Discrete confidence levels revealed by sequential decisions

Matteo Lisi[1,*], Gianluigi Mongillo[2,4], Andrei Gorea[3,4],

1 Centre for Applied Vision Research, City, University of London
2 Centre National de la Recherche Scientifique (CNRS), UMR8119, Paris, France
3 Centre National de la Recherche Scientifique (CNRS), UMR8242, Paris, France
4 Université Paris Descartes, Sorbonne Paris Cité, Paris, France

* matteo.lisi@city.ac.uk

Abstract

Whether humans are optimal decision makers is still a debated issue in the realm of perceptual decisions. Taking advantage of the direct link between an optimal decision-making and the confidence in that decision, we offer a new dual-decisions method of inferring such confidence without asking for its explicit valuation. Our method circumvents the well-known miscalibration issue with explicit confidence reports as well as the specification of the cost-function required by 'opt-out' or post-decision wagering methods. We show that observers' inferred confidence in their first decision and its use in a subsequent decision (conditioned upon the correctness of the first) fall short of both an optimal and an under-sampling behavior and are significantly better fitted by a model positing that observers use no more than four confidence levels, at odds with the continuous confidence function of stimulus level prescribed by a normative behavior.

Introduction

How do organisms select behavioral responses when interacting with a given environment? According to a popular theory – the rational choice theory – the organism evaluates, for each decision to be taken, its expected utility and selects thereafter the one yielding the highest rank. Despite the fact that the *descriptive* adequacy of rational choice theory has long been challenged on empirical as well as on theoretical grounds, mainly questioning its biological/psychological plausibility [14, 15, 17, 29], the debate is far from being settled [2, 3, 16]. In order to maximize expected utility, an organism has to be able to associate a *subjective* probability to each possible consequence of its actions. The characterization of this process requires (i) the quantitative assessment of the organism's ability to attach probabilities to events, and (ii) the appraisal of the measured subjective probabilities against those a Bayesian observer, with the same prior knowledge as the organism, would assign to the same events. The experimental instantiation of such comparisons has been hindered by serious methodological problems with assessing subjective probabilities.

Subjective probability is tantamount to the organism's confidence in the occurrence of an event E [6]. It is formally defined as a marginal rate of substitution [6, 26], that is, the organism's subjective probability about the event E occurring or having occurred is $p(E)$, if the organism is indifferent to gaining one unit of utility contingent on E against gaining $p(E)$ units of utility for sure. Current methods of measuring subjective probabilities hence confidence in, e.g., perceptual decisions (about the choice being correct) using opt-out or post-decision wagering techniques are straightforward operationalizations of the above definition. A major, well and long known (see, e.g., [26]), problem with these methods is that

Significance statement
A normative decision behavior requires that one's confidence be a continuous function of the difficulty of decision to be made (the posterior probability that the decision was correct given the evidence). People however concede that they could not discriminate between two indefinitely close confidence states. Using a new experimental paradigm that does not require participants' explicit confidence evaluation nor knowledge of their utility function, we show that people can discriminate only two and at most four confidence states, the implication of which is that they are not optimal decision-makers. This finding counters a plethora of studies suggesting the contrary and also rejects an alternative stand according to which people's decisions are based on a sample of an available posterior function.

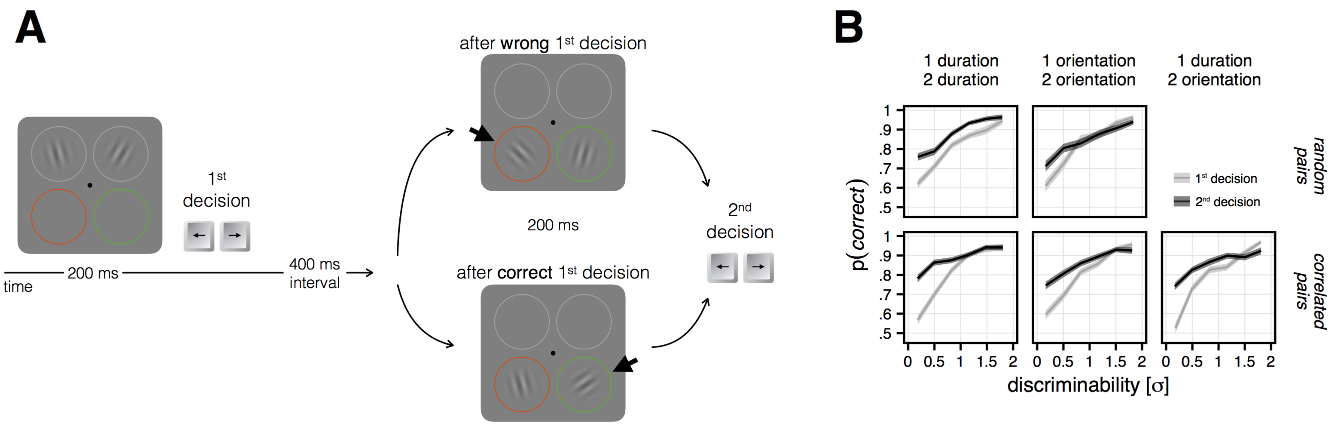

Figure 1. Experimental paradigm and results. Orientation-orientation (O-O) paradigm (**A**). In this task the participants had to make two consecutive decisions on which of two Gabor-patches was more tilted from the vertical (hence the signal on which the participant had to base the decision was the difference in orientation between the two Gabors). While the absolute difference in tilt between the two Gabors (the discriminability) is drawn independently from a uniform distribution in each of the two Gabor-pairs , the location of the more tilted Gabor in the second Gabor-pair (left/right; indicated by the small thick black arrows in the right-hand panel) was made depended on the accuracy of participant's first response. Faced with this task an ideal observer would assign a prior probability equal to its confidence – of being correct in the first decision – that the more tilted Gabor in the second Gabor-pair was displayed in the right-hand (green) placeholder. Performance (proportion of correct responses) plotted as a function of discriminability (measured in units of internal noise) separately for 1st and 2nd decisions/Gabor-pairs in all tasks and conditions. **B**) Note how performance in the second decision is higher, especially for very small orientation differences (where performance would be at chance in the absence of prior information). Performance is averaged over 6 equally spaced discriminability bins. The thickness of the traces represents bootstrapped standard errors across participants.

they rely on unverifiable assumption about the utility function (which cannot be measured) of the participant. Such methods cannot disentangle subjective probability from factors such as opportunity cost in waiting-time paradigms. More frequently used, methods requiring *explicit* confidence valuation by the decision-maker suffer from well-documented miscalibrations and response biases (see, e.g., [10, 11, 21]).

Here, we present a novel approach of estimating subjective probabilities which overcomes all the problems above. Human participants were presented with two consecutive signals and asked to decide whether they were above or below some reference value. The key innovation was that the statistics of the second signal was made contingent upon the decision-maker having made a correct decision on the first signal (explicit feedback is not provided): correct/incorrect first decisions resulted into signals above/below the reference value for the second decision. Differences in performance between the second and the first decisions, at signal parity, allow the estimation of the subjective probability of being correct on the first decision (i.e., the confidence). Using this approach we show that humans are quite accurate in assessing confidence, yet they exhibit systematic deviations from optimality, most commonly in the form of under-confidence. These systematic deviations cannot be accounted for by a sample-based approximation of the optimal Bayesian strategy (see Supplemental information). An alternative non-Bayesian model, characterized by a finite number of discrete confidence levels, provides the best and most parsimonious description of the empirical patterns of humans' choices.

Results

The dual-decision paradigm

The paradigm is illustrated in Figure 1A with an example of the orientation discrimination task (orientation-orientation or O-O condition, see Material and methods and Supporting information for details), in which the participant is presented with two consecutive Gabor-pairs, and for each pair must decide which of the two Gabors was more tilted from the vertical; correct/incorrect first decisions result into displaying the more tilted Gabor in the right/left

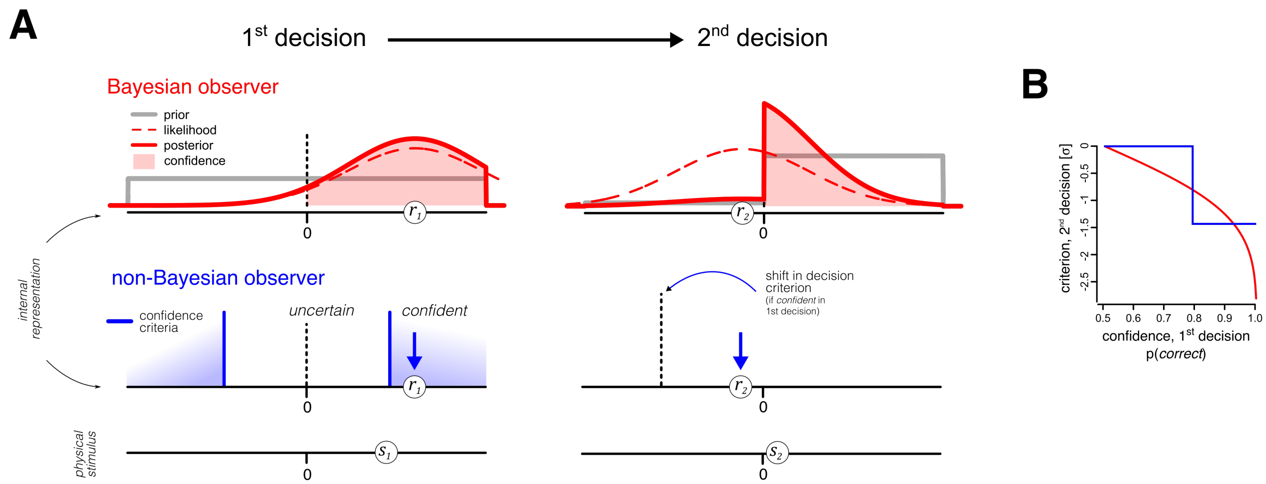

Figure 2. Bayesian and a class of non-Bayesian models of confidence and of sequential decision-making. (**A**) Internal beliefs of the Bayesian (top panels) and non-Bayesian observer (bottom panels), for first and second decision (from left to right).The lower abscissae represent the state of the world, i.e. the actual difference in tilt between the two Gabors in a pair, while the internal belief distributions about the decision variable are referred to the upper two abscissae. The first stimulus in this ad-hoc trial (s_1) evokes a noise contaminated internal response ($r_1 = s_1 + \eta$, where η is Gaussian noise with variance σ^2), which is the only information available to the participant. The Bayesian observer has full knowledge about the statistics of the internal noise, and can accurately compute a likelihood function which provides the probability of observing that specific r_1 for each possible values of the stimulus. Because for the first decision the prior is flat (grey lines in the upper panels), the likelihood (dashed curves) agrees with the posterior distribution (red continuous curves) about the real location of s_1. The Bayesian probability that represents the confidence c_1 in the first choice ('right', the correct response) corresponds to the area of the posterior distribution filled in red ($c_1 = p\,(s_1 \geq 0|\mathrm{s}_1)$). The decision variable for the second response is represented in the right upper panel. Note that the prior distribution (grey line) now assigns a larger probability (equal to the confidence in the first choice) to the possibility that the more tilted Gabor appeared in the right placeholder. Inasmuch as the response to the first task was correct, the true value of s_2 for the second decision is indeed positive (lower mid-panel). s_2 being however very small (as illustrated), it evokes by chance (due to internal noise) a negative internal response r_2 that would (in the absence of prior information) lead to the erroneous conclusion that the more tilted Gabor was on the left. Nevertheless, because of the asymmetrical prior distribution, the posterior distribution is still favoring the correct response. The lower left panel illustrates how the non-Bayesian observer (who does not have knowledge about the statistics of the internal noise, and only "sees" point estimates) could reach the same response as the Bayesian observer by comparing the internal response r_1 with an additional confidence criterion (the two vertical blue lines symmetrically placed about the decision criterion, dashed vertical line). By means of this heuristic rule the non-Bayesian observer discriminates 'confident' decisions, from 'uncertain' or 'non-confident' decisions (without assigning a numerical probability to his confidence). Only in trials where this observer is confident, he will shift the second decision criterion, with the result of increasing the frequency of responses 'right'. It can be demonstrated that the Bayesian observer's use of prior expectations for the second decision also amounts to shifting its the decision criterion (see Supporting Information), and that the optimal criterion in the second decision is a function of the confidence c_1 in the first decision. The optimal criterion shift, plotted as a function of the confidence c is represented in panel **B** (red curve), together with an example of criterion shift for the non-Bayesian observer (staircase blue line).

place-holder, respectively. The very same experimental format was applied to a duration discrimination task (duration-duration or D-D condition) where participants had to decide which of two Gaussian blob flashes (presented sequentially) was displayed for a longer duration. These two conditions were tested both in a first experiment where the difficulties of first and second decision were independently drawn from a uniform distribution (random-pairs experiment) and, on a different group of subjects, in another experiment where these were not independent (correlated-pairs experiment): more specifically difficult/easy first decisions were more likely to be followed by easy/difficult second decisions, respectively (see Material and methods). This correlation was introduced to encourage participants to exploit the statistical dependance between the signal in the two decision. Additionally, we also tested a condition where the two tasks were combined, duration-orientation or D-O condition, to test within our paradigm the hypothesis that confidence may work as a 'common currency' between different perceptual judgments [8, 9].

Given the specific stimuli presented in the first decision and participant's trial-by-trial response, participant's performance in the dual-decision paradigm can be compared with that of a Bayesian ideal observer that accurately estimates the probability of being correct in the

Table 1. Results of the logistic analysis that measured the influence of the correctness of the first decision on the probability of choosing 'right' in the second decision (after controlling for the stimulus).

Condition	Experiment	χ^2	p	β	odds-ratio
D-D	correlated-pairs	13.73	0.0002	2.26	9.59
D-D	random-pairs	15.70	<.0001	3.54	34.53
O-O	correlated-pairs	11.69	0.0006	1.85	6.34
O-O	random-pairs	0.90	0.46	0.90	2.46
D-O	correlated-pairs	13.30	0.0003	2.16	8.64

first decision and uses it as prior information for the second decision, according to the rules of Bayesian decision theory (see Figure 2A). We have also compared human's behavior with that of two other classes of non-Bayesian models, one characterized by a finite number of discrete confidence levels; see Figure 2A bottom), and another representing a sample-based approximation of the Bayesian model (where the posterior distribution is approximated by finite number of samples; see Supplemental information).

Model-free analyses

In figure 1B we plotted the proportion of correct responses as a function of discriminability (measured in units of internal noise) in the first and second decision. As it can be seen, second decision performance (darker traces) is higher than first decision performance, especially for the most difficult trials (where the difference between the two stimuli in a pair was very small and first decision performance was close to chance). Since the more tilted or longer stimuli appeared more often in the right placeholder for the second decision (given that task difficulty was adjusted so as to yield average first-decision performance above chance), an increase in second-decision performance could be the result of a fixed bias (i.e., participants having chosen a 'right' response more frequently), without necessarily involving a trial-by-trial monitoring of uncertainty. To control for this possibility we performed a logistic analysis to measure the influence of the correctness of the first decision on the probability of choosing 'right' on the second decision. For each experiment (correlated- and random-pairs) and condition (O-O, D-D, and D-O) we fitted a hierarchical (mixed-effects) logistic regression, using **R** [23] and the lme4 package [1], with the absolute difference between the two stimuli (in units of σ) and the accuracy of the first response as fixed effect predictors, and the participant as a grouping factor. We evaluated statistically the effect of the correctness of the first response with a likelihood ratio test between the fitted model and a reduced model where the effect of the first response was set to 0. This test resulted significant for all the experiments and conditions (see Table 1), with the exception of the O-O condition in the random-pairs experiment. In order to check whether a simpler fixed-bias model would really suffice to describe performance in this latter case, we performed a Monte Carlo simulation. For each trial we estimated the expected probability of a 'right' second response on the basis of the stimuli presented and the psychometric function fitted to the first decision responses. This results in a set of Bernoulli trials with different probabilities of success, which we simulated 10^5 times in order to estimate a 95% confidence interval (using the percentile method) on the expected proportion of second responses 'right' given the stimuli and the first response. The results revealed that after a wrong first response none of the participants responded 'right' more often than what would be expected given the stimuli: all the observed proportions resulted within the confidence intervals (which were calculated at the Bonferroni corrected alpha level 0.005). Instead, after a correct response, the observed proportion of responses 'right' exceeded the confidence interval for 4 out of the 5 participants in this experiment. This result suggests that also in this condition, where the fixed bias hypothesis could not be rejected at the group level, we find evidence that most participants have monitored the confidence in their first response on a trial-by-trial basis.

Model comparison

On average, the cross-validated log-likelihoods were higher for the non-Bayesian class of models, regardless of the number of confidence levels (Figure 3). The log-likelihoods were computed on the left-out samples, meaning that the non-Bayesian observer was better at predicting the actual behavior of the participants. With very few exceptions (see Supporting Information, Figure S2), the non-Bayesian class of models also outperformed the Bayesian models for each individual subject. It must be noted that the cross-validated log-likelihoods show very little change with the number of confidence levels (and thus of the number of free parameters). Still some participants' data were slightly better predicted by the models with 3-4 levels of confidence (see Supplemental information, Figure S2). The finding that the cross-validated log-likelihood did not decrease nor increase systematically with the increasing number of free parameters is most likely due to the limited number of trials run in each experiment and condition. Indeed, the variance of the estimated parameters increases drastically with the number of implemented confidence levels (see Supplemental information, Figure S3).

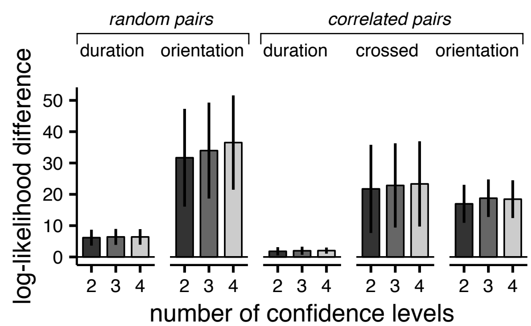

Figure 3. Cross-validation results. Log-likelihood differences (non-Bayesian minus Bayesian), averaged across participants, are positive, indicating that the non-Bayesian class of models had a better predictive performance. Error bars represent bootstrapped standard errors of the mean (SEM).

Given the results of the cross-validation, it is important to understand the differences between the predictions of the Bayesian and non-Bayesian models. Figure 4 displays the observed proportion of 'right' second decision choices and the corresponding correct responses (top and bottom panels, respectively) as a function of the difficulty of the first decision averaged across observers (see Supporting Information, Figure S3 for individual data). It can be seen that the easier the first decision, the more do subjects tend to choose a 'right' in the second decision, and the more correct are their responses. While both models, Bayesian and non-Bayesian, capture the trend in participants' responses, the non-Bayesian model (here with 2 confidence levels) is much closer to the empirical data. On average, the data show an under-confidence bias (with respect to optimal): participants responded 'right' in the second decision less frequently than what would be expected given the stimuli presented, indicating that they underestimated their probability of being correct (see Figure 4, upper panels). Based on this observation we rejected also sampling-based approximations of the Bayesian model as a viable explanation of the data, as such models predict systematic overconfidence biases (see Supplemental information). Indeed individual data revealed an over-confidence bias (e.g. see participant 4, D-D condition, Supplemental information, Figure S4B, or participant 8, D-O condition, Supplemental information, Figure S4A) only in in few cases (5 out of 37 combinations of participant and condition).

Discussion

We developed a novel approach where in a sequence of two dependent perceptual decisions humans could improve their second decision performance by taking advantage of the fact that the statistics of the stimuli presented for this second decision depended on the correctness of their first decision. This experimental protocol can be regarded as a laboratory proxy of more

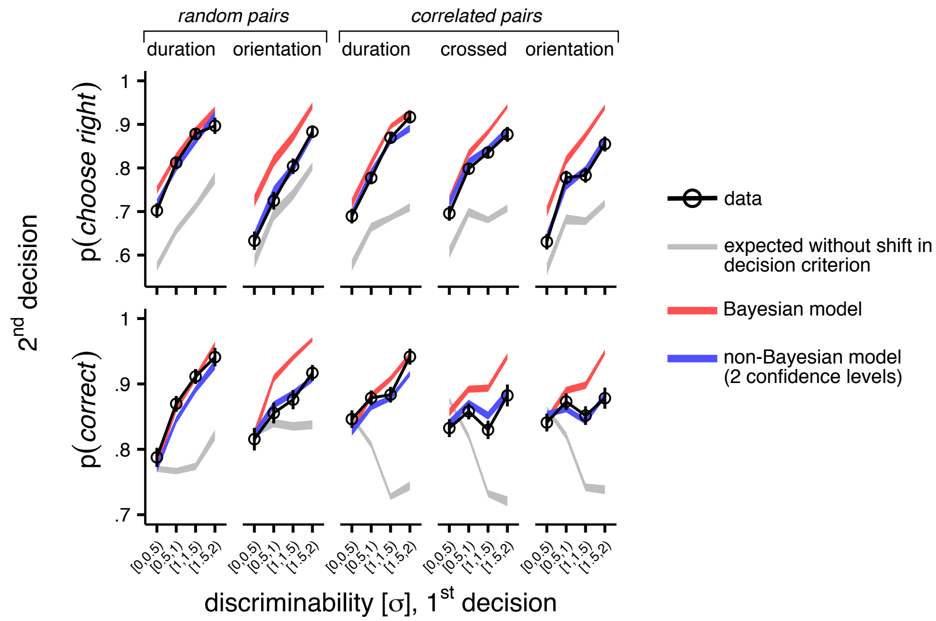

Figure 4. Comparison of model predictions and empirical data. See main text for details. The grey bands represent the expected proportion of 'right' (i.e. reporting $s_2 > 0$) and correct choices that should have been observed if the participants did not adjusted their criteria for the second decision. Error bars, and the width of the bands represents $\pm$1bootstrapped SEM.

complex environments in which confidence is used to guide behavior in cases where a current decision depends on the unknown outcome of a previous one. We developed a normative Bayesian observer model of this task, i.e., of an ideal observer who performs Bayesian inference to estimate the posterior probability of being correct in the first decision (the confidence), and use it optimally (i.e., maximizing the probability of correct second decision) as prior probability distribution for the making the second decision. By comparing participants' performance with the predictions of the Bayesian model, we were able to test the hypothesis that human observers are able to evaluate probabilistically their own uncertainty and make perceptual decisions according to the rules of Bayesian inference. More specifically, our paradigm allows the simultaneous test of two main tenets of Bayesian theories of perceptual decision-making, namely whether humans: (1) can and do compute the posterior probability that their choice is correct, and (2) use this posterior probability and combine it optimally as a prior for incoming sensory observations. Previous studies have repeatedly supported the hypothesis that human observers can combine prior information (e.g. due to environmental statistics) with sensory observations optimally or nearly-optimally [18, 27] and use this prior to adjust decision criteria [12, 13], even when these priors are learned on very short timescales [5]. These results would suggest that any deviation from optimality in our tasks should be attributed to the internal assessment of confidence, rather than to its use as a prior for the next perceptual decision.

The present results revealed clearly that the participants engaged in this task exhibit systematic deviations from the predictions of the normative Bayesian model. While they were clearly able to take trial-by-trial uncertainty into account (a simple fixed bias model was not sufficient to account for the observed behavior, see Results, Model-free analyses), their pattern of second decisions revealed the presence of a global miscalibration of confidence, specifically under-confidence, as indexed by lower rates of responses 'right' in the second decision with respect to ideal. These biases could be accounted for by our non-Bayesian class of models. From a psychological point of view, the non-Bayesian class of models posits that confidence is discretized in a number of distinct levels, as also suggested by previous work [31]. For example, in the one-criterion variant of our model there would be two discrete 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 wrong (see Figure 2). This one confidence criterion variant of our model provided the best and most parsimonious description of the empirical data (see Results, Model comparison). Because the heuristic model does not assume any knowledge about the variability of the internal decision variable, these levels are defined only on an ordinal scale and do not convey a precise numerical information about the probabilities involved. Taken together, our experimental and modelling results indicate that humans can use sensory evidence to perform comparative probabilistic judgments, but ultimately cannot assign numerically precise subjective probabilities to the two events. These comparative probability judgments, not linked to precise numerical values, are the essence of qualitative probability reasoning [25], a weakened but more pragmatic and intuitive counterpoint of classical probability theory.

The notion of qualitative probability lies at the foundations of the notion of subjective (or Bayesian) probability since its early formulation [7, 25]. Much of the early work on the theory of Bayesian probability has been dedicated to identifying the conditions that allow the departure from the qualitative probability toward the quantitative (numerically precise) probability, defined according to the classical Kolmogorov axioms [20]. Several propositions have been put forward, but all of them ultimately assume that the decision-maker's knowledge allows the partition the probability space associated with the event space into a uniform and arbitrarily large collection of disjoint states or events [25]. While this assumption can be used to provide a formal framework for exact reasoning under uncertainty, it may be too fine-grained for a realistic, biological decision-maker. Indeed, in the real world assigning exact numerical probabilities is often difficult or impossible, and the ability to compare the likelihood of two events without having to provide exact probabilities may be sufficient. Our results support this idea, as the performance of the heuristic observer show a marked increase in the probability of correct second decisions relatively to a simpler model which does not consider any prior information (compare the blue with the grey traces in the lower panel of Figure4, while the ensuing performance benefit obtained by the Bayesian model may not be enough to justify its increased computational complexity and costs.

Ackowledgments

This work was supported by Grant ANR-12-BSH2-0005 from the French National Research Agency to A. Gorea.

Material and methods

Participants (5 for the random-pairs experiment, and 9 for the correlated-pairs experiment) performed a sequence of two decisions on each trial, with the statistics of the second decision stimuli being conditioned on the correctness of the first decision. All experiments were performed in accordance with French regulations and the requirements of the Helsinki Convention. The protocol of the experiments was approved by the Paris Descartes University Ethics Committee for Non-invasive Research (CERES).

We describe here the general structure of the protocol, with the details of the implementation and of the analysis being provided in the Supplemental information. At the beginning of each trial, two stimuli were presented in two placeholders on the left and right of the fixation point. The two stimuli differed along one physical dimension [orientation of the Gabor-patches – spatial frequency 1.5 cycles/dva (degrees of visual angle), standard deviation of the envelope 0.7 dva, contrast 25%, – or duration – Gaussian blobs with a pick intensity of ≈ 25.3 cd/m^2 and a standard deviation of 0.65 dva]. The participant was required to indicate which of the two stimuli was characterized by a higher value along the given dimension by pressing the left/right arrow keys. The difference between the two stimuli was uniformly distributed within 2 JNDs (just noticeable differences), measured in preliminary sessions (see below). 400 ms after providing the first response, a second pair of stimuli was presented and the participant was again asked to indicate which of the two has a higher value. The difference in value between the stimuli in the second pair was also randomly sampled from a uniform distribution. However the location of the higher-value stimulus depended this time on the correctness of the first response: if the first response was correct, the higher-value stimulus was presented on the right, and on the left otherwise. Participants were informed about this rule, and were asked to use it in order to achieve the best possible second decision accuracy. Before starting the experimental trials, participants were explained the rule and were familiarized with a version of the task where the difference between the two stimuli to be compared could go up to very high values (up to 45° in the orientation task, and up to 1 second in the duration task, uniformly distributed). The large differences in the practice session were intended to make the rule clear and unambiguous for all participants. In one first experiment discrimination difficulties in the first and second decisions were drawn independently (random-pairs). In a second experiment we biased the probability of association between discrimination difficulties in the first and second decision (correlated-pairs, see Supplemental information, Figure S5). Specifically, when the difference in intensity in the first pair was less than 1 JND, there was a 0.7 probability that the difference in the second pair would be larger than 1 JND, and vice versa. This was intended to encourage participants to make use of the rule. The random-pairs experiment was tested in two conditions, run in different session on different days (order balanced), where the two decisions involved both an orientation discrimination (O-O) or a discrimination of duration (D-D). Each of these sessions comprised 500 trials. The correlated-pairs experiment was declined in three different conditions, O-O, D-D and D-O (where the two decisions involved a duration discrimination followed by an orientation discrimination). Each of the three sessions comprised 300 trials, each consisting of two consecutive perceptual decisions. The different number of total trials in the correlated- and random- pairs experiments was designed so that they resulted in similar number of easy (difficult) decisions followed by difficult (easy) decisions. Each testing session was divided in 10 blocks of trials. At the end of each block participants were given a feedback about the overall accuracy of their second decisions in that block. Additionally, to help participants keep track of their performance, starting with the end of the second block they were also informed on whether their accuracy had increased or decreased with respect to the previous block.

Analysis

For each participant we estimated the standard deviation of the internal noise, σ, by fitting a cumulative Gaussian psychometric function on proportion of 'right' choices in the first decision (individual psychometric functions are reported in Supporting information, Figure S6). Next, we used the estimated σ to transform the values of s from raw units (e.g. degrees and seconds) to units of internal noise. Finally, we fitted the non-Bayesian models using maximum likelihood estimation, and compared the predictions of Bayesian and non-Bayesian models. Because these models differs in the number of free parameters, in order to prevent overfitting we performed a leave-one-out cross-validation procedure, and compared the models on the basis of the cross-validated log-likelihoods summed over all the left-out trials. All analyses were performed in the open-source software **R** [23]; the data and the code of the analysis are available upon request. The mathematical details of the computational models are provided in the Supporting Information.

References

1. D. Bates, M. Maechler, B. Bolker, and S. Walker. lme4: Linear mixed-effects models using Eigen and S4, 2014.
2. J. S. Bowers and C. J. Davis. Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138(3):389–414, 2012.
3. J. S. Bowers and C. J. Davis. Is that what Bayesians believe? reply to Griffiths, Chater, Norris, and Pouget (2012). *Psychological Bulletin*, 138(3):423–426, 2012.
4. D. H. Brainard. The Psychophysics Toolbox. *Spatial Vision*, 10(4):433–436, jan 1997.
5. M. Chalk, A. R. Seitz, P. Seriès, and P. Series. Rapidly learned stimulus expectations alter perception of motion. *Journal of vision*, 10(8):2, jul 2010.
6. B. De Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.
7. B. de Finetti. *Theory of Probability*. Wiley, New York, 1975.
8. V. De Gardelle, F. Le Corre, and P. Mamassian. Confidence as a common currency between vision and audition. *PLoS ONE*, 11(1), 2016.
9. V. de Gardelle and P. Mamassian. Does Confidence Use a Common Currency Across Two Visual Tasks? *Psychological Science*, 25(6):1286–1288, 2014.
10. J. Drugowitsch, R. R. Moreno-Bote, and A. Pouget. Relation between Belief and Performance in Perceptual Decision Making. *PLoS ONE*, 9(5):e96511, 2014.
11. S. M. Fleming and H. C. Lau. How to measure metacognition. *Frontiers in human neuroscience*, 8(443):1–9, 2014.
12. N. Gekas, M. Chalk, A. R. Seitz, P. Seriès, and P. Series. Complexity and specificity of experimentally-induced expectations in motion perception. *Journal of vision*, 13(4):8, mar 2013.

13. N. Gekas, A. R. Seitz, and P. Seriès. Expectations developed over multiple timescales facilitate visual search performance. *Journal of vision*, 15(9):10, 2015.

14. S. J. Gershman, E. J. Horvitz, and J. B. Tenenbaum. Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245):273–278, 2015.

15. G. Gigerenzer and W. Gaissmaier. Heuristic decision making. *Annual Review of Psychology*, 62:451–482, 2011.

16. T. L. Griffiths, N. Chater, D. Norris, and A. Pouget. How the bayesians got their beliefs (and what those beliefs actually are): comment on bowers and davis (2012). *Psychological Bulletin*, 138(3):415–422, 2012.

17. R. J. Herrnstein and D. Prelec. Melioration: A theory of distributed choice. *The Journal of Economic Perspectives*, 5(3):137–156, 1991.

18. M. Jazayeri and M. N. Shadlen. Temporal context calibrates interval timing. *Nature Neuroscience*, 13(8):1020–1026, jun 2010.

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

20. A. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeits- rechnung*. Springer, Berlin, 1933.

21. P. Mamassian. Visual Confidence. *Annual Review of Vision Science*, 2(1):annurev–vision–111815–114630, 2016.

22. D. G. Pelli. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4):437–442, jan 1997.

23. R Core Team. R: A language and environment for statistical computing, 2015.

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

25. L. J. Savage. *The foundations of statistics*. John Wiley & Sons, Inc., 1954.

26. L. J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

27. P. Seriès and A. R. Seitz. Learning what to expect (in visual perception). *Frontiers in Human Neuroscience*, 7(October):668, 2013.

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

29. H. A. Simon. A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1):99, feb 1955.

30. 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.

31. H. Zhang, N. D. Daw, and L. T. Maloney. Human representation of visuo-motor uncertainty as mixtures of orthogonal basis distributions. *Nature Neuroscience*, 18(8):1152–1158, 2015.

Supplemental information

Material and methods

Apparatus

Participants sat in a quiet, dimly lit room, with the head positioned on a chin rest at a distance of 60 cm from the display screen, a gamma-linearized Mitsubishi Diamond Plus 230SB CRT monitor (screen resolution 1600x1200, vertical refresh rate 85 Hz). Stimuli were generated by a computer running Matlab (Mathworks) with the Psychophysics Toolbox [4, 22].

Stimuli and procedure

All experiments were run with the same visual display, consisting of a central fixation point and four placeholders, continuously visible on a uniform gray background (luminance ≈ 13.6 cd/m^2). The four placeholders were circles measuring 2.8 dva (degrees of visual angle) in diameter, whose centers were placed at 1.8 dva from the fixation point. Two placeholders were grey (≈ 15.5 cd/m^2), and were placed above the horizontal midline; the other two were placed below the midline and were colored in red the one on the left, and in green the one on the right (their luminance was matched with the grey placeholders, ≈ 15.5 cd/m^2).

In the orientation (O-O) task the stimuli were two Gabor gratings (sinusoidal luminance modulation presented within a Gaussian contrast envelope) of different orientations, presented for 200 ms. The spatial frequency of the Gabors was set at 1.5 cycles/dva, the phase was drawn randomly, and the standard deviation of the Gaussian envelope was 0.7 dva. The Gabor displayed in the left placeholder was always tilted to the left, and the one appearing in the right placeholder was always tilted to the right. The task of the participants was to indicate which Gabor was more tilted from the vertical; the less tilted of the two Gabors was always tilted by 15°; the minimum difference was 0.1°.

In the duration (D-D) task the two stimuli consisted of white Gaussian blobs (standard deviation 0.65 dva), 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 vertical refresh of the monitor (≈ 12 ms). The minimum duration difference was one single monitor refresh interval.

Pre-test JND measurement

Before the orientation and duration task, we measured individual JNDs using a weighted up-down staircase procedure [19]. The purpose of this pre-test was to quickly obtain a measure of the JND in order to adapt the range of stimuli in the main experiment to individual sensitivities. 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 2° in the orientation task and 4 refresh intervals in the duration task (≈ 50ms), and was diminished to 0.5° and 1 refresh after the second reversal. Stimuli in these pre-test measurements were presented only in the top placeholders.

Participants

5 subjects (2 female; mean age 30.8, standard deviation 3.1) participated in both conditions of the random-pairs experiment (D-D, and O-O; see Main text, Material and methods). 9 participants (4 female, 2 authors; mean age 33.9, standard deviation 9.6) participated in the 3 conditions of the correlated-pairs experiment (D-D, O-O, D-O). All participants (except the author) were naïve to the specific purpose of the experiment. All conditions were performed in separate session on different days. The order of the D-D and O-O conditions was counterbalanced across subjects, while the D-O condition was always performed in the last session. All participants had normal or corrected-to-normal vision and gave their informed consent to perform the experiments.

Computational models

To model the performance in our task, we considered that the observer forms a decision variable based on the difference in the intensity between the left and right stimuli, $s = |s_{\text{right}}| - |s_{\text{left}}|$, where s_{right} and s_{left} indicates the deviation from vertical of the two Gabor gratings in the orientation task, or the duration of the two blobs in the duration task. Hence the task amounts to deciding whether the signal s is greater or less than 0. We assumed that the observer has only access to a corrupted version of s, $r = s + \eta$, where η is Gaussian noise with variance σ^2. In the first decision s is uniformly chosen from an interval centered 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. The probability $p(+|s_1)$ of the observer reporting $s_1 > 0$, is given by

$$p\left(+|s_1\right) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{s_1}{\sigma\sqrt{2}}\right)\right] = \Phi\left(\frac{s_1}{\sigma}\right) \tag{1}$$

where Φ is the cumulative distribution function of the standard normal distribution. We fitted this function by maximum likelihood to estimate the value of σ, for each participant and task (duration and orientation), taking into account only the first decisions (see Supplemental Figure S6).

Bayesian observer

The ideal Bayesian observer has full knowledge of the statistics of the internal noise. Because we assumed the internal noise to be Gaussian, the likelihood function, that is the probability of observing r_1 given s_1 is

$$p\left(r_1|s_1\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{r_1 - s_1}{\sigma}\right)^2} \tag{2}$$

In first decision, the prior probability of s_1 is uniform with the range $(-R, R)$, that is $p(s_1) = 1/(2R)$ if $-R \leq s \leq R$ and 0 otherwise. Taking the prior into account, the unconditioned probability of observing r_1 can be computed from

$$p\left(r\right) = \int_{-R}^{R} p\left(r_1|s_1\right)\, p\left(s_1\right)\, ds_1 \tag{3}$$

$$= \frac{1}{4R}\left[\operatorname{erf}\left(\frac{R + r_1}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{R - r_1}{\sigma\sqrt{2}}\right)\right] \tag{4}$$

The posterior probability of s_1 having observed r_1 can be obtained applying Bayes rule

$$p\left(s_1|r_1\right) = \frac{p\left(r_1|s_1\right) p\left(s_1\right)}{p\left(r_1\right)} \tag{5}$$

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

$$c_1^+\left(r_1\right) = \int_{0}^{R} p\left(s_1|r_1\right) ds_1 = \frac{\operatorname{erf}\left(\frac{r_1}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{R - r_1}{\sigma\sqrt{2}}\right)}{\operatorname{erf}\left(\frac{R + r_1}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{R - r_1}{\sigma\sqrt{2}}\right)} \tag{6}$$

When $c_1^+\left(r_1\right) \geq 1/2$ the observer reports that $s_1 > 0$, otherwise he reports that $s_1 < 0$. The probability of being correct in the first decision, that is the confidence of the ideal observer, is given by $c_1 = \max\left[c_1^+\left(r_1\right),\ 1 - c_1^+\left(r_1\right)\right]$. The ideal observer would use his confidence 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 is drawn from the positive interval

$$p\left(s_2\right) = \begin{cases} \frac{c_1}{R}, & \text{if } 0 < s_2 \leq R \\ \frac{1 - c_1}{R}, & \text{if } -R \leq s_2 < 0 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

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

$$c_2^+\left(r_2\right) = \frac{c_1\left[\operatorname{erf}\left(\frac{r_2}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{R - r_2}{\sigma\sqrt{2}}\right)\right]}{\left(1 - c_1\right)\operatorname{erf}\left(\frac{R + r_2}{\sigma\sqrt{2}}\right) + c_1 \operatorname{erf}\left(\frac{R - r_2}{\sigma\sqrt{2}}\right) + \left(2c_1 - 1\right)\operatorname{erf}\left(\frac{r_2}{\sigma\sqrt{2}}\right)} \tag{8}$$

This equation reduces to the one for the first decision variable when $c_1 = 1/2$, as it should. Note also that the range R on which s_1 and s_2 takes values is immaterial. It is possible to simplify equations and by taking the limit $R \to \infty$. In this limit one obtains

$$c_1^+\left(r_1\right) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{r_1}{\sigma\sqrt{2}}\right)\right] \tag{9}$$

$$c_2^+\left(r_2\right) = \frac{c_1\left[1 + \operatorname{erf}\left(\frac{r_2}{\sigma\sqrt{2}}\right)\right]}{1 + \left(2c_1 - 1\right)\operatorname{erf}\left(\frac{r_2}{\sigma\sqrt{2}}\right)} \tag{10}$$

The decision rules described by these equations amount to comparing the internal signal r to a criterion θ, and decide accordingly (i.e., if $r \geq \theta$, chose $s > 0$, otherwise chose $s < 0$). The criterion for the first decision, θ_1, is given by

$$c_1^+\left(\theta_1\right) = \frac{1}{2} \quad \Rightarrow \quad \operatorname{erf}\left(\frac{\theta_1}{\sigma\sqrt{2}}\right) = 0 \tag{11}$$

which is satisfied for $\theta_1 = 0$. Similarly, the criterion for the second decision, θ_2, is given by

$$c_2^+(\theta_2) = \frac{1}{2} \quad \Rightarrow \quad \mathrm{erf}\left(\frac{\theta_2}{\sigma\sqrt{2}}\right) = 1 - 2c_1 \tag{12}$$

which indicates that θ_2 is a function of c_1 (see Main text, Figure 2B).

Non-Bayesian observer

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 with the internal observation r. These non-Bayesian models perform similarly to the Bayesian model for the first decision, that is when $r \geq 0$ they chose $s > 0$, otherwise $s < 0$. However, they cannot estimate a full probability distribution over the values of s, and therefore can assess confidence only by comparing the internal response r (which can be described as a point-estimate) to a set of one or more fixed criteria (Main text, Figure 2). In the case of a single confidence criterion, the non-Bayesian observer is confident in the response when the internal signal exceed the confidence criterion, and uncertain otherwise. When confident about the first response, he shifts the decision criterion for the second decision by a fixed amount (thereby increasing the probability of choosing $s_2 > 0$). If only one criterion is used, then the model has 2 discrete confidence levels (e.g., uncertain vs confident). In such a model, if the confidence criterion is w_1, the probability of the observer being confident about his first decision, after having responded 'right' (+), is

$$p\left(\text{confident}|s_1, +\right) = \frac{1 - \Phi\left(\frac{s_1 - w_1}{\sigma}\right)}{1 - \Phi\left(\frac{-w_1}{\sigma}\right)} \tag{13}$$

Note that this probability do not denote the confidence of the observer about his choice, which instead is assumed here to be a discrete binary state. The probability of the observer reporting $s_2 > 0$ in the second decision is then

$$\begin{aligned} p\left(+|s_2\right) = {} & p\left(\text{confident}|s_1, +\right)\Phi\left(\frac{s_2 - \theta_2}{\sigma}\right) \\ & + \left[1 - p\left(\text{confident}|s_1, +\right)\right]\Phi\left(\frac{s_2}{\sigma}\right) \end{aligned} \tag{14}$$

where θ_2 is the shift in criterion for the second decision applied by the observer when he is confident in his first decision. w_1 and θ_1 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. In our analysis we considered models with 2, 3, 4 discrete levels of confidence, which had 2,4 and 6 free parameters, respectively. The parameters w_1, w_2, w_3 and θ_2, θ_3, θ_4 were constrained so that $0 \leq w_1 \leq w_2 \leq w_3$ and $0 \geq \theta_2 \geq \theta_3 \geq \theta_4$.

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 his internal signals, but bases his 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 $p\left(s|r\right)$. 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 small number of samples [24]. Here we show, and confirm by simulation, that the sampling-based approximation of the Bayesian observer will display a systematic over-confidence bias, that is in the opposite direction with respect to what we found for most of our subjects, and is thus inconsistent with our behavioral results.

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 probability of overestimation, that is when $\hat{p} > p$, or underestimation, $\hat{p} < p$, depends in a complex way on both the probability p and the sample size n [28]. This is however for estimating a single fixed probability p. What would be instead the predominant bias, over many 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 predominant bias is one of over-estimation.

For a given n and p the probabilities of over- and under- estimation are

$$p\left(\hat{p} > p|n,p\right) = \sum_{k=\lfloor np \rfloor+1}^{n} \binom{n}{k} p^k (1-p)^{n-k} \tag{15}$$

$$p\left(\hat{p} < p|n,p\right) = \sum_{k=1}^{\lceil np \rceil+1} \binom{n}{k} p^k (1-p)^{n-k} \tag{16}$$

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 [28] we consider the difference between these two, denoted as *probability estimation bias*

$$\Delta p = p\left(\hat{p} > p|n,p\right) - p\left(\hat{p} < p|n,p\right) \tag{17}$$

When Δp is positive, it indicates that the probability p is more likely to be overestimated than underestimated, and viceversa 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 $[0.5, 1]$)

$$\mathbf{E}\left[\Delta p\right] = \int_{0.5}^{1} \frac{\Delta p}{0.5}\, dp \tag{18}$$

A positive value of the expected probability estimation bias (that is $\mathbf{E}\left[\Delta p\right] > 0$) indicates that, on average, the probabilities in this interval are more frequently overestimated rather than underestimated. This integral can be evaluated numerically, and in Figure S1A 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.

Fixed-n Bayesian sampler

Here we confirm by simulation the intuition developed in the previous paragraph. 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 favor of one of the options), these have been shown elsewhere to result in very similar performances as the fixed-n policy [30].

We start by providing the mathematical details of the model. Similarly to the previous cases, we assume that the observer has only access to r_1, a corrupted version of the stimulus s, $r_1 = s_1 + \eta$, where η is Gaussian noise with variance σ^2. In the first decision the prior is flat and, taking the limit of the stimuli range $R \to \infty$, the posterior distribution $p\left(s|r\right)$ results in a Gaussian distribution centered on the internal observation r_1. The probability that a sample from this distribution is above 0 (the criterion for the first decision) is therefore given by

$$p\left(x_n > 0|r_1\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{r_1}{\sigma\sqrt{2}}\right)\right] = \Phi\left(\frac{r_1}{\sigma}\right) \tag{19}$$

The conditional probability that the observer chooses (+) after having observed 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 is given by

$$p\left(+|r_1\right) = \sum_{k=\lceil \frac{n}{2} \rceil}^{n} \binom{n}{k} p\left(x_i > 0|r_1\right)^k \left[1 - p\left(x_i > 0|r_1\right)\right]^{n-k} \tag{20}$$

where $p\left(x_i > 0|r\right)$ is the probability that a single sample x_i is above 0. In other words the observer respond (+) when the majority of samples is above 0. The observer's confidence in his decision, c, is given by the proportion of samples in favor of the choice made

$$c = \max\left[c^+ \,,\, 1 - c^+\right] \tag{21}$$

where

$$c^+ = \frac{1}{n}\sum_{i=1}^{n} 1_{x_i > 0} \tag{22}$$

The probability that the observer chooses (+) with respect to s_1 can be calculated by integrating the probability $p\left(+|r_1\right)$ over all possible values of r_1

$$p\left(+|s_1\right) = \sum_{k=\lceil \frac{n}{2} \rceil}^{n} \binom{n}{k} \int_{-\infty}^{\infty} \Phi\left(\frac{r_1}{\sigma}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma}\right)\right]^{n-k} p\left(r|s_1\right) dr_1 \tag{23}$$

where

$$p\left(r_1|s_1\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(r_1-s_1)^2}{2\sigma^2}} \tag{24}$$

is the likelihood function that gives the probability of the internal observation r_1 given the stimulus s_1.

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

$$\theta = \sigma\sqrt{2} \cdot \mathrm{erf}^{-1}\left(1 - 2c\right) \tag{25}$$

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

$$p\left(c_j|s_1,+\right) = \binom{n}{k} \int\limits_{-\infty}^{\infty} \Phi\left(\frac{r_1}{\sigma}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma}\right)\right]^{n-k} p\left(r_1|s_1\right) dr_1 \tag{26}$$

where $j \in \left(1, ..., \lceil \frac{n+1}{2} \rceil\right)$ is an index linked to the number of samples above the criterion, k, according to $k = \lfloor \frac{n-1}{2} \rfloor + j$.

Conversely, if the observer's first decision was instead $(-)$, then the probability of the confidence level c_j is

$$p\left(c_j|s_1,-\right) = \binom{n}{k} \int\limits_{-\infty}^{\infty} \left\{1 - \Phi\left(\frac{r_1}{\sigma}\right)^k \left[1 - \Phi\left(\frac{r_1}{\sigma}\right)\right]^{n-k}\right\} p\left(r_1|s_1\right) dr_1 \tag{27}$$

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

A level of confidence c_j would result in a shift in decision criterion θ_j, calculated according to equation 25. The probability of choosing (+) in the second decision is

$$p\left(+|s_2,s_1,c_j\right) = \sum_{k=\lceil \frac{n}{2} \rceil}^{n} \binom{n}{k} \int\limits_{-\infty}^{\infty} \Phi\left(\frac{r_2-\theta_j}{\sigma}\right)^k \left[1 - \Phi\left(\frac{r_2-\theta_j}{\sigma}\right)\right]^{n-k} p\left(r_2|s_2\right) dr_2 \tag{28}$$

Taking everything together, the probability of the observer choosing (+) in the second decision conditioned on s_2 and s_1 (and on the first decision +/-, omitted here for simplicity) can be calculated as

$$p\left(+|s_2,s_1\right) = \sum_{j=1}^{\lceil \frac{n+1}{2} \rceil} p\left(c_j|s_1\right) \cdot p\left(+|s_2,s_1,c_j\right) \tag{29}$$

Simulation

We simulated the model for values of n ranging from 2 to 9. In order to compare the model with the full Bayesian observer, the value 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 5000 iterations we: (1) generated a random set of stimuli for 500 trials; (2) simulated the Bayesian model on those trials; (3) estimated σ for the sampling models based on the set of first responses produced by the Bayesian model (this was done using maximum likelihood estimation and equation 23); (4) simulated the sampling model. The average values of σ obtained are shown in table S1.

This approach ensured that all models resulted in similar proportion of first correct decisions, see figure S1B. The proportion of responses 'right' (+) in the second decision are plotted in Figure S1C. As expected, they show a pattern of marked over-confidence: all sampling models tended to respond + more often than the optimal model, despite similar accuracy in the first decision. The bias is 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 showed on average an under-confidence bias (see Main text, Results).

Table S1. Estimated values of σ (and their standard errors) that result in similar performance as the Bayesian model with $\sigma = 1$.

n. of samples	σ	se
2	0.74	0.06
3	0.83	0.06
4	0.84	0.06
5	0.88	0.07
6	0.89	0.07
7	0.91	0.07
8	0.91	0.07
9	0.93	0.07

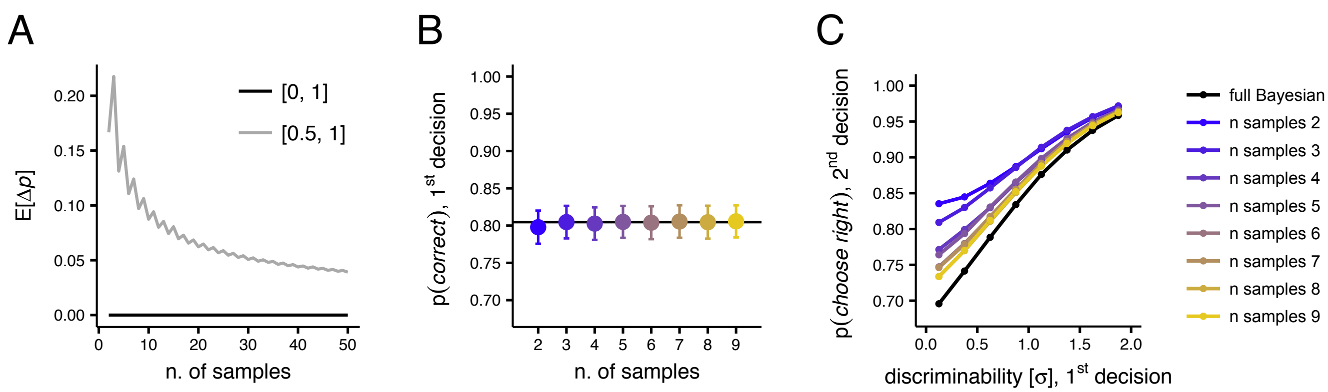

Figure S1. Sampling-based approximation of the Bayesian model. Expected probability estimation bias Δp plotted as a function of the number of samples when the probability p varies randomly (uniformly) either in the range $[0.5, 1]$, plotted in grey, or in the range $[0, 1]$ (**A**). It can be seen that 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). Proportion of correct first decision in the sampling model as a function of the number of samples n; the horizontal black line indicates the average performance of the full Bayesian model; error bars represents SEM (**B**). Proportions of responses "right" (+) in the second decision as a function of the difficulty of the first decision, as predicted by the full Bayesian model (black line) and the fixed-n Bayesian sampler model (**C**).

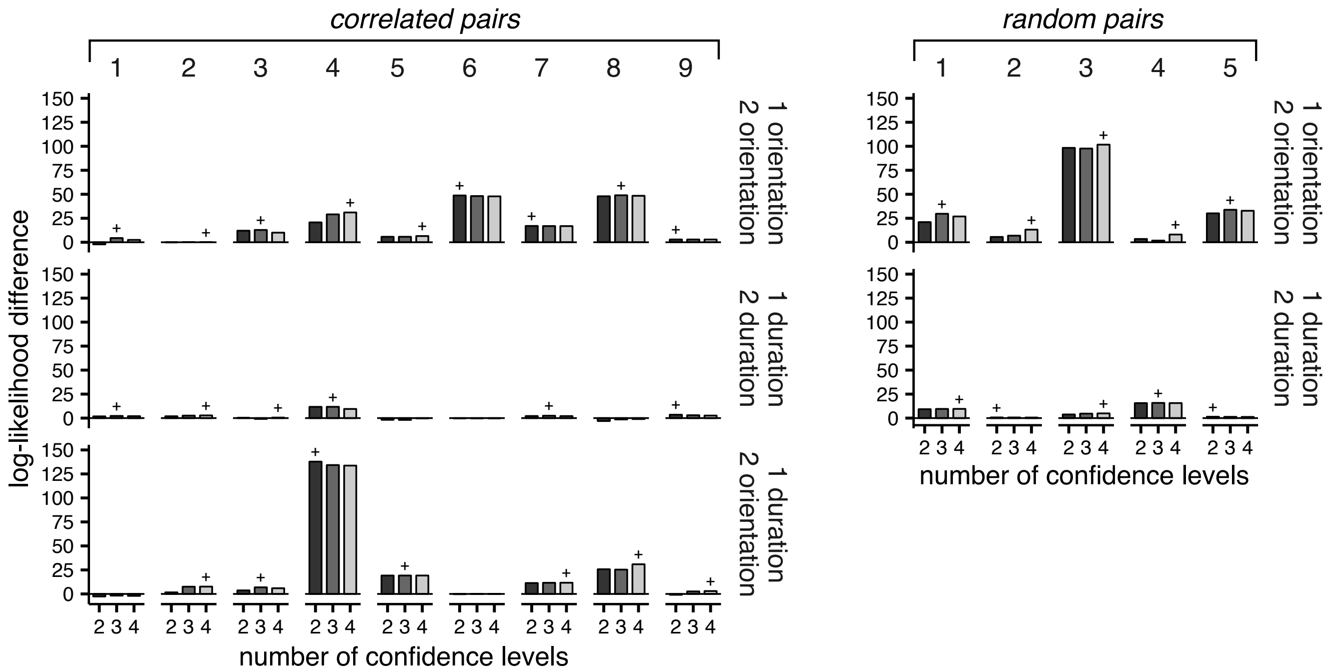

Figure S2. Cross-validation results for individual participants. For clarity, we marked with a + symbol the models with highest cross-validated likelihood for each participant and condition (the symbol is absent for conditions where the Bayesian model obtained the highest likelihood;).

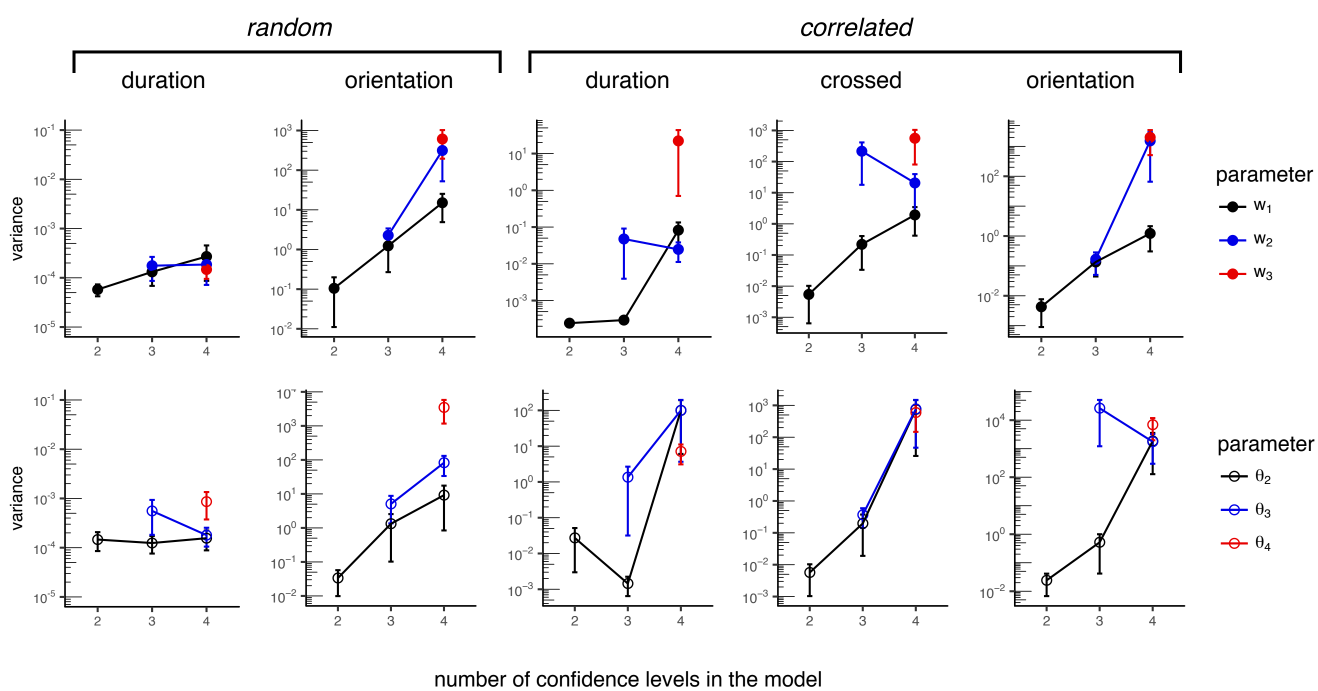

Figure S3. Variance of the hold-one-out estimates of the parameters of the non-Bayesian models, averaged across participants. Error bars represents bootstrapped standard errors. Despite similar predictive performance, the variance of the estimated parameters increase drastically in the model with more than 2 confidence levels, suggesting overfitting.

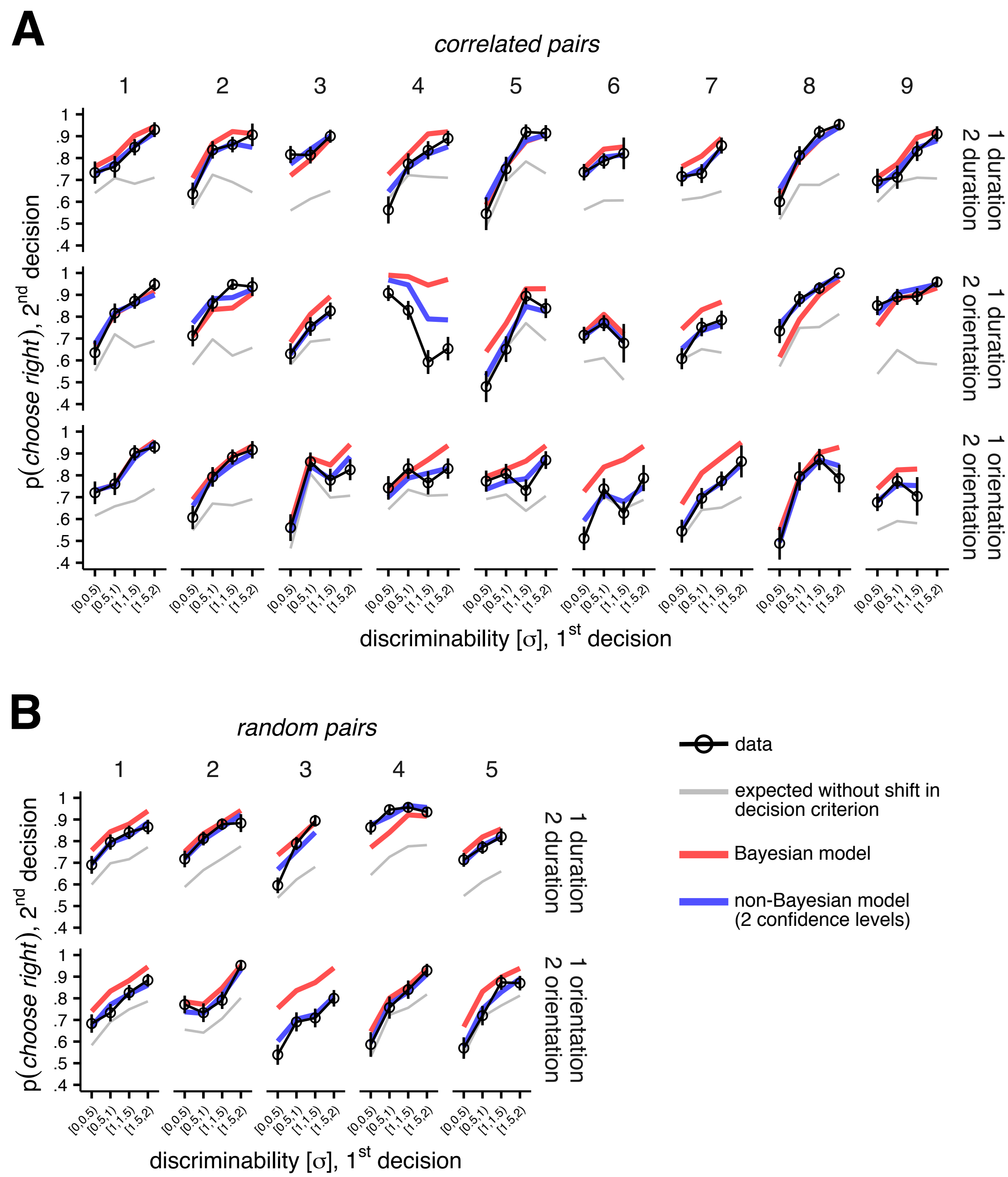

Figure S4. Comparisons of individual participants' data and model predictions.

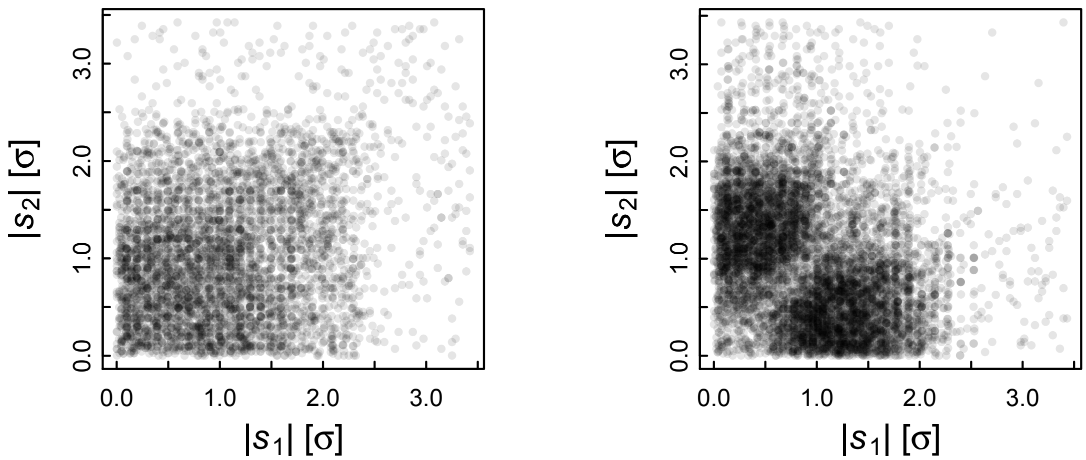

Figure S5. Distribution of absolute differences between stimuli (in units of noise, σ),in the random-pairs experiment, on the left and in the correlated-pairs experiment, on the right (both plots represents the stimuli for all participant and conditions of the respective experiments). It can be seen that there is a negative correlation in the correlated-pairs experiment, so that for example easy first decisions ($|s_1| > 1\sigma$), are more often associate with difficult second decisions ($|s1| < 1\sigma$). Note that while we set the range of the stimuli to be within 2JNDs as measured in the pre-test, the stimuli are plotted here in units of internal noise that were estimated taking into account all experimental trials; the variability of the range as represented here reflect the variability of the initial JND estimate.

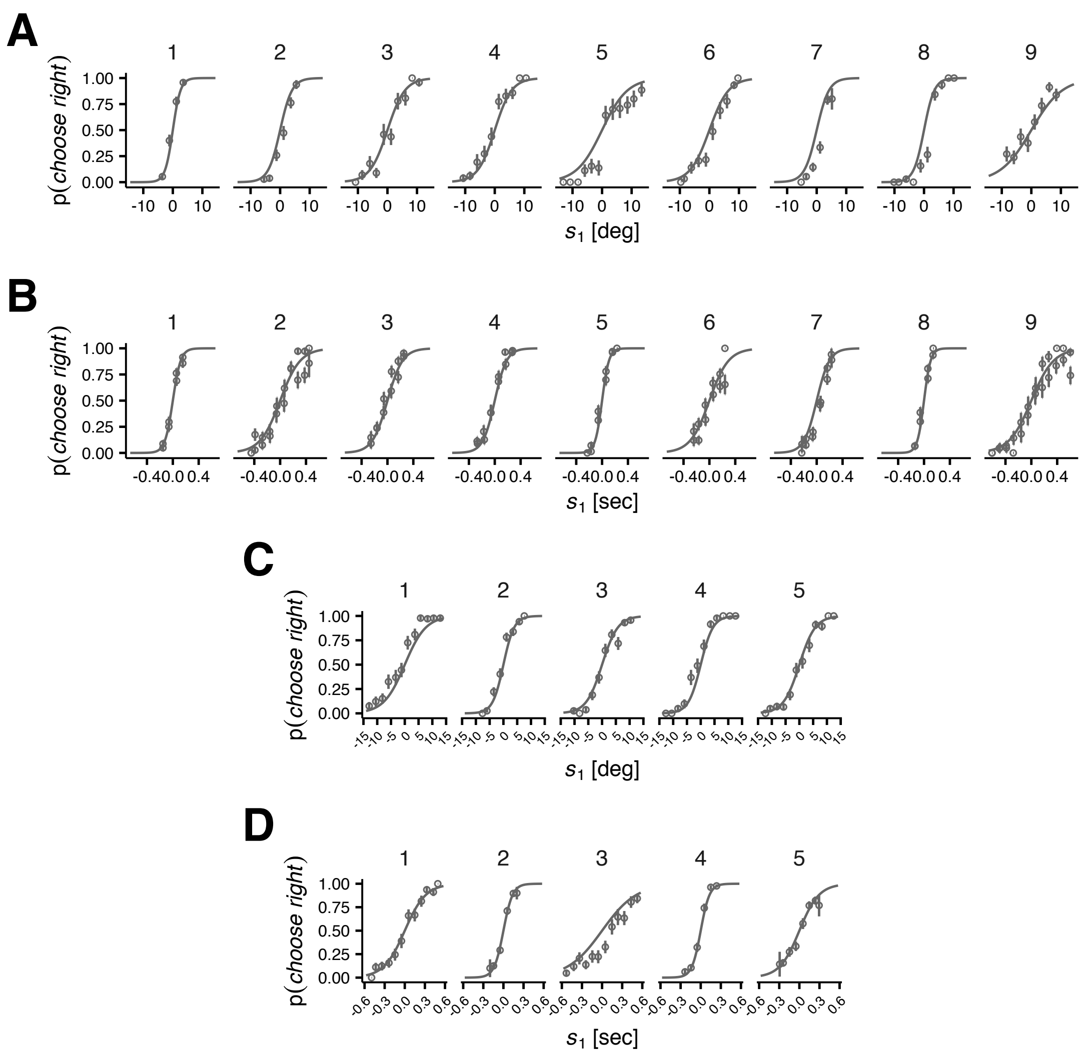

Figure S6. The psychometric functions for the first decisions,which were used to estimated, for each participant, the standard deviation of the noise (σ). (**A**) Orientation task, correlated pairs. (**B**) Duration task, correlated pairs. (**C**) Orientation task, random pairs. (**D**) Duration task, random pairs