarly, the criterion for the second decision, θ_2, can be expressed as:

$$c_2^+(\theta_2) = \frac{1}{2} \quad \Rightarrow \quad \mathrm{erf}\left(\frac{\theta_2}{\sigma_2\sqrt{2}}\right) = 1 - 2c_1 \tag{4}$$

which indicates that θ_2 is a function of c_1 (see Main text, Figure 2B). Using Eqs (1) and (4), we can express θ_2 as a function of the internal response, r_1, as:

$$\theta_2 = -\frac{\sigma_2}{\sigma_1}|r_1|$$

Note that, if the noise is the same (i.e., $\sigma_1 = \sigma_2$), the optimal strategy amounts to shifting the criterion by the same amount as the internal response, that is $\theta_2 = -|r_1|$. Thus, when both decisions involve the same perceptual judgment, the optimal strategy does not necessarily imply the knowledge of the variability of the internal response. On the contrary, the optimal strategy does indeed require the knowledge of σ_1 and σ_2 (at least their ratio) when the decisions are made in different domains, as, e.g., in our experiment involving motion and orientation judgments (see Main text, Figure 1A).

The likelihood of the Bayesian observer choosing d_2^+ (after having chosen d_1^+ in the first decision) can be expressed as:

$$p\left(d_2^+ \mid s_2, d_1^+, s_1\right) = \int_0^{+\infty} p\left(d_2^+ \mid \theta_2 = -\sigma_2 r_1/\sigma_1, s_2\right)\, p\left(r_1 \mid d_1^+, s_1\right) dr_1$$

$$= \frac{1}{2} + \frac{1}{2p\left(d_1^+ \mid s_1\right)} \cdot \frac{1}{\sqrt{2\pi\sigma_1^2}} \int_0^{+\infty} \mathrm{erf}\left[\frac{1}{\sqrt{2}}\left(\frac{r_1}{\sigma_1} + \frac{s_2}{\sigma_2}\right)\right] \exp\left[-\left(\frac{r_1 - s_1}{\sigma_1\sqrt{2}}\right)^2\right] dr_1. \tag{5}$$

3

If instead the observer has chosen d_1^- in the first decision the likelihood for the second decision takes the form:

$$p\left(d_2^+|s_2, d_1^-, s_1\right) = \int_{-\infty}^{0} p\left(d_2^+ \mid \theta_2 = \sigma_2 r_1/\sigma_1, s_2\right) \, p\left(r_1 \mid d_1^-, s_1\right) dr_1$$

$$= \frac{1}{2} - \frac{1}{2p\left(d_1^- \mid s_1\right)} \cdot \frac{1}{\sqrt{2\pi\sigma_1^2}} \int_{-\infty}^{0} \mathrm{erf}\left[\frac{1}{\sqrt{2}}\left(\frac{r_1}{\sigma_1} - \frac{s_2}{\sigma_2}\right)\right] \exp\left[-\left(\frac{r_1 - s_1}{\sigma_1\sqrt{2}}\right)^2\right] dr_1.$$

$$(6)$$

The integrals in the Eqs (5) and (6) were evaluated numerically using the adaptive quadrature algorithm as implemented in the function integrate() in **R** [10].

1.1.2 *Biased*-Bayesian observer

Similarly to the ideal Bayesian observer presented in the previous section, the biased-Bayesian observer also makes decisions based on internal signals corrupted by Gaussian noise. However, contrary to the ideal Bayesian observer, this model does not assume exact knowledge of the values of σ_1 and σ_2. In this case, the adjusted criterion for the second decision is given by

$$\theta_2 = -\frac{1}{m}\frac{\sigma_2 \, |r_1|}{\sigma_1} = \frac{1}{m}\theta_2^{(opt)},$$

where $\theta_2^{(opt)}$ is the criterion of the Bayesian observer. When the ratio of the variances of the internal noises is overestimated ($m > 1$), the criterion is shifted away from zero less than in the optimal Bayesian model. This leads to an under-confidence bias, i.e., in the second decision the observer chooses *as if* the first decision was correct more frequently than expected if their performance was optimal. Conversely, when the ratio of the variances of the internal noises is underestimated ($m < 1$), the criterion is shifted more than in the optimal Bayesian model, leading to an opposite over-confidence bias.

1.1.3 *Discrete* confidence model

As an alternative to the optimal Bayesian model we considered a class of models that do not assume any knowledge about the nature of the internal stochastic process linking the stimulus s_i with the internal observation r_i. These non-Bayesian observers perform similarly to the Bayesian model for the first decision, that is when $r_1 \geq 0$ they choose d_1^+, and d_1^- otherwise. However, they cannot estimate a full probability distribution over the values of s_i, and therefore can assess confidence only by comparing the internal response r_i - which can be described as a point-estimate - to a set of one (or possibly more) fixed criteria. In the case of a single confidence criterion, this non Bayesian observer is in a high-confidence state when the internal signal exceeds the confidence criterion, and uncertain otherwise. When confident about the first response, the observer shifts the criterion for the second decision by a fixed amount (thereby increasing the probability of choosing d_2^+). If only one criterion is used the model has 2 discrete confidence levels (e.g., 'uncertain' vs 'confident'). In such a 'discrete' model, if the confidence criterion is w_1, the probability of the observer being confident about their first decision, after having responded d_1^+, can be expressed as:

$$p\left(\text{confident} \mid s_1, d_1^+\right) = \frac{1 - \Phi\left(\frac{s_1 - w_1}{\sigma_1}\right)}{1 - \Phi\left(\frac{-w_1}{\sigma_1}\right)}$$

4

where Φ is the cumulative distribution function of the standard normal distribution. Note that this probability does not denote the confidence of the observer about their choice, which instead is assumed here to be a discrete binary state. Applying the law of total probability, the probability of the observer reporting d_2^+ in the second decision can be expressed as:

$$p\left(d_2^+ \mid s_2\right) = \ p\left(\text{confident} \mid s_1, d_1^+\right) \Phi\left(\frac{s_2 - \theta_2}{\sigma_2}\right) + \left[1 - p\left(\text{confident} \mid s_1, d_1^+\right)\right] \Phi\left(\frac{s_2}{\sigma_2}\right)$$

where θ_2 is the shift in criterion for the second decision applied by the observer when they are confident in their first decision. w_1 and θ_2 are free parameters that we fit to the data by maximum likelihood estimation. It is straightforward to extend the model in order to have more than two confidence levels, however we find that increasing the number of additional levels in the model provides, in the current dataset, only a negligible increase in likelihood which is not sufficient to justify the additional complexity. The difference in AIC score between the simple model with only 2 confidence levels outlined above and a model with one additional confidence criterion (and therefore 3 total confidence levels) was 102 ± 5 (summed over subjects) favouring the simpler model with only 2 confidence levels.

1.2 Sampling-based approximation of Bayesian observer

An interesting alternative to the models presented in the previous sections is represented by models where the observer does not have access to the full probability distribution of their internal signals, but bases their decision on a limited number of samples. In these models the posterior probability that the choice is correct (the confidence) is approximated on the basis of a fixed number n of samples $x_1, x_2, ..., x_n$ drawn from the posterior distribution $p\left(s_i \mid r_i\right)$. The performance of these sampling-based models will approach the optimal Bayesian model as $n \to \infty$, however they are expected to display systematic biases and deviations from the optimal model for a small number of samples [12]. Here we show that the sampling-based approximation of the Bayesian observer will display a systematic over-confidence bias, that is in the opposite direction to that displayed by most participants in our study, and is thus inconsistent with the behavioural pattern found in our data.

1.2.1 Sampling bias and over-confidence

It has been shown that when a probability p is estimated from a small sample, as the empirical frequency of successes k out of n random trials, $\hat{p} = k/n$, the risk of over-estimation (that is $\hat{p} > p$) or under-estimation, $(\hat{p} < p)$ depends in a complex way on both the probability p and the sample size n [13]. This is however for estimating a single, fixed probability p. What would instead be the average (expected) bias, over many repeated estimations, when the probability p varies randomly within a given range? In our experiments confidence is the posterior probability that a binary choice is correct, and as such it varies from complete uncertainty, $p = 0.5$, to complete certainty, $p = 1$, hence $p \in [0.5, 1]$. We show here that when a set of probabilities $p_1, ..., p_m$ uniformly distributed in the interval $[0.5, 1]$ is estimated using a limited number of samples n, the expected bias is one of over-estimation.

For a given n and p the probabilities of over- and under- estimation can be expressed as:

$$p\left(\hat{p} > p \mid n, p\right) = \sum_{k=\lfloor np \rfloor + 1}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$p\left(\hat{p} < p | n, p\right) = \sum_{k=1}^{\lceil np \rceil + 1} \binom{n}{k} p^k (1-p)^{n-k}$$

where $\lceil . \rceil$ and $\lfloor . \rfloor$ are the ceiling and the floor operators, i.e. functions that map a number to the smallest following integer or the largest previous integer, respectively. Following Shteingart and Loewenstein [13] we consider the difference between these two, denoted as *probability estimation bias*, which takes the form:

$$\Delta p = p\left(\hat{p} > p | n, p\right) - p\left(\hat{p} < p | n, p\right).$$

When Δp is positive, it indicates that the probability p is more likely to be overestimated than underestimated, and vice versa for negative values. Assuming that all values of p in the interval are equally likely, the expected bias can be computed by integrating Δp over the range of p (that is $[0.5, 1]$):

$$\mathbb{E}\left[\Delta p\right] = \frac{1}{2} \int_{0.5}^{1} \Delta p \, dp.$$

A positive value of the expected probability estimation bias (that is $\mathbb{E}\left[\Delta p\right] > 0$) indicates that, on average, the probabilities in this interval are more frequently overestimated rather than underestimated. This integral can be evaluated numerically, and in Supplementary Figure 1A we plotted the expected probability estimation bias as a function of the number of samples, for two different ranges. When p varies within the range of confidence, $[0.5, 1]$, the value of the probability estimation bias is always positive, although modulated by the number of samples, indicating that in the range $[0.5, 1]$ overestimation is more likely than underestimation.

1.2.2 Fixed-n Bayesian sampler

Here we present the details of a model based on a sampling approximation of the Bayesian observer, and confirm by simulation that it will results in an over-confidence bias. We consider a fixed-n policy, where the observer draws a fixed number of samples for each decision. Although alternative decision policies are possible (such as an accumulator policy, where the decision is taken after a minimum number of samples is accumulated in favour of one of the options), these have been shown elsewhere to result in very similar performances as the fixed-n policy [15].

Similarly to the other models, we assume that the observer has only access to r_1, a corrupted version of the stimulus s_1, $r_1 = s_1 + \eta_1$, where η_1 is Gaussian noise with variance σ_1^2. In the first decision the prior is flat and, taking the limit of the stimuli range $R \to \infty$, the posterior distribution $p\left(s_1 | r_1\right)$ results in a Gaussian distribution centred on the internal observation r_1. The probability that a sample from this distribution is above 0 (the criterion for the first decision) can be computed as:

$$p\left(x_n > 0 | r_1\right) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{r_1}{\sigma_1 \sqrt{2}}\right)\right] = \Phi\left(\frac{r_1}{\sigma_1}\right).$$

The conditional probability that the observer chooses d_1^+ given r_1 is obtained by summing the probability of all the set of samples with at least $\lceil \frac{n}{2} \rceil$ samples above 0 and can be expressed as:

$$p\left(d_1^+ \mid r_1\right) = \sum_{k=\lceil \frac{n}{2} \rceil}^{n} \binom{n}{k} \Phi\left(\frac{r_1}{\sigma_1}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma_1}\right)\right]^{n-k}.$$

In other words the observer choses d_1^+ when the majority of samples is above 0. The observer's confidence in his decision, $c_1 = \max\left[c_1^+, 1 - c_1^+\right]$, is given by the proportion of samples in favor of the choice made:

$$c^+ = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{x_i > 0} \tag{7}$$

where $\mathbb{1}_{x_i > 0}$ denotes the indicator function ($\mathbb{1}_{x > 0} = 1$ if $x > 0$, and 0 otherwise). The probability that the observer chooses d_1^+ given s_1 can be calculated by integrating the probability $p\left(d_1^+ \mid r_1\right)$ over all possible values of r_1:

$$p\left(d_1^+ \mid s_1\right) = \sum_{k=\left\lceil \frac{n}{2} \right\rceil}^{n} \binom{n}{k} \int_{-\infty}^{\infty} \Phi\left(\frac{r_1}{\sigma_1}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma_1}\right)\right]^{n-k} p\left(r_1 \mid s_1\right) dr_1. \tag{8}$$

In the second decision the prior probability of the stimulus s_2 is different for stimuli above or below 0. From the point of view of the observer, the prior probability that the stimulus for the second decision s_2 is above 0 corresponds to the confidence c_1 that the first decision was correct and, conversely, $p\left(s_2 < 0\right) = 1 - c_1$. Updating the prior probability for the second decision amounts to a shift in the decision criterion, as demonstrated for the full Bayesian model. The shift in criterion for the second decision θ_2 is a function of the confidence in the sampling model as in the full Bayesian model

$$\theta_2 = \sigma_2 \sqrt{2} \ \mathrm{erf}^{-1}\left(1 - 2c_1\right). \tag{9}$$

However, and differently from the full Bayesian model, the confidence is limited to a finite set of values determined by the number of samples n. The number of possible confidence levels is $\left\lceil \frac{n+1}{2} \right\rceil$, and each of these corresponds to a value of θ_2. If the observer's first decision was d_1^+, then the probability of the confidence level c_j

$$p\left(c_j \mid s_1, d_1^+\right) = \binom{n}{k} \int_{-\infty}^{\infty} \Phi\left(\frac{r_1}{\sigma_1}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma_1}\right)\right]^{n-k} p\left(r_1 | s_1\right) dr_1$$

where $j \in \left(1, ..., \left\lceil \frac{n+1}{2} \right\rceil\right)$ is an index linked to the number of samples above the criterion, k, according to $k = \left\lfloor \frac{n-1}{2} \right\rfloor + j$. Conversely, if the observer's first decision was instead d_1^-, then the probability of the confidence level c_j is

$$p\left(c_j | s_1, d_1^-\right) = \binom{n}{k} \int_{-\infty}^{\infty} \left\{1 - \Phi\left(\frac{r_1}{\sigma_1}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma_1}\right)\right]^{n-k}\right\} p\left(r_1 | s_1\right) dr_1 \tag{10}$$

and the relation between the index of confidence level j and the number of samples above 0, k, becomes $k = n - \left(\left\lfloor \frac{n-1}{2} \right\rfloor + j\right)$.

A level of confidence c_j would result in a shift in decision criterion θ_j, calculated according to Eq (9). The probability of choosing d_2^+ in the second decision is

$$p\left(d_2^+ \mid s_2, s_1, c_j\right) = \sum_{k=\left\lceil \frac{n}{2} \right\rceil}^{n} \binom{n}{k} \int_{-\infty}^{\infty} \Phi\left(\frac{r_2 - \theta_j}{\sigma_2}\right)^k \left[1 - \Phi\left(\frac{r_2 - \theta_j}{\sigma_2}\right)\right]^{n-k} p\left(r_2 | s_2\right) dr_2.$$

Taking everything together, the probability of the observer choosing d_2^+ given s_2, s_1 and the first decision can be expressed as:

$$p\left(d_2^+ \mid s_2, d_1^+, s_1\right) = \sum_{j=1}^{\left\lceil \frac{n+1}{2} \right\rceil} p\left(c_j | s_1, d_1^+\right) \ p\left(d_2^+ \mid s_2, s_1, c_j\right).$$

1.2.3   Simulation of sampling model

We simulated the model for values of n ranging from 2 to 20. In order to compare the model with the full Bayesian observer, the values of σ for each n are adjusted so as to obtain the same proportion of correct responses in the first decision. For each value of n and for 100 iterations we first estimated the σ for the sampling models based on the set of first decisions made by the participants (this was done using maximum likelihood estimation and equation 8) and simulated second decisions from the sampling model, conditioned on the actual set of observed first decisions made by the participants. The proportion of responses 'right' (d_2^+) in the second decision produced by the sampling models are plotted in Supplementary Figure 1C. As expected, they show a pattern of marked over-confidence: all sampling models tended to respond d_2^+ more often than the optimal model, despite similar accuracy in the first decision. The bias is generally larger for models with smaller number of samples, and decreases approaching the optimal Bayesian model as n increases. Importantly, this bias is incompatible with the observed behavioral data, which instead showed a marked under-confidence bias.
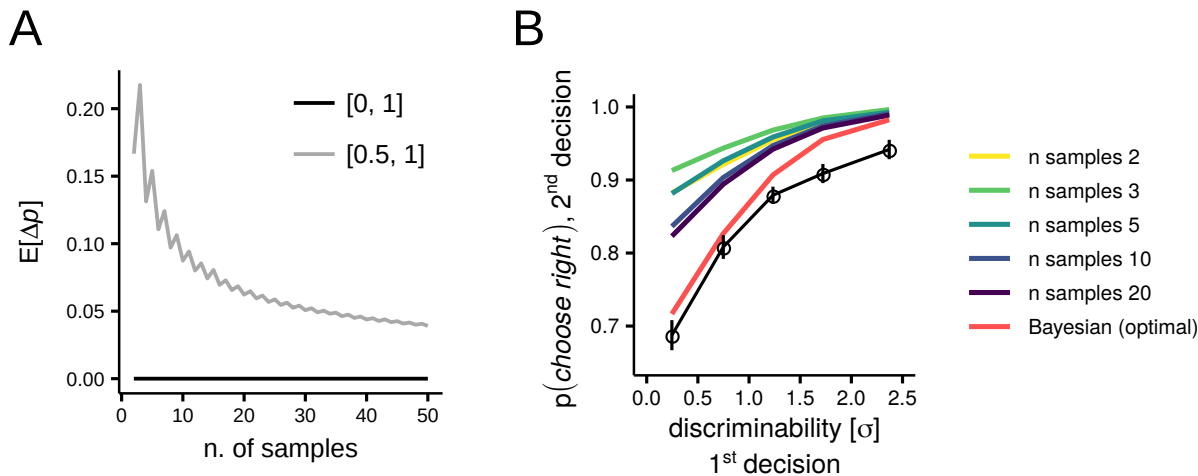
Supplementary Figure 1: Sampling-based approximation of the Bayesian model. A. Expected probability estimation bias Δp plotted as a function of the number of samples when the probability p varies randomly (uniformly) either in the range $[0.5, 1]$, plotted in grey, or in the range $[0, 1]$. It can be seen that when p varies within the range of confidence, from chance to certainty $[0.5, 1]$ the predominant bias is one of over-estimation (because Δp is always positive, grey line). Only when p varis over the whole domain of probability, $[0, 1]$, the expected bias is on average zero and over-estimation and under-estimation are equally likely (black line). **B.** Proportions of responses 'right' (d_2^+) in the second decision as a function of the difficulty of the first decision, as predicted by the full Bayesian model (red line) and the fixed-n Bayesian sampler model. The black dots and lines indicate the choices made by human participants (error bars are bootstrapped standard errors).

1.3 Model estimation and comparison

In our analyses we compared each model's ability to explain performance in the second decision. The first step was to estimate the standard deviation of the internal noise, using only decisions where the prior probability was equal for the two alternatives (that is first decisions and control conditions). In fact in the first decision s_1 is uniformly chosen from an interval centred around 0 (i.e., s_1 will be above 0 with probability $1/2$), and there is no prior information about the sign of s_1. Therefore, for all the models presented in the previous sections (with the exception of the sampling approximation of the ideal Bayesian observer discussed in the previous section), the probability $p\left(d_1^+ \mid s_1\right)$ of the observer choosing d_1^+ can be expressed as:

$$p\left(d_1^+ \mid s_1\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{s_1}{\sigma_1\sqrt{2}}\right)\right] = \Phi\left(\frac{s_1}{\sigma_1}\right).$$

We estimated by maximum likelihood $\hat{\sigma}_1$ for each participant and task. We also estimated other psychometric models, that extended this simple model to account for the possibility that the observer was biased, $p(d_1^+ \mid s_1) = \Phi\left(\frac{s_1-b}{\sigma_1}\right)$ (where b indicates the bias), or made stimulus-independent errors (e.g. attention lapses) with non-zero probability, $p(d_1^+ \mid s_1) = \lambda + (1-2\lambda)\Phi\left(\frac{s_1}{\sigma_1}\right)$ (where λ is the lapse probability); or both $p(d_1^+ \mid s_1) = \lambda + (1-2\lambda)\Phi\left(\frac{s_1-b}{\sigma_1}\right)$. Since we were interested in estimating the internal noise, rather than selecting one of these models, we averaged the estimates of $\hat{\sigma}$ obtained with each of these psychometric models according to their Akaike weights [1, 16]. In more detail, we first calculated the Akaike Information criterion (AIC) of each psychometric models m as

$$\mathrm{AIC}_m = 2\mathcal{V} - 2\ln\hat{\mathcal{L}}_m$$

where $\hat{\mathcal{L}}_m$ is the maximized value of the likelihood function of model m and $\mathcal{V}$ its number of free parameters. The AIC values are first transformed into differences with respect to the best model, $\Delta_m = \mathrm{AIC}_m - \min\mathrm{AIC}$. These differences are then transformed on a likelihood scale and normalized to obtain the Akaike weights w_m, which can be directly interpreted as conditional probabilities for each model [16]

$$w_m = \frac{\exp\left(-\frac{1}{2}\Delta_m\right)}{\sum\limits_{k=1}^{K}\exp\left(-\frac{1}{2}\Delta_k\right)}.$$

Finally, the model-averaged estimate of the noise parameter is computed as the average of each estimated weighted by the Akaike weight of the respective models

$$\bar{\sigma}_1 = \sum_{k=1}^{K} w_k\hat{\sigma}_1^{(k)}.$$

The advantage of this approach is that it avoids the need to choose and commit to a particular form of the psychometric function. Furthermore, because the Akaike weights take into account the complexity of the psychometric functions (i.e. the number of free parameters $\mathcal{V}$), the resulting estimate optimises the balance between model fidelity to the data and the need to avoid overfitting (i.e. of fitting a model that is too complex to be supported by the data). Indeed the model-averaged estimate is more likely to perform better than the alternatives in accounting for new data [1].

In our analysis we used the model-averaged estimate for transforming the stimuli from physical units to units of internal noise. This simplifies the analysis (since first and second decisions involved different perceptual judgements) and guarantees that any differences between models would be due only to differences in explaining performance in the second decision. The same was done for the estimates of the bias term, $\hat{b}$, which was then subtracted from the stimuli value to take into account subjective decision biases in the computation of models' likelihoods. The remaining free parameters for the biased-Bayesian and the non-Bayesian observer were estimated by numerically maximising the (log) likelihood of the data (the observed patterns of second decisions). Maximum likelihood fits were obtained via the BOBYQA algorithm [9], implemented in the optimx package [7, 8] in **R** [10]. The relative quality of models' fit was assessed with the AIC, which we report pooled over subject in the main text; individual differences in AIC values are represented in Extended Data Figure 1.

In addition to the AIC, we compared the models using a random-effects approach [11, 14] which allows for the possibility that different participants could have used different strategies (corresponding to different models) to generate their behaviour. In this approach models are treated as random effects that could differ between subjects and have a fixed (unknown) distribution in the population. The uncertainty around model frequencies is described by a Dirichlet distribution (a multivariate generalisation of the Beta distribution), whose parameters are estimated according to an iterative variational Bayesian procedure, described in [14][1], starting from a measure of the evidence supporting each model (here we used the AIC). Having obtained the optimised parameters of the Dirichlet distribution a useful way to rank the model is according to their exceedance probability, that is the second-order probability that participants were more likely to choose a certain model to generate their behaviour rather than any other alternative model. If $r_1, \ldots, r_K$ are the estimated frequencies of the K models, the exceedance probability (EP) of model k is formally defined as $EP_k = p(r_k > r_j \mid \text{data})$, with $k \neq j$. Since we are considering more than two models, the exceedance probabilities need to be approximated via simulation [14]. We simulated 10^6 multivariate Dirichlet samples $\tilde{r}_1, \ldots, \tilde{r}_K$ and calculated the exceedance probability of model k as

$$EP_k = \frac{\sum \mathbb{1}_{\tilde{r}_k > \tilde{r}_j}}{10^6}, \forall j \in \{1, \ldots, K, j \neq k\}$$

where $\mathbb{1}$ is the indicator function ($\mathbb{1}_{x>0} = 1$ if $x > 0$ and 0 otherwise), summed over the total number of multivariate samples drawn. *Protected* exceedance probabilities (PEP) [11] are a further extension of exceedance probabilities corrected for the possibility that the observed model differences are due to chance. They are calculated as $PEP_k = (1 - \text{Bor}) EP_k + \text{Bor}/K$, where Bor is the Bayesian omnibus risk, that is the posterior probability that all models have the same frequency [11]. See Main Text and Figure 3B for a plot of the estimated model frequencies with the associated exceedance probabilities and Bayesian omnibus risk.

1.4 Model-agnostic analyses

We report here the details of analyses that did not involve fitting the decision models presented in Section 1.1. We analysed performance by means of a series of logistic regressions. For the analysis of response accuracy, we estimated logistic regression of the form

$$\text{logit}\left[p(\text{correct})\right] = \beta_0 + \beta_1 \left|s\right| + \beta_2 \, \text{DecisionOrder} + \beta_3 \, \text{Dual} + \beta_4 \left(\text{Dual} \times \text{DecisionOrder}\right) \quad (11)$$

---

[1]Implemented in an **R** package available at `http://github.com/mattelisi/bmsR`

where logit $[p] = \log \frac{p}{1-p}$, $|s|$ is the absolute value of the stimulus (i.e. the evidence discriminability, expressed in units of internal noise, σ) and DecisionOrder (first vs second) is a dummy variable that was set to 1 in the case of second decisions and 0 otherwise. Dual was also a dummy variable, set to 1 for the dual-decision condition, and 0 for the control condition. In the main text we report 95% CI on the odds-ratio of a correct response in the second decision relative to that in the first decision; for example, in equation (11) this would be calculated as e^{β_2} for the control condition and as $e^{\beta_2 + \beta_4}$ for the dual-decision condition.

We performed further analyses to assess the possibility that the sub-optimality of human behaviour could be explained by a fixed-bias model which does not adjust decision criterion dynamically depending on the stimulus observed at the first decision. We analysed the proportion of 'right' alternatives chosen in the second decision as a function of the stimulus value in the second decision and the accuracy of the first decision

$$\text{logit} \left[p(\text{choose right}) \right] = \beta_0 + \beta_1 s_2 + \beta_2 \, \text{Acc}_1$$

where Acc_1 is a dummy variable set to 1 in the case of correct first decisions. This logistic regression showed that participants were more likely to choose 'right' after correct first decisions (see Main Text). This logistic regression was also extended to include the interaction between accuracy and evidence discriminability in the first decision, that is

$$\text{logit} \left[p(\text{choose right}) \right] = \beta_0 + \beta_1 s_2 + \beta_2 \, \text{Acc}_1 + \beta_3 s_1 + \beta_4 \left(|s_1| \times \text{Acc}_1 \right).$$

This analysis showed that the proportion of 'right' choices (in the second decision) significantly increased or decreased with the discriminability of the first stimulus ($|s_1|$) depending on whether the first decision was correct or wrong (see Main Text and Figure 1B).

2 Supplementary results

2.1 Analysis of response times

We conducted additional analyses to investigate how the response time (RT, an approximation of the deliberation time required to form the decision) in the second decision was related to the confidence in the first decision. Note that this analysis is not aimed at discriminating between models. In fact our models do not consider the dynamical aspects of the decision process, and do not make specific predictions about how degree of confidence and speed of decision should be related. Moreover, in our experiments we explicitly asked participants to prioritise accuracy, without worrying about the speed of their responses. Nevertheless, it is known that the confidence in a decision is usually inversely correlated with the response time [5]. Our goal here is therefore to provide a description of the relationship between confidence and speed of response in our dual decision paradigm, in order to provide a more complete description of the data and inspire future studies on the relationship between speed and confidence in the context of sequential decisions. The rationale of the analysis is the following: we begin by partialling out the effect of stimulus discriminability (both first and second) on the RT of the second decision. Next we calculate the expected confidence of the first decision (according to the ideal-Bayesian model, which corresponds to the objective probability of a correct response) and examine whether and how the expected confidence can account for the residual variations in RTs.

Given that participants had as much time as they wanted to respond, the RT distributions are likely to be contaminated by outliers. Therefore, prior to this analysis we removed trials with RTs longer than 5 sec after the offset of the stimulus ($\approx 0.28\%$ of total trials). The average RT (measured from the onset of the stimulus) in the first decision was 0.74 sec (SD across participants 0.19 sec), slightly slower than that of the second decision, 0.68 sec (SD 0.19 sec), $t(28) = 2.73$, $p = 0.01$ (see Supplementary Figure 2A). This difference was found only in the dual-decision condition: in the control condition RTs in the first and second decision had similar duration: first decision, mean 0.80 sec, SD 0.23 sec; second decision, mean 0.81 sec, SD 0.22 sec; $t(28) = 0.57$, $p = 0.57$, the Bayes factor (null/alternative) was 5.96, providing moderate support for the null hypothesis.

To control for the effect of stimulus discriminability on RTs, we fit a linear model to the logarithm of RTs, which included as predictors both the stimuli of first and second decisions as well as their interactions. Furthermore, given that RTs also tend to differ between correct and wrong responses, we also included the accuracy of first and second decisions as predictors. The regression model thus had the form

$$\log\left(\text{ResponseTime}\right) = \beta_0 + \beta_1 \left|s_1\right| + \beta_2 \left|s_2\right| + \beta_3 \left|s_1\right| \times \left|s_2\right| + \beta_4 \text{Acc}_1 + \beta_5 \text{Acc}_2 + \epsilon$$

where $\left|s_1\right|, \left|s_2\right|$ denote the absolute values of the stimuli in first and second decisions (the evidence discriminability), Acc_1 and Acc_2 the accuracy of first and second decisions, and ϵ a normally distributed residual error. The model was fit individually for each participant. The residuals of this model (i.e. the trial-by-trial differences between the observed and predicted logarithm of RTs) represent the residual variation in response times that was not accounted for by the stimuli levels and response accuracy.

Next, we computed the expected confidence according to the ideal observer. We do not have access to the noisy internal response of the participants (r_1), however we can calculate its expected value given the entire sequence of stimuli (s_1, s_2) and decisions (d_1, d_2). The distribution of the noisy internal response is a Gaussian centered on the true stimulus value. After observing the first decision of the participants, we can infer on which side of the decision criterion the internal response fell.
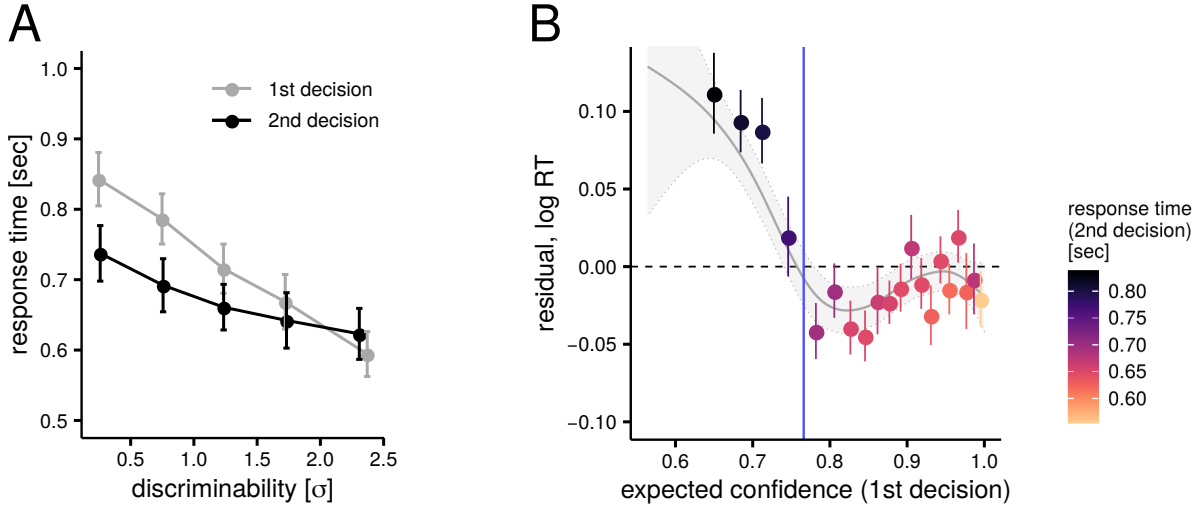
Supplementary Figure 2: Analysis of response times. A. Average response times in the dual-decision task, split by decision order and plotted as a function of the stimulus value (in units of internal noise). This plot shows that response times tend to be faster in the second decision than in the first, and particularly so for difficult decisions, mimicking the effect found in response accuracies (see Figure 1, Main text). Furthermore, this plot shows that response times decreases approximately linearly with the discriminability of the stimulus (i.e. with the average distance of the stimulus from the decision criterion). **B**. The residual variation in the response times of the second decision (after partialling out the effects of the two stimuli level and response accuracies, see text for details) is plotted as a function of the expected confidence of the Bayesian model in the first decision. Each point is the average residual (in logarithmic units) across participants for a given interval of confidence levels; error bars represent standard error of the mean. Bins are chosen such that each bin contains approximately the same number of trials. The smooth grey line represents a smooth GAM fit to the unbinned residuals (with 95% confidence interval). The color of each dot reprents the average response time in that bin, in seconds. Finally, the vertical blu line is the confidence level corresponding to the location at which participants placed the confidence criterion, according to the discrete model.

Therefore, the probability density function of the internal responses becomes a truncated Gaussian distribution. For example, when the decision is d_1^+, the distribution of r_1 can be expressed as

$$p\left(r_1 \mid d_1^+, s_1\right) = \frac{\phi\left(\frac{r_1 - s_1}{\sigma_1}\right)}{\sigma_1 \left(1 - \Phi\left(\frac{-s_1}{\sigma_1}\right)\right)}$$

which is a truncated Gaussian bounded from below at 0. (Φ and ϕ are the cumulative distribution function and the probability density function of the standard normal distribution, respectively.) The distribution of the internal response is then obtained by applying Bayes theorem

$$p\left(r_1 \mid d_1, s_1, d_2, s_2\right) = \frac{p\left(d_2 \mid s_2, r_1\right) \; p\left(r_1 \mid d_1, s_1\right)}{\int\limits_{-\infty}^{\infty} p\left(d_2 \mid s_2, r_1\right) p\left(r_1 \mid d_1, s_1\right) dr_1}.$$

For each trial, we computed the expected value of the internal response, that is $\mathbb{E}\left[r_1 \mid d_1, d_2, s_1, s_2\right] = \int_{-\infty}^{\infty} r_1 \; p\left(r_1 \mid d_1, s_1, d_2, s_2\right) dr_1$, and plugged it into equation (3) to obtain the expected confidence level.

We found that the expected confidence level (*in the first decision*) can explain the residual variation of RTs *in the second decision*. Specifically, second decisions are faster or slower - relative to the conditional mean of the log RT predicted by the regression model - when confidence in the

the first decision was higher or lower, respectively, as shown in Supplementary Figure 2B. Thus, this findings indicates that dynamical aspects of the second decision were adjusted based on trial-by-trial confidence in the first decision. Furthermore, we find that the RT residuals do not seem to decrease linearly with confidence, but rather are above the marginal mean when confidence is below a certain level (approximately corresponding to the average confidence level at which participants placed their confidence criterion w_1, as estimated by the discrete model, ≈ 0.77) then decrease rapidly below the conditional mean when confidence exceeds this level. This suggests a possible chronometric correlate of the discrete confidence states that we inferred from modelling participants' decisions. Although serendipitous, we believe this is an interesting result that could inspire future studies to apply dynamical models of decision making (e.g. drift diffusion models) in a dual-decision setting.
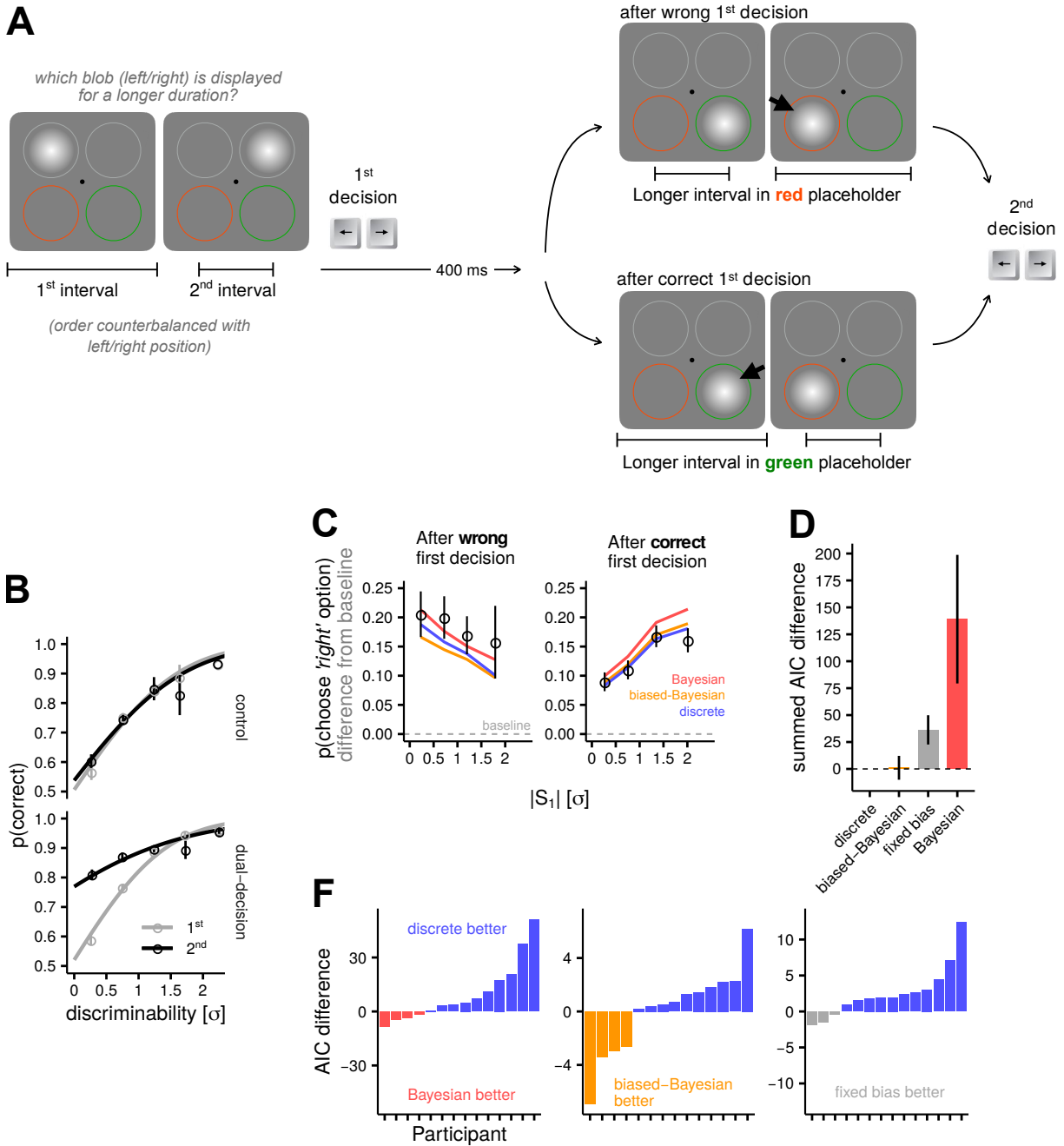
2.2 Dual-decision within the same perceptual task

Here we report the results of an additional experiment in which both decisions involved the same perceptual task, specifically a judgment about the duration of a visual stimulus (Supplementary Figure 3A). Although overall the results show a pattern similar to that found in our main 'crossed-task' experiment, they are also suggestive of slightly more efficient use of confidence information when the two decisions involve the same ('within-task'), as opposed to different ('crossed-task'), perceptual judgments. At first sight, a benefit in within-task relative to crossed-task experimental designs would seem to challenge the notion of confidence as a common currency [2, 3] (i.e. represented on a common scale) between different perceptual judgments. However, as explained below, within-task designs present some limitations. This, combined with differences in procedure across the twp experiments, makes strong conclusions based unwarranted.

One issue with within-task designs is that they are less constrained in terms of the possible strategies adopted by participants and, as a result, are less informative about the nature of the probability representations underlying performance in the dual decision task. To see that, recall the relationship that we found between the internal response r_1, elicited by the first signal and the optimal criterion θ_2 for the second decision, i.e.,

$$\theta_2 = -\frac{\sigma_2}{\sigma_1}\left|r_1\right|. \tag{12}$$

If the two decisions involve the *same* preceptual judgement, it is not unreasonable to assume that $\sigma_1 = \sigma_2$, and then $\theta_2 = -\left|r_1\right|$. An observer aware of this mapping could achieve optimal performance having available only a point-estimate of the first signal, without needing any knowledge about the internal noise. Such an observer would be indistiguishable from the optimal Bayesian model whilst using an entirely non-probabilistic strategy [6].

Our main crossed-task experiment instead used 2 different perceptual judgments. Since the signals upon which these judgments are based are different in nature (mean tilt from vertical versus signed motion coherence), they are processed by distinct neural mechanisms and likely encoded in different internal representations. Optimal performance in this task requires comparing the uncertainty in these two representations, and therefore implies some knowledge about their noise statistics (at the very least the ratio between the internal noises in the two tasks, σ_1 and σ_2, since the optimal mapping becomes $\theta_2 = -\frac{\sigma_2}{\sigma_1} \times \left|r_1\right|$). Our results from the crossed-task experiment indicate that participants do not use this mapping, but rather seem to rely on the simpler strategy of shifting

Supplementary Figure 3: Dual-decision within the same perceptual task. A. Experimental protocol of the within-task experiment on duration judgments. **B**. Probability of correct first and second decisions in control and dual-decision condition; error bars are bootstrapped standard errors across observers, and smooth lines are probit functions fit to the aggregated data. **C**. Probability of choosing 'right' in the second decision broken down by the accuracy of the first response. **D**. Difference in AIC relative to the discrete confidence model, summed over subjects; error bars are boostrapped standard errors. **E**. Ranked individual AIC differences.

the second decision's criterion by a fixed amount whenever their point-estimate of the first signal exceed a fixed confidence criterion (i.e. the discrete model). In the case of a within-task design, however, participants may use the trial-by-trial point-estimate of the first signal (already available to them) directly as criterion in the second decision. Furthermore, they could use this optimal (but heuristic) mapping only in a fraction of trials while falling back on the simpler strategy in other trials. This creates an interpretational confound that complicate disentangling if performance in the within-task case is truly underpinned by a Bayesian decision strategy.

2.2.1 Methods

The experiment was run at the Université Paris Descartes, and the protocol of the experiments was approved by the Paris Descartes University Ethics Committee for Non-invasive Research (CERES). A total of 14 in Experiment 2 (mean age 33, standard deviation 7; 6 females, 2 authors) participated in the experiment. An additional group of 5 subjects participated in the control condition (mean age 27, standard deviation 3; 3 females). All participants had normal or corrected-to-normal vision and gave their informed consent in written form prior to participation. All participants (except the two authors) were naïve to the specific purpose of the experiment. Naïve participants were compensated with either 10€ for each hour of experiment.

The visual display of the within-task design consisted of a central fixation point and four placeholders, Supplementary Figure 3A, continuously visible on a uniform grey background. The four placeholders were circles (anti-aliased) measuring 2.8 dva in diameter, whose centers were placed at 1.8 dva from the fixation point. Two placeholders were light grey, and were placed above the horizontal mid-line; the other two were placed below the mid-line and were colored in red the one on the left, and in green the one on the right. Stimuli consisted of white Gaussian blobs (standard deviation 0.65 deg), presented sequentially in the two placeholders (left/right). The order of presentation (left/right stimulus first) was balanced with respect to the longer/shorter duration of presentation. Participants were asked to indicate the location (left/right) of the longer duration blob. The shorter duration was always set to 600 ms, and the difference between shorter and longer durations was discretized in bins determined by the refresh rate of the monitor (100 Hz). The minimum duration difference was one single monitor refresh interval ($\approx$ 12 ms). The display screen was a 22 inches CRT monitor (screen resolution 1600×1200); viewing distance was 60 cm, controlled by means of a chin rest. Before the dual decision task, we measured individual sensitivities using a weighted up-down staircase procedure [4]. The purpose of this pre-test was to quickly obtain an estimate of the internal noise (σ) to adapt the range of stimuli in the dual-decision condition. The staircase procedure continued until 30 reversals were counted. The initial step size (the size of the decrease/increase of the difference between the two stimuli) was 4 refresh interval and was diminished 1 refresh after the second reversal. Stimuli in these pre-test measurements were presented only in the top placeholders. 5 participants run a version of the experiments where discrimination difficulties in the first and second decisions were drawn independently. The remaining 9 participants run a version of the experiment where we biased the probability of association between discrimination difficulties in the first and second decision. Specifically when the discriminability of the first stimulus was less than half the maximum possible value, there was a 0.7 probability that the discriminability of the second stimulus would be greater than half, and vice-versa. This was intended to encourage participants to make use of the rule, however we did not find any differences in performance across these two versions of the task and therefore data were pooled together in subsequent analyses. The number of trials were 300 and 500 for the version with biased and

random association between discrimination difficulties, respectively, to ensure balanced numbers of trial with transitions from easy to difficult decisions and viceversa.

2.2.2 Results

We applied the same exact analysis procedure used for the crossed-task experiment (see Main Text and Supplementary Methods). We found that decision order had no statistically significant effect in the control condition: 1.05, 95% CI [0.85, 1.31], $t(13) = 0.35$, $p = 0.74$; and that it had a significant effect in the dual-decision condition (Supplementary Figure 3B): the mean odds-ratio that participants responded correctly in the second decision relative to the first one was greater than one, 2.05, 95% CI [1.72, 2.51], $t(13) = 7.17$, $p = 7.12 \times 10^{-6}$. The size of this effect however was found to be larger than in the crossed-task experiment, $t(41) = 3.18$, $p = 0.003$.

When including the accuracy of the first decision as predictor, we found that also in this experiment participants were significantly more likely to choose 'right' in the second decision after a correct first decision than after an incorrect one: odds-ratio 2.36, 95% CI [1.69, 3.36], $t(13) = 4.72$, $p = 7.12 \times 10^{-6}$. The magnitude of this effect did not differ significantly from the crossed-task experiment, $t(41) = 0.58$, $p = 0.56$; the estimated Bayes factor (null/alternative) was 3.63, providing moderate support for the null hypothesis.

Moreover, by adding the discriminability of the first stimulus as a predictor, we found that it significantly influences the odds-ratio of choosing 'right' in the second decision: for each unitary increase in discriminability, the odds-ratio increases multiplicatively by a factor of 1.58, 95% CI [1.30, 1.95], $t(13) = 4.27$, $p = 9.13 \times 10^{-4}$, after a correct response and decreases by a factor of 0.74, 95% CI [0.57, 0.98], $t(13) = 2.07$, $p = 0.05$, after an incorrect response. There was no statistically significant difference in the size of both effects across the experiments: $t(41) = 0.62$, $p = 0.54$; and $t(41) = 0.32$, $p = 0.75$, respectively (the estimated Bayes factors, 3.55 and 4.02 respectively, providing moderate support for the null hypothesis).

Overall, these model-agnostic analyses indicate that although participants in the within-task experiment displayed a greater boost in performance (in decision 2 relative to decision 1), the extent by which the probability of choosing 'right' in the second decision changed as a function of first decision difficulty and accuracy did not differ significantly (in fact the Bayes factor provided evidence in favour of the null hypothesis of no difference between the two experiments). It thus remains unclear why participant in the within-task experiment achieved a greater boost in performance in the second decision: this may be due to the two decisions involving more similar judgments, or other differences in experimental procedures. For example, in the within-task experiment the range of stimuli was adjusted by running a staircase procedure, instead of a separate control condition, and on average it resulted in a slightly narrower range of stimuli when expressed in units of noise: the average range of stimuli in the within-task condition was $\pm 2.05\sigma$, whereas it was $\pm 2.67\sigma$ in the crossed-task experiment, $t(41) = 2.78$, $p = 0.008$.

Breaking down the trials according to the correctness of the first decision (Supplementary Figure 3C) reveals a pattern of deviations form the optimal Bayesian model very similar to what was observed in the crossed-task experiment, characterized by under-confidence after correct first decisions, albeit less evident, and with relatively higher-confidence errors. This pattern was again captured better by the discrete confidence model (compare data and model fit in Supplementary Figure 3C). However, in this case we find that the quality of model fit, as quantified by the AIC, did

not differ significantly between the discrete model and the biased-Bayesian model (Supplementary Figure 3D and E). The difference in AIC (summed over subjects) was 1.04 ± 11.26 in favour of the discrete model; the estimated Bayes factor (null/alternative) was 4.97, providing moderate evidence for the null hypothesis. These differences in quality of model fit, together with the greater boost in performance in the second decision, may indicate a more efficient use of confidence information in the within-task experiment. However, as explained above, there may be different explanations for this pattern that cannot be readily disentangled within the current set of results. On the one hand, it could indicated greater 'transmission fidelity' between internal states supporting decisions and confidence, i.e. a more precise and less noisy mapping between sensory evidence, confidence information and prior expectations when all decisions involve the same kind of sensory signals. On the other hand, it could indicate a greater heterogeneity in the strategies adopted by participants, since for this task the optimal strategy can be achieved effectively without needing additional information about noise or uncertainty. Since these difference computations may give rise to the same predicted behavioural outcome disentangling between them may require recording of neural activity in addition to modelling of behavioral data.

References

[1] K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* Springer New York, New York, US, 2nd editio edition, 2002.

[2] V. De Gardelle, F. Le Corre, and P. Mamassian. Confidence as a common currency between vision and audition. *PLoS ONE*, 11(1), 2016.

[3] V. de Gardelle and P. Mamassian. Does Confidence Use a Common Currency Across Two Visual Tasks? *Psychological Science*, 25(6):1286–1288, 2014.

[4] C. Kaernbach. Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics*, 49(3):227–229, may 1991.

[5] R. Kiani, L. Corthell, and M. N. Shadlen. Choice certainty is informed by both evidence and decision time. *Neuron*, 84(6):1329–1342, 2014.

[6] W. J. Ma. Organizing probabilistic models of perception. *Trends in cognitive sciences*, 16(10):511–8, oct 2012.

[7] J. C. Nash. On Best Practice Optimization Methods in {R}. *Journal of Statistical Software*, 60(2):1–14, 2014.

[8] J. C. Nash and R. Varadhan. Unifying Optimization Algorithms to Aid Software System Users: {optimx} for {R}. *Journal of Statistical Software*, 43(9):1–14, 2011.

[9] M. J. D. Powell. The BOBYQA algorithm for bound constrained optimization without derivatives. Technical report, Centre for Mathematical Sciences, University of Cambridge, Cambridge, UK, 2009.

[10] R Core Team. R: A language and environment for statistical computing, 2015.

[11] L. Rigoux, K. E. Stephan, K. J. Friston, and J. Daunizeau. Bayesian model selection for group studies - Revisited. *NeuroImage*, 84:971–985, 2014.

[12] A. N. Sanborn and N. Chater. Bayesian Brains without Probabilities. *Trends in Cognitive Sciences*, 20(12):883–893, dec 2016.

[13] H. Shteingart and Y. Loewenstein. The Effect of Sample Size and Cognitive Strategy on Probability Estimation. *Decision*, 2(2):107–117, 2015.

[14] K. E. Stephan, W. D. Penny, J. Daunizeau, R. J. Moran, and K. J. Friston. Bayesian model selection for group studies. *NeuroImage*, 46(4):1004–1017, 2009.

[15] E. Vul, N. Goodman, T. L. Griffiths, and J. B. Tenenbaum. One and done? Optimal decisions from very few samples. *Cognitive Science*, 38(4):599–637, 2014.

[16] E.-J. Wagenmakers and S. Farrell. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11(1):192–196, 2004.