LETTERS

# Neural correlates, computation and behavioural impact of decision confidence

Adam Kepecs[1], Naoshige Uchida[1,2], Hatim A. Zariwala[1,3] & Zachary F. Mainen[1,4]

**Humans and other animals must often make decisions on the basis of imperfect evidence[1,2]. Statisticians use measures such as *P* values to assign degrees of confidence to propositions, but little is known about how the brain computes confidence estimates about decisions. We explored this issue using behavioural analysis and neural recordings in rats in combination with computational modelling. Subjects were trained to perform an odour categorization task that allowed decision confidence to be manipulated by varying the distance of the test stimulus to the category boundary. To understand how confidence could be computed along with the choice itself, using standard models of decision-making[3–6], we defined a simple measure that quantified the quality of the evidence contributing to a particular decision. Here we show that the firing rates of many single neurons in the orbitofrontal cortex match closely to the predictions of confidence models and cannot be readily explained by alternative mechanisms, such as learning stimulus–outcome associations[7–10]. Moreover, when tested using a delayed reward version of the task, we found that rats' willingness to wait for rewards increased with confidence, as predicted by the theoretical model. These results indicate that confidence estimates, previously suggested to require 'metacognition'[11,12] and conscious awareness[13,14], are available even in the rodent brain, can be computed with relatively simple operations, and can drive adaptive behaviour. We suggest that confidence estimation may be a fundamental and ubiquitous component of decision-making.**

Rats were trained on a two choice odour mixture categorization task (Fig. 1a). On each trial, a binary mixture of two pure odorants (A, caproic acid; B, 1-hexanol) was delivered at one of several concentration ratios (Fig. 1b), which were randomly interleaved from trial-to-trial[15]. Choices were rewarded at the left choice port for mixtures A/B > 50/50 and at the right choice port for A/B < 50/50 (Fig. 1b). By varying the distance of the stimulus to the category boundary (50/50) we could vary the difficulty of the decision (Fig. 1c, d). Although the reward contingencies were deterministic, subjects experienced varying degrees of decision uncertainty due to imperfect perception of stimuli and/or knowledge of the category boundary.

To explore the neural correlates of decision confidence, we recorded single neuron activity in the orbitofrontal cortex (OFC; Supplementary Fig. 1), a brain region implicated in decision-making under uncertainty[16–20]. We reasoned that neural activity related to the subject's confidence in the outcome of a choice should occur while the subject is anticipating the trial outcome, and therefore focused our analysis on this delay period (Fig. 2a). The firing rates of many OFC neurons were modulated by stimulus difficulty during the anticipation period. Figure 2b, c shows the activity of a neuron that fired more intensely following more difficult decisions. By replotting the same data as a function of the choice accuracy associated with each stimulus type, it can be seen that this neuron fired more vigorously when the likelihood of an upcoming reward was lower (Fig. 2d). A large fraction of OFC neurons, like this example, fired more intensely for stimuli closer to the category boundary (120/563 at $P < 0.05$, Wilcoxon signed-rank test). A smaller fraction (66/563) showed the opposite tuning, firing at a higher intensity for easy stimuli, those far from the category boundary (Fig. 2e, f).

The observed modulation of firing rate by stimulus difficulty is consistent with previous findings that the response of many OFC neurons correlates with the expected values associated with reward predictive cues[7–10]. Surprisingly, however, when we compared correct and incorrect choices for the same stimulus (for example, the 68/32 mixture), we found that many neurons showed different firing rates even before the outcome was delivered. Figure 3a, b shows an example of a neuron that tended to fire more when the rat had

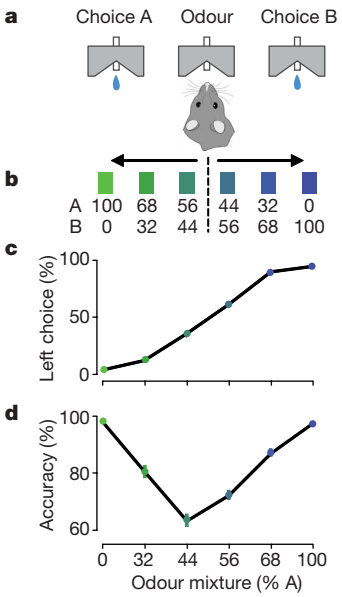


**Figure 1 | Odour mixture categorization task. a**, Schematic of the behavioural paradigm. To initiate a trial, the rat enters the central odour port and after a pseudorandom delay of 0.2–0.5 s a mixture of odours is delivered. Rats respond by moving to the left or right choice port, where a drop of water is delivered after a 0.3–2 s waiting period for correct choices. **b**, Stimulus design. **c**, Performance of one rat discriminating between mixtures of caproic acid (A) and 1-hexanol (B) in a single session. Error bars (s.e.m.) are hidden by markers. Colours are used to represent odour mixtures, with different blue and green blends representing different odour mixture ratios. **d**, Choice accuracy as a function of odour mixture. Data across three rats are plotted as mean ± s.e.m.

[1]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. [2]Department of Molecular and Cellular Biology and Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA. [3]Allen Institute for Brain Science, Seattle, Washington 98103, USA. [4]Champalimaud Neuroscience Programme, Instituto Gulbenkian de Ciência, 2780-901 Oeiras, Portugal.

committed an error than when it was correct, despite the fact that the outcome was not yet revealed to the subject. The same phenomenon could also be seen as a difference in the average behavioural accuracy when the neuron was firing at high compared to low rates (see Supplementary Fig. 2a). Similar to this example, a large fraction of neurons fired at a higher rate in incorrect trials ('error trials') compared to correct trials within a given stimulus type (46/317 neurons for 56/44 mixtures and 86/563 for 68/32 mixtures at $P < 0.05$, permutation test, Fig. 3d–f; Supplementary Figs 2b and 3c). Interestingly, for easier stimuli the difference in firing rates between correct and error trials was larger (Fig. 3; Supplementary Fig. 3d). A second, smaller population of neurons (21/317 for 56/44 mixtures and 50/563 for 68/32 mixtures at $P < 0.05$, permutation test) had an analogous pattern of activity, but fired more in anticipation of correct rather than incorrect outcomes (Supplementary Fig. 4).

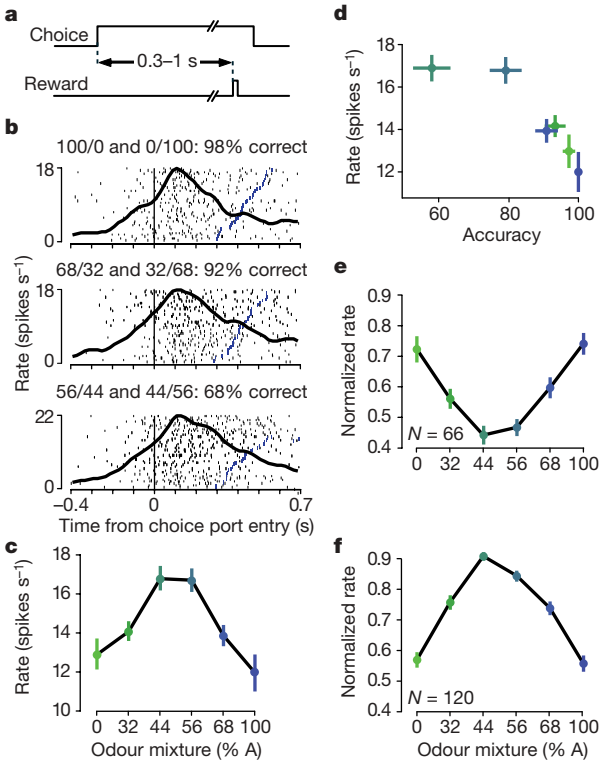


**Figure 2 | Graded representation of stimulus difficulty in orbitofrontal cortex.** **a**, Timing of outcome anticipation period. Entry into the choice port is recorded using the interruption of the photo-beams within each port. The delivery of water is pseudo-randomly delayed, with the earliest onset varying between 0.3 s and 1 s and the latest offset from 0.8 s to 2 s after entry, according to a uniform distribution with varying parameters in each session. The anticipation period ends at the first possible time of reward delivery, and thus ranges from 0.3 s to 1 s across sessions. Firing rates are calculated either during the initial 0.4 s of the anticipation period or the entire period if it was shorter. **b**, Activity of an example neuronal unit. Raster plots represent neural activity, with each row corresponding to a single trial and each tick mark to a spike. Forty trials are shown in each plot with the post-stimulus time histogram (PSTH) overlaid (smoothed with a Gaussian filter, s.d. = 25 ms). Neural activity is aligned to the timing of entry into the choice port. Blue ticks represent the time of reward delivery. Trials for different stimuli were interleaved in the sessions but grouped into different panels according to stimulus difficulty, with stimuli and performance indicated above. **c**, Mean firing rate of cell in **b** as a function of stimulus identity. Rates are calculated during the outcome anticipation period (0.3 s window beginning at the time of entry into the choice port). Error bars, s.e.m. across trials. **d**, Mean firing rate as function of mean accuracy grouped by stimulus identity. **e**, Mean-normalized firing rate as a function of stimulus identity for the population of neurons with higher firing rates in error trials (Wilcoxon test, $P < 0.05$). **f**, As **e** but for the population of neurons with higher firing rates in correct trials (Wilcoxon test, $P < 0.05$).

These firing patterns appear paradoxical for a prediction made on the basis of overall stimulus–outcome associations. However, reward predictions may be generated by a dynamic learning process based on recent reinforcement history[21–23]. To test this idea, we used a more powerful multiple linear regression model to try to predict the firing rate of a given trial based on the history of recent reward outcomes and other externally observable variables (the stimulus and choice direction). This analysis revealed that although a subset of OFC neurons do carry information about past trial events, these account for a relatively small fraction of the firing rate variance compared to what can be explained by the anticipated current trial outcome (Supplementary Fig. 5; for details see Methods). Therefore, the signals we observed in OFC neurons could not be readily explained as reward expectancy based on either a simple average stimulus–reward association or more complex predictions based on reinforcement history.

In principle, the probability of a correct trial outcome could be estimated based on a subjective measure of confidence about the decision. We hypothesized that a useful confidence metric could be calculated by measuring the reliability and consistency of the values

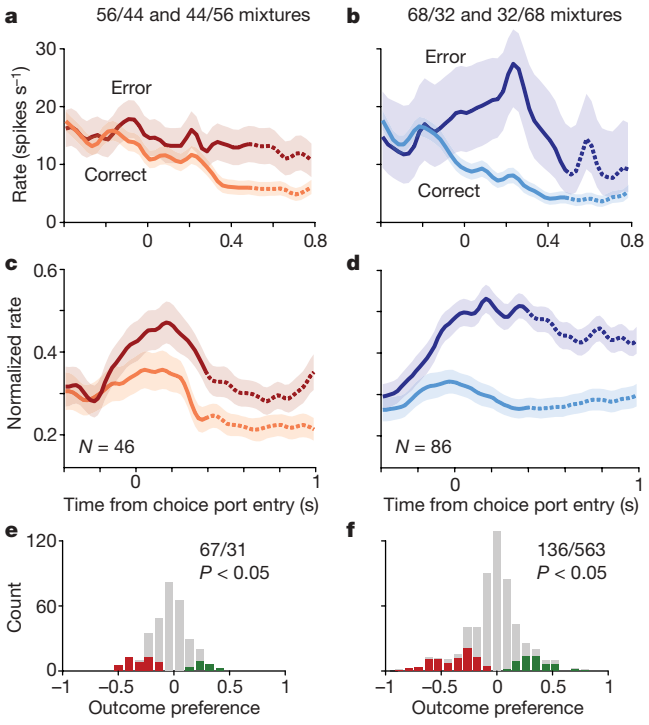


**Figure 3 | Orbitofrontal neurons anticipate trial outcome.** **a**, **b**, Firing rate of a single neuron aligned to the time of entry into the choice port. Trials are grouped by stimulus difficulty (**a**, 44/56 and 56/44 odour mixture ratio; **b**, 32/68 and 68/32) and trial outcome (correct, orange and cyan; error, red and blue). Shading represents s.e.m.; note there are few 68/32 error trials. Only activity occurring before the onset of water delivery and choice port exit is averaged into the PSTH. After the outcome anticipation period (0.5 s in this session) the PSTH curves are dashed, signifying a time period when in some trials rats experienced reward delivery, although post-reward firing is never actually included. Note that the separation between correct and error trials begins before entry into the choice port but after the animal leaves the odour sampling port. **c**, **d**, Mean-normalized firing of negative outcome selective neurons (those with increased firing rate in error trials during the anticipation period) is plotted the same way as **a**, **b**. Shading represents s.e.m. across neurons. Dashed curves as in **a**, **b**. **e**, **f**, Outcome preference for the population of OFC cells during the outcome anticipation period. Outcome preference is calculated using ROC analysis (see Methods). Colour bars represent significant selectivity (permutation test, $P < 0.05$); red indicates neurons with increased firing rates in incorrect ('error') trials (negative outcome selectivity, 46/317 neurons); green indicates neurons with increased firing rates in correct trials (positive outcome selectivity, 22/317 neurons); grey bars, not significant.

of the internal variables that contributed to the decision. To explore this idea, we constructed a simple model for the categorization task based on the comparison of the perceived stimulus value and the recalled category boundary (Fig. 4a; see Methods for details). In this model, the choice depends on whether the stimulus sample, $s_i$, is smaller or larger than the category boundary, $b_i$. This comparison yielded an average choice function similar to that observed behaviourally (Fig. 4b; compare Fig. 1c). To estimate the confidence about this choice, we propose to measure the quality of the evidence in this model using the distance between the stimulus and memory samples, $d_i = |s_i - b_i|$; the larger the distance, the more reliable should be the decision. We found that after a simple transformation, $d_i$ can indeed provide a veridical prediction of the likelihood of a successful outcome, 'decision confidence', $\delta_i = f(d_i)$, or the likelihood of a failure, 'decision uncertainty', $\sigma_i = 1 - \delta_i$ (Fig. 4c). Similar algorithms can also yield useful confidence estimates in other decision models. For example, in a two-alternative 'race' model, an instance of a class of models based on the accumulation of evidence[4–6], decision confidence can be calculated from the difference between two decision variables at the time a decision is reached (Supplementary Fig. 6; Supplementary Information). These modelling results demonstrate that confidence estimates derived solely from the decision variables in the current trial can provide good estimates of the expected decision outcome across trials.

We next looked for specific predictions—patterns of firing rates—that would arise from theoretical confidence estimates. We noticed that, when plotted as a function of stimulus type and trial outcome, decision uncertainty, $\sigma_i$, shows a characteristic and somewhat counterintuitive pattern, namely opposing V-shaped curves for correct and error choices (Fig. 4d): (1) for correct choices, $\sigma_i$ decreases with distance from the category boundary; (2) for a given stimulus, error trials are associated with higher $\sigma_i$ than correct trials; (3) the difference in $\sigma_i$ for error and correct trials increases as the stimulus becomes easier. These patterns are robust to model details and do not depend on the relative contributions of stimulus versus memory noise or on the precise choice of the transform function, $f$ (Supplementary Fig. 7). In addition, the same pattern of confidence estimates are produced by decision models based on integration of evidence (Supplementary Fig. 6).

The dependence of OFC neuronal activity on stimulus type and trial outcome closely matched the predictions of confidence estimates derived from decision models (Fig. 4e–h). First, individual OFC neurons showed the predicted dependence on the distance of the stimulus to the category boundary as well as the predicted difference between correct and error trials (Fig. 4e). A similar pattern held at the population level (Fig. 4g, 133/563 negatively-tuned neurons, all stimuli pooled at $P < 0.05$, permutation test; see also Supplementary Figs 3, 8). These patterns were qualitatively different from those expected from left/right modulation of stimulus selectivity (Supplementary Fig. 3). Second, the probability of correct trial outcome varied with the firing rate of individual neurons (Fig. 4f), and at the population level (Fig. 4h), as predicted (Fig. 4c). This analysis also showed that the highest firing rates were associated with near chance performance (50% reward probability), as expected if these neurons signalled lack of confidence rather than incorrect performance (0% reward probability; see Methods for details). The opposite patterns held for the positive outcome selective OFC population (105/563 neurons for all stimuli pooled at $P < 0.05$, permutation test; Supplementary Fig. 4).

It is possible for the experimenter observing OFC neurons to predict individual trial outcomes, but can rats use such information behaviourally? We tested the ability of rats to provide a behavioural report of confidence using a modified version of the task in which we encouraged rats to give up waiting for uncertain rewards by increasing the delay to reward delivery and permitting subjects to reinitiate a

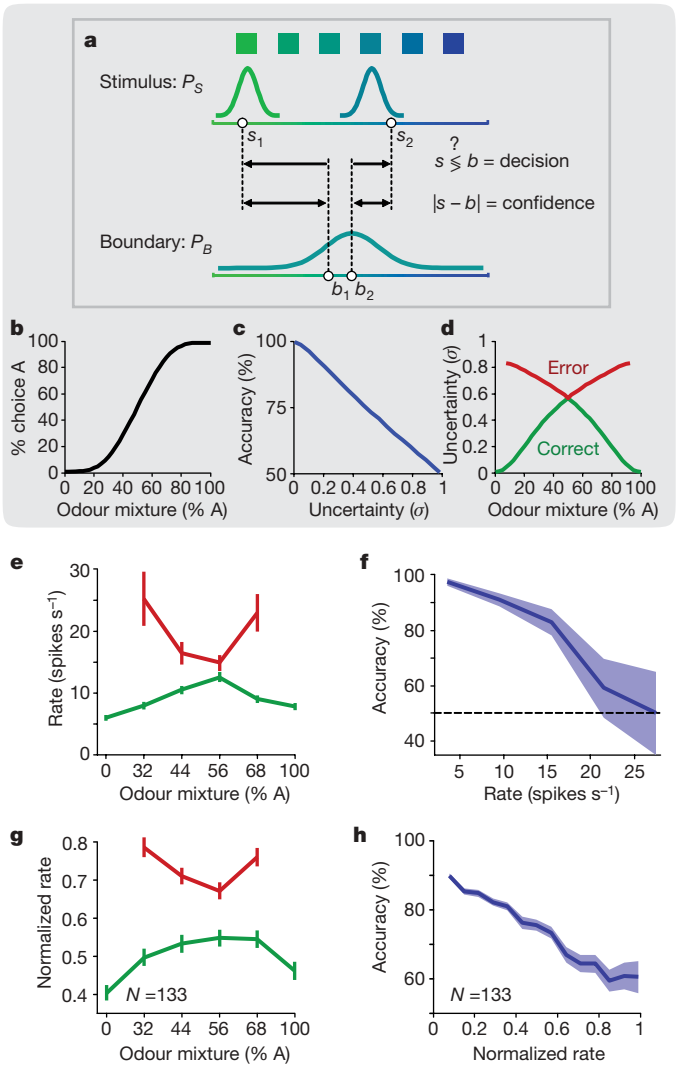


**Figure 4 | Confidence estimation in a decision model and by OFC neurons.** **a**, Schematic of a model for category decisions. Each odour mixture stimulus, as well as the memory for the category boundary, is encoded as a distribution of values. In each trial a stimulus, $s_i$, and memory of the boundary, $b_i$, are drawn from their respective distributions. A choice is calculated by comparing the two samples ($s_i < b_i$), and a confidence value is estimated by calculating their distance ($|s_i - b_i|$). Incorrect choices result from noise, represented in the model by the width of the stimulus and category boundary distributions. See Methods for details. **b**, Example psychometric function of the model, replicating the high choice accuracy of rats for pure odours and decreased accuracy for mixtures near the imposed the category boundary. **c**, Mean accuracy of model choices as a function of decision uncertainty. The uncertainty estimate, $\sigma$, is transformed from the distance between the stimulus and boundary samples ($\sigma_i = 1 - \tanh(|s_i - b_i|)$), see Methods). **d**, Mean decision uncertainty estimates generated by the model as a function of stimulus and trial outcome. Note that the model (or a subject) has access only to a stimulus sample and not the stimulus type (for example, 56/44) (see Supplementary Information for an explanation of the pattern of uncertainty estimates.). **e**, Firing rate of an example neuron (same unit as Fig. 3a, b) during the outcome anticipation period as a function of odour stimulus and trial outcome. Error bars are s.e.m. across trials. **f**, Mean choice accuracy as a function of the firing rate for the same unit in **e**. Firing rates were binned and the mean accuracy was calculated for each range of firing rates. Error bars represent standard errors based on the binomial distribution of outcomes. **g**, Mean normalized firing rate of negative outcome selective population (negative outcome preference index across trials with all stimuli pooled at $P < 0.05$, permutation test) during the anticipation period. **h**, Mean accuracy as a function of the firing rate for the same neuron population as in **g**. Firing rates were binned for individual neurons and the mean accuracy was calculated for each range of firing rates. These curves were normalized to a maximal firing rate of 1 and averaged. Error bars represent s.e.m. across neurons.

trial (Fig. 5a). While waiting at the choice port, the decision whether to stay and wait for a possible reward or to go and reinitiate the trial could benefit from an estimate of the confidence in the original decision. Indeed, we found that rats preferentially aborted uncertain trials. Like the neural responses in OFC, these response patterns closely agreed with the predictions of the decision confidence model (Figs 5b, c and 4d). Therefore rats not only show a neural correlate of decision confidence but they can use such information in subsequent decisions to guide adaptive behaviour.

The patterns of neural activity and behaviour we observed suggest that when a decision is made the brain not only makes a choice but also generates an evaluation about the quality of evidence that contributed to the decision. We liken this to the way *P* values are assigned to statistical statements. Our interpretation of the data rests on two results: first, we defined a mechanism for computing confidence in simple decision models and showed that this produced a close fit to a non-trivial pattern of neural and behavioural data; second, we ruled out alternative models for the data, principally ones based on learning. Confidence estimates based on internal decision variables provide useful information that is not readily gained by observing the past relationships between externally observable stimulus, response and outcome variables. Intuitively, this is possible because the observable result of a decision, the choice, is only a partial distillation of the information entering the internal decision process. Computing decision confidence essentially requires calculating how 'close a call' was the choice or how well the evidence was in agreement. When decision 'noise' arises from sources internal to the brain, this process is inherently subjective (accessible only to the subject). More formally, decision confidence can be expressed as the variance measured across the set of decision variables contributing to a single trial (see Supplementary Information). Two different classes of decision model yielded very similar results, suggesting a degree of generality to our description. Nevertheless, it will be important to examine the properties of other methods for estimating confidence.

A variety of results suggests that a key function of OFC is to generate reward predictions based on stimulus–reward associations[7–10]. Our data support and extend this idea by showing that OFC neurons signal outcome predictions derived from a different source, specifically, from internal variables contributing to a perceptual decision on a given trial. In addition to predicting expected rewards, OFC has also been implicated in signalling outcome risk or variance[16–20]. Because in a two-alternative psychophysical decision task the expected reward and its variance are closely related, our data are consistent with both functions and further experiments will be needed to distinguish between these alternatives. It also remains to be determined whether OFC neurons drive the reinitiation behaviour displayed by rats (Fig. 5) or other behaviours contingent on confidence estimates. Indeed, decision confidence signals could be useful for a variety of functions, including controlling exploration[24,25], modulating learning rates[26] and focusing attention[27,28].

Bayesian theory suggests that uncertainty estimates must be incorporated into neural computations for optimal behaviour[29]. Humans and other primates clearly have the ability to assess and act on the degree of uncertainty or confidence in their beliefs about the world[1,11,30], but it has been argued that this might be a sophisticated 'metacognitive' capacity requiring self-awareness[13,14] and a neural architecture specific to primates[11]. Our results show that rodents possess the ability to act on their degree of belief in a decision[12] and demonstrate that estimating the confidence in a choice is little more complex than calculating the choice itself. It is likely that confidence estimates for memories or other beliefs[11,30] could be derived in an analogous fashion. We suggest that the computation of subjective confidence may be a core component of decision-making that, like subjective value signals[7–10,21–23], is important to a wide range of behaviours and their neural substrates.

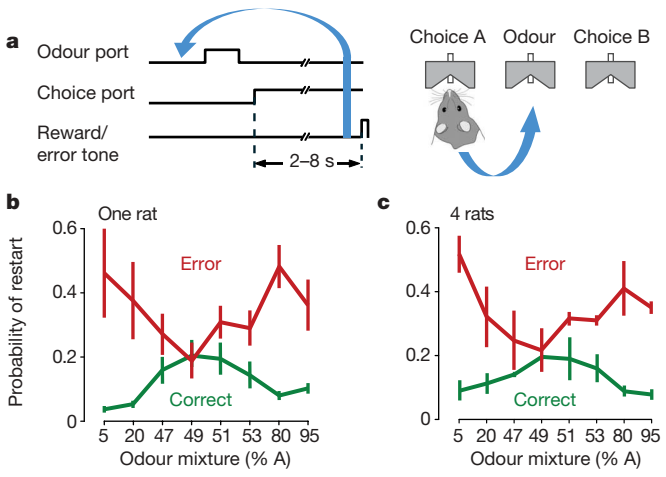


**Figure 5 | Behavioural use of decision confidence.** **a**, Schematic of the reinitiation task. Reward delivery was pseudo-randomly delayed between 2 and 8 s (uniform distribution) after the rat's choice was registered. Incorrect choices were signalled with an error tone delivered at the end of the 8 s delay. There was a minimum delay of 2 s from the time of the choice before rats could initiate a new trial. **b**, Probability of reinitiation for a single rat plotted as a function of odour stimulus and trial outcome. Error bars represent s.e.m. across trials. Entry into the odour port within 2 s of aborting was considered a reinitiation. **c**, Mean probability of reinitiation for 4 rats as a function of odour stimulus and trial outcome. Error bars represent s.e.m. across rats.

## METHODS SUMMARY

Male Long-Evans hooded rats were trained to perform an odour categorization task for water reward. Behavioural testing was controlled by custom software written in Matlab (Mathworks) using data acquisition hardware (National Instruments) to record the port signals and control the valves of the olfactometer and water-delivery[15].

Rats were implanted with custom-made microdrives in the left orbitofrontal cortex (3.5 mm anterior to bregma and 2.5 mm lateral to midline). Extracellular recordings were obtained with six independently movable tetrodes using the Cheetah system (Neuralynx) and single units were isolated by manually clustering spike features with MClust (A. D. Redish).

We focused our analysis on the 'reward anticipation period' while rats remained at one of the choice ports. This excluded spikes that occurred during or after water valve actuation on correct trials; on error trials, no feedback was present. To determine how well neural activity predicted the upcoming outcome (reward/no reward), we used receiver operating characteristics (ROC) analysis to calculate an outcome preference index (OP) that measures how well an ideal observer can predict the outcome from the knowledge of the firing rate from trial to trial. This index varies from −1 to 1 with the sign denoting whether a neuron fires more for rewarded (correct, +) or unrewarded (error, −) decisions:

$$\mathrm{OP} = 2(\mathrm{ROC}_{\mathrm{area}} - 0.5); \quad \mathrm{ROC}_{\mathrm{area}} = \int_0^\infty P(f_{\mathrm{correct}} = f)P(f_{\mathrm{error}} < f)\mathrm{d}f \quad \text{where} \quad f_{\mathrm{correct}}$$

and $f_{\mathrm{error}}$ refer to the distribution of firing rates during the reward anticipation period in correct and error trials respectively. Statistical significance was evaluated using a permutation test, where trial order was pseudo-randomly shuffled 200 times to yield a *P* value.

All procedures involving animals were carried out in accordance with National Institutes of Health standards and were approved by the Cold Spring Harbor Laboratory Institutional Animal Care and Use Committee.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 28 February; accepted 26 June 2008.**
**Published online 10 August 2008.**

1. Kahneman, D., Slovic, P. & Tversky, A. *Judgment under Uncertainty: Heuristics and Biases* (Cambridge Univ. Press, 1982).
2. Glimcher, P. W. *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (MIT Press, 2003).
3. Kim, J. N. & Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neurosci.* **2**, 176–185 (1999).

4. Bogacz, R. *et al.* The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113,** 700–765 (2006).
5. Mazurek, M. E., Roitman, J. D., Ditterich, J. & Shadlen, M. N. A role for neural integrators in perceptual decision making. *Cereb. Cortex* **13,** 1257–1269 (2003).
6. Ratcliff, R. & Smith, P. L. A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* **111,** 333–367 (2004).
7. Schoenbaum, G., Chiba, A. A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neurosci.* **1,** 155–159 (1998).
8. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398,** 704–708 (1999).
9. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441,** 223–226 (2006).
10. Wallis, J. D. Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci.* **30,** 31–56 (2007).
11. Smith, J. D., Shields, W. E. & Washburn, D. A. The comparative psychology of uncertainty monitoring and metacognition. *Behav. Brain Sci.* 26, 317–339 340–373 (2003).
12. Foote, A. L. & Crystal, J. D. Metacognition in the rat. *Curr. Biol.* **17,** 551–555 (2007).
13. Persaud, N., McLeod, P. & Cowey, A. Post-decision wagering objectively measures awareness. *Nature Neurosci.* **10,** 257–261 (2007).
14. Koch, C. & Preuschoff, K. Betting the house on consciousness. *Nature Neurosci.* **10,** 140–141 (2007).
15. Uchida, N. & Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nature Neurosci.* **6,** 1224–1229 (2003).
16. Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. Deciding advantageously before knowing the advantageous strategy. *Science* **275,** 1293–1295 (1997).
17. Critchley, H. D., Mathias, C. J. & Dolan, R. J. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* **29,** 537–545 (2001).
18. Grinband, J., Hirsch, J. & Ferrera, V. P. A neural representation of categorization uncertainty in the human brain. *Neuron* **49,** 757–763 (2006).
19. Hsu, M. *et al.* Neural systems responding to degrees of uncertainty in human decision-making. *Science* **310,** 1680–1683 (2005).
20. Tobler, P. N., O'Doherty, J. P., Dolan, R. J. & Schultz, W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* **97,** 1621–1632 (2007).
21. Barraclough, D. J., Conroy, M. L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neurosci.* **7,** 404–410 (2004).
22. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Matching behavior and the representation of value in the parietal cortex. *Science* **304,** 1782–1787 (2004).
23. Lau, B. & Glimcher, P. W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84,** 555–579 (2005).
24. Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton Univ. Press, 1986).
25. Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nature Neurosci.* **10,** 1214–1221 (2007).
26. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46,** 681–692 (2005).
27. Dayan, P., Kakade, S. & Montague, P. R. Learning and selective attention. *Nature Neurosci.* **3** (Suppl), 1218–1223 (2000).
28. Luck, S. J., Hillyard, S. A., Mouloua, M. & Hawkins, H. L. Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction? *J. Exp. Psychol. Hum. Percept. Perform.* **22,** 725–737 (1996).
29. Knill, D. C. & Pouget, A. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci.* **27,** 712–719 (2004).
30. Hampton, R. R. Rhesus monkeys know when they remember. *Proc. Natl Acad. Sci. USA* **98,** 5359–5362 (2001).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank J. Paton, A. Pouget, S. Raghavachari, G. Turner and members of the Mainen laboratory for comments on the manuscript. Support was provided by the National Institutes of Health (NIDCD) (Z.F.M.), the Center for the Neural Mechanisms of Cognition at Cold Spring Harbor Laboratory (Z.F.M.), and the Swartz Foundation (A.K., N.U., Z.F.M).

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.K. (kepecs@cshl.edu) or Z.F.M. (zmainen@igc.gulbenkian.pt).

## METHODS

Here we describe the behavioural and physiological methods used in this study and explain the analyses presented in the main text.

**Behavioural task.** The behavioural box contains a panel of three ports: the central port for odour delivery ('odour port'), and two ports on each side ('choice ports') for water delivery (Fig. 1a). Entry and exit from the ports was detected based on an infrared photo-beam located inside each port. Odours were mixed with pure air to produce a 1:20 dilution at a flow rate of 1 l $min^{-1}$ using a custom-built olfactometer[15].

Rats self-initiated each experimental trial by introducing their snout into a central port where odour was delivered (Fig. 1a). After a variable delay, drawn from a uniform random distribution of 0.2–0.5 s, a binary mixture of two pure odorants, caproic acid and 1-hexanol, was delivered at one of 4–6 concentration ratios (100/0, 68/32, 56/44, 44/56, 32/68, 0/100; Fig. 1b) in pseudorandom order within a session. After a variable odour sampling time up to 1 s, rats responded by withdrawing from the central port, which terminated the delivery of odour, and moved to the left or right choice port (Fig. 1a). Choices were rewarded according to the dominant component of the mixture, that is, at the left port for mixtures A/B < 50/50 and at the right port for A/B > 50/50 (Fig. 1b). We introduced a variable reward delay period after entry into the choice port. For correct choices, reward was delivered between at least 0.3 s after entry into the choice port and sometimes up to 2 s (in individual sessions the delays were uniformly distributed with the onset ranging from 0.3–0.8 s and the offset to 1–2 s). Outcome selectivity calculations used firing rates calculated over the first 0.4 s of the reward anticipation period. In a few sessions the reward anticipation was 0.3 s (e.g. Fig. 2c, d); in those sessions the entire reward anticipation period was used.

This task allowed us to control the distance of each stimulus to the category boundary and hence systematically manipulate the difficulty of individual categorization problems (Fig. 1d). Intuitively, this task is analogous to categorizing colours along a continuous spectrum (for example, blue/green, Fig. 1b). For colour blends in the middle, the answer depends on a semi-arbitrary convention of colour category boundaries. Similarly, our training protocol enforced the 50/50 odour category boundary, which is semi-arbitrary, as the pure odours do not have equal intensity.

**Reinitiation task.** In this version of the task, the delay to reward was increased to between 2 and 8 s (uniform random distribution). Errors were signalled with an auditory beep at 8 s and punished with an additional 4 s time-out. After a 2 s mandatory wait from the entry into a choice port and before water or auditory feedback was provided, subjects were allowed to abort trials by exiting the water port. Entry into the odour port within 2 s of aborting was considered as 'reinitiation'. The stimulus ensemble consisted of 75% easy (95/5, 80/20 mixtures: 92 ± 4% accuracy, s.e.m across rats) and 25% difficult (53/47, 51/49 mixtures: 55 ± 2% accuracy) stimuli so that rats could expect to encounter an easier stimulus after reinitiating a new trial. The expectation of a rat to receive reward by staying at the choice port should be proportional to its confidence about the first choice (Fig. 4d) while the expectation to receive reward by reinitiating a new trial should be fixed (because the new stimulus is not predictable). Therefore the relative value of reinitiating is predicted to increase as confidence drops, with approximately the same dependence on stimulus and outcome as given by the model (Fig. 4c). The exact value depends on the actual delays and the subject's temporal discounting function.

**Neural data collection and analysis.** Rats were implanted with custom-made microdrives in the left orbitofrontal cortex (3.5 mm anterior to bregma and 2.5 mm lateral to midline) as described previously[31] (Supplementary Fig. 1). Extracellular recordings were obtained using six independently adjustable tetrodes for recording. Electrodes were advanced each recording day to sample an independent population of cells across sessions. The placement of electrodes was estimated by depth and confirmed with histology. Neural and behavioural data were synchronized by acquiring time-stamps from the behavioural system along with the electrophysiological signals. Data analysis was performed using Matlab (Mathworks).

For Fig. 2e, f, confidence-modulated neurons were selected by performing a non-parametric, Wilcoxon signed-rank test on firing rates during the reward anticipation period for correct versus error trials. Neurons with significant ($P < 0.05$) firing rate differences were separated into two populations based on whether their mean firing rate was higher for correct or error trials. We then plotted the maximum normalized firing rate averaged for each neural population as a function of stimulus mixture ratio. We used this selection criterion because by not using information about the stimulus it does not impose a specific shape on the tuning curves. Other selection criteria, such as significant rate-accuracy correlations (for example, Fig. 2d), yielded similar results.

**Multiple linear regression analysis.** We considered the possibility that a prediction of upcoming trial outcome might be made on the basis of recent reward history[32–35] and other observable task variables. For example, if the average performance fluctuated due to changes in attention or motivation and OFC neurons tracked the recent history of trial outcomes, it could lead to a differential prediction of correct versus error trials when averaged over the entire session. In this scenario, outcome selectivity would arise because the present trial's expected outcome is correlated with the recent trials' outcomes. Although we did not observe prominent performance fluctuations, we wanted to test this and related possibilities directly. We used multiple linear regression in an attempt to predict the firing rate of a given trial based on the history of recent reward outcomes and experimental variables (stimulus type and choice direction). Specifically we fitted the firing rates during the reward anticipation period to the following model:

$$\mathrm{RATE}_{t=0} = \alpha_1 S_{t=0} + \alpha_2 C_{t=0} - \sum_{k=0}^{-3} \beta^{\mathrm{L}}_{t=k} O^{\mathrm{L}}_{t=k} - \sum_{k=0}^{-3} \beta^{\mathrm{R}}_{t=k} O^{\mathrm{R}}_{t=k} + \gamma$$

where $S_{t=0}$ represents the stimulus difficulty of the current trial ($t = 0$), which is assumed to be learned through long-term experience with a given stimulus; $C_{t=0}$ represents the choice of sides (left or right, L or R) in the current trial, which is known to influence the firing rate of OFC neurons[31,36]. The variable $O^{\mathrm{SIDE}}_{t=k}$ represents outcomes of the current trial and past three trials ($t = -1, -2, -3$), separated according to the side where the reward was received, again to account for the known selectivity of rodent OFC neurons[31,36]. The coefficients $\alpha_1$ and $\alpha_2$ measure the influence of the stimulus difficulty and the choice, $\beta^{\mathrm{L}}_{t=k}$ and $\beta^{\mathrm{R}}_{t=k}$ measure the influence of current and past trial outcomes, and $\gamma$ captures the mean rate not accounted for by other variables.

The model was fitted using a least-square error criterion with singular value decomposition (SVD). In some cases the problems were ill-conditioned and therefore we also tried ridge regression to obtain more stable solutions. For this analysis, the optimal regularization parameter was chosen by generalized cross-validation[37]. The results of both analyses essentially agreed and therefore we report the results from SVD estimated regression models. The statistical significance of regression coefficients was determined using a permutation test by pseudo-randomly shuffling trial order for the variable of interest[38]. The data were shuffled 1,000 times to yield a $P$ value for the permutation test.

Supplementary Fig. 5a shows the coefficients of this model fit to the neuron shown in Fig. 3a, b. Error bars show standard deviations estimated using leave-one-out-bootstrap[37] and filled circles show significant values at $P < 0.05$ based on a permutation test. This neuron had significant selectivity for the upcoming outcome, $\beta^{\mathrm{L,R}}_{t=0}$, for both choice sides, as well as for the previous outcome, $\beta^{\mathrm{L,R}}_{t=-1}$, to a much smaller degree, while the influence of past outcomes, $\beta^{\mathrm{L,R}}_{t=-2,-3}$, was not significant. Leaving out all past outcomes, $\beta^{\mathrm{L,R}}_{t=-1,-2,-3} = 0$, did not significantly increase the prediction error ($P < 0.05$, permutation test).

This analysis was repeated on the population of 133 neurons (Fig. 4g, h) that were deemed to be negative outcome selective (pooling trials across all stimuli) based on ROC analysis at $P < 0.05$. Supplementary Fig. 5b shows the number of neurons (grey bars) and the mean value of significant regression coefficients (circles, $P > 0.05$). Overall, 121 neurons had significant $\beta^{\mathrm{L,R}}_{t=0}$ coefficients for the current outcome and 70 neurons had significant $\beta^{\mathrm{L,R}}_{t=-1}$ coefficients for the outcome of the previous trial for at least one side. Only four neurons carried past outcome information for at least one side for all three trials back. Comparison of the average value of the significant coefficients for current and past trial outcomes (Supplementary Fig. 5b, circles) shows that even when past trial outcomes had significant coefficients the average value of their weights was only half those for the current trial.

We also performed an analysis to test whether including the history of recent outcomes improves the model fit. To do this, we compared the full model to one in which the coefficients $\beta^{\mathrm{L,R}}_{t=-1,-2,-3}$ were set to zero and used a permutation test to compare the mean prediction errors for the full and reduced model. To obtain a conservative estimate (that is, allow the best chance for inclusion of history terms to increase performance) we did not compensate for the increased complexity of the full model. This analysis showed that for only 12 of 116 neurons did the inclusion of past outcome information, $\beta^{\mathrm{L,R}}_{t=-1,-2,-3}$, significantly reduce the prediction error ($P < 0.05$, permutation test). Moreover, the reduction in error was small, with an average <3% improvement for the full compared to the reduced (current-trial-only) model.

In summary, we conclude that although a subset of OFC neurons do carry information about past outcomes, past trial events account for a relatively small fraction of the firing rate variance compared to what can be explained by the anticipated current trial outcome.

**Outcome selectivity analysis.** Orbitofrontal cortex is known to signal outcome expectations[39–42], and an apparent prediction of outcome might arise from a combination of stimulus and side selectivity. If firing rates encoded the stimulus difficulty (Fig. 2) and in addition were modulated by the choice side[31,35] one

would expect (1) outcome preference would be inverted across choice sides, and (2) outcome selectivity would be equal or weaker for easier compared to more difficult stimuli. A cartoon of this scenario is shown in Supplementary Fig. 3b, with both an additive and a multiplicative component to the choice side modulation. In contrast, the uncertainty model makes the opposite predictions (Supplementary Fig. 3a and Fig. 4d). Although the average tuning curve for negative outcome selective neurons are similar to what is expected for a representation of uncertainty (Fig. 4g), we wanted to test these predictions on a neuron-by-neuron basis. We used the outcome preference index (OP) to measure whether the firing rates are higher or lower for error trials, and the unsigned version of this measure, the outcome selectivity index (OS = |OP|), to measure whether how strongly firing rates signal different outcomes. These measures are based on signal detection theory and quantify the difference between the firing rates for error and correct trials (see Methods Summary for details). Statistical significance was estimated using a 200-fold permutation test[43] at $P < 0.05$. Note that for these analyses trials had to be subdivided according to several stimulus types and for many neurons there were few error trials available to reliably compare conditions. An insufficient number of error trials can result in either spurious selectivity values due to noise and/or low significance values.

First we tested whether the direction of outcome preference was concordant across sides (that is, regular arrows in Supplementary Fig. 3a, b). We used 310 out of 563 neurons for which there were more than 5 error trials for each of 32/68 and 68/32 stimuli. From these neurons 116 showed outcome selectivity across all stimuli, but only 19 were significantly selective for both 32/68 and 68/32 mixtures when considered separately. 85% (16/19) of neurons had concordant outcome preference values, and the preference values were significantly correlated across sides ($r^2 = 0.66$, $P < 0.05$; Supplementary Fig. 3c). Next we tested whether outcome selectivity was stronger for easy stimuli (32/68 and 68/32 mixtures) compared to more difficult ones (44/56 and 56/44 mixtures; see dashed arrows in Supplementary Fig. 3a, b). Out of 317 neurons with 56/44 trials, 131 were selective across all stimuli but only 23 were significant for both easy and difficult mixtures when considered separately. For 91% (21/23) of these neurons, outcome selectivity was stronger for easier stimuli (Supplementary Fig. 3d). These analyses support the uncertainty model (Supplementary Fig. 3a) and are not consistent with the hypothesis that choice side-modulation of stimulus encoding neurons produces an apparent outcome selectivity (Supplementary Fig. 3b).

Next we conducted an additional analysis to show how well individual neurons conform to the firing patterns expected for decision confidence across the entire recorded OFC population. We used OP to measure whether the firing rates are higher or lower for error versus correct trials across 32/68, 44/56, 56/44 and 68/32 stimuli. In addition, we calculated a stimulus difficulty selectivity index (DI) to measure whether firing rates are higher or lower for correct choices in difficult trials (32/68, 44/56, 56/44 and 68/32 stimuli) compared to easy trials (0/100 and 100/0 stimuli). Again, both measures are derived from the area under the ROC from signal detection theory and statistical significance was estimated using a 200-fold permutation test at $P < 0.05$. Supplementary Fig. 8 shows DI as a function of OP across the entire population. Out of 563 neurons, 83 were significant for both measures, 85 for OP alone, 105 for DI alone and 290 were not significant at $P < 0.05$. The selectivity measures were correlated (CC = 0.75 at $P < 0.05$) across the entire population. This analysis shows that across the population without any preselection there is a good correlation between outcome preference (selectivity for correct/error choices) and stimulus difficulty preference (selectivity for more/less difficult stimuli) as expected for a decision confidence signal.

**Interpretation of negative outcome selectivity: error signal or uncertainty?** The observed selectivity of neural activity for the upcoming outcome might arise if, after executing a choice, extra sensory or memory information enters decision-making circuits and causes the realization that an error occurred even before obtaining feedback. According to this interpretation the negative outcome selective population of OFC neurons would signal error[44] instead of uncertainty. In contrast, the highest observed firing rates were associated with near chance level performance and not errors (Fig. 4g, f). To test this more rigorously, we asked whether an ideal observer could obtain better performance than the experimental subject if it could switch choices based on the firing rate after the choice and before feedback is provided. In all but one negative outcome selective neuron (1/133), the highest firing rates (top 5% of trials) were associated with chance level performance (within the 95% confidence interval). Therefore negative outcome selectivity does not imply that OFC neurons are actually able to predict error trials but rather that high firing rates predict near chance level performance consistent with an uncertainty signal.

**Confidence model.** We model the stimulus as the log ratio of the odour mixture with additive Gaussian noise: $s_i = \log\frac{[A]}{[B]} + \eta_{stim}$ in each trial $i$, where $\eta_{stim} \in N(0,\sigma_{stim})$. The boundary is fixed at 0 with additive noise, $b^i = \eta^{bound}$, where $\eta_{bound} \in N(0,\sigma_{bound})$. The choice is computed by comparing stimulus and boundary, $choice_i = \{left | s_i < b_i; right | s_i \geq b_i\}$ The distance between the stimulus and boundary, $d_i = |s_i - b_i|$, provides an estimate of decision confidence. Other distance metrics, such as Euclidian distance, are also suitable. This distance measure can be calibrated and linearized to produce a veridical estimate of outcome probabilities[45]. We did not attempt to systematically calibrate confidence but found that sigmoid functions provide a good approximation (see also Supplementary Information). Therefore we define 'decision confidence', $\delta_i = f(d_i) = \tanh(d_i)$ and its opposite 'decision uncertainty' as $\sigma_i = 1 - \delta_i$. For the simulations in Fig. 4 we chose the stimulus and boundary noise to be equal, $\sigma_{bound} = \sigma_{stim} = 0.5$, but we note that the results are dependent only on the total noise (sum of the variances) not their relative contribution (see Supplementary Fig. 7). Therefore, the model has a single effective parameter, $\sigma_{noise} = \sqrt{\sigma_{bound}^2 + \sigma_{stim}^2}$, that determines the slope of the psychometric function, leaving no free parameters with respect to confidence estimates (Supplementary Fig. 7).


31. Feierstein, C. E. *et al.* Representation of spatial goals in rat orbitofrontal cortex. *Neuron* **51,** 495–507 (2006).
32. Barraclough, D. J., Conroy, M. L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neurosci.* **7,** 404–410 (2004).
33. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Matching behavior and the representation of value in the parietal cortex. *Science* **304,** 1782–1787 (2004).
34. Lau, B. & Glimcher, P. W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84,** 555–579 (2005).
35. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. D Cortical substrates for exploratory decisions in humans. *Nature* **441,** 876–879 (2006).
36. Roesch, M. R., Taylor, A. R. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51,** 509–520 (2006).
37. Hansen, P. C. *Rank-deficient and Discrete Ill-posed Problems: Numerical Aspects of Linear Inversion* (SIAM, 1998).
38. Davison, A. C. & Hinkley, D. V. *Bootstrap Methods and Their Application* (Cambridge Univ. Press, 1997).
39. Hikosaka, K. & Watanabe, M. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb. Cortex* **10,** 263–271 (2000).
40. Wallis, J. D. & Miller, E. K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18,** 2069–2081 (2003).
41. Gottfried, J. A. O'Doherty, J. & Dolan, R.J. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301,** 1104–1107 (2003).
42. Simmons, J. M. Ravel, S. Shidara, M. & Richmond, B.J. A comparison of reward-contingent neuronal activity in monkey orbitofrontal cortex and ventral striatum: guiding actions toward rewards *Ann. N.Y. Acad. Sci.* **1121,** 376–394 (2007).
43. Efron, B. & Tibshirani, R. *An Introduction to the Bootstrap* (Chapman & Hall, 1993).
44. Laubach, M., Wessberg, J. & Nicolelis, M. A. Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature* **405,** 567–571 (2000).
45. Keren, G. On the calibration of probability judgments. Some critical comments and alternative perspectives. *J. Behav. Decis. Making* **10,** 269–278 (1997).