Review

A computational framework for the study of confidence in humans and animals

Adam Kepecs[1],* and Zachary F. Mainen[2]

[1]*Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA*
[2]*Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Av. Brasília s/n, 1400-038 Lisbon, Portugal*

Confidence judgements, self-assessments about the quality of a subject's knowledge, are considered a central example of metacognition. Prima facie, introspection and self-report appear the only way to access the subjective sense of confidence or uncertainty. Contrary to this notion, overt behavioural measures can be used to study confidence judgements by animals trained in decision-making tasks with perceptual or mnemonic uncertainty. Here, we suggest that a computational approach can clarify the issues involved in interpreting these tasks and provide a much needed springboard for advancing the scientific understanding of confidence. We first review relevant theories of probabilistic inference and decision-making. We then critically discuss behavioural tasks employed to measure confidence in animals and show how quantitative models can help to constrain the computational strategies underlying confidence-reporting behaviours. In our view, post-decision wagering tasks with continuous measures of confidence appear to offer the best available metrics of confidence. Since behavioural reports alone provide a limited window into mechanism, we argue that progress calls for measuring the neural representations and identifying the computations underlying confidence reports. We present a case study using such a computational approach to study the neural correlates of decision confidence in rats. This work shows that confidence assessments may be considered higher order, but can be generated using elementary neural computations that are available to a wide range of species. Finally, we discuss the relationship of confidence judgements to the wider behavioural uses of confidence and uncertainty.

Keywords: metacognition; uncertainty; Bayesian; psychophysics; model-based; post-decision wagering

1. INTRODUCTION

Uncertainty is ubiquitous. Both natural events in the world and the consequences of our actions are fraught with unpredictability, and the neural processes generating our percepts and memories may be unreliable and introduce additional variability. In face of this pervasive uncertainty, the evaluation of confidence in one's beliefs is a critical component of cognition. As humans, we intuitively assess our confidence in our percepts, memories and decisions all the time, seemingly automatically. Nevertheless, confidence judgements also seem to be part of a reflective process that is deeply personal and subjective. Therefore, a natural question arises: does assessing confidence—knowledge about subjective beliefs—constitute an example of the human brain's capacity for self-awareness? Or is there a simpler explanation that might suggest a more fundamental role for confidence in brain function across species?

One approach to the study of confidence is rooted in metacognition, traditionally defined as the knowledge and experiences we have about our own cognitive processes [1–3]. Confidence judgements have been long studied as a central example of metacognition. In this light, confidence judgements are viewed as a monitoring process reporting on the quality of internal representations of perception, memory or decisions. Because self-reports appear prima facie the only way to access and study confidence, confidence judgements have been taken to be a prime example of a uniquely human cognitive capacity [1–3] and supposed to require the advanced neural architecture available only in the brains of higher primates [1,4,5]. Moreover, as subjective reports about our beliefs, confidence judgements have even been used as indices of conscious awareness. Against this backdrop of studies emphasizing the apparent association of confidence with the highest levels of cognition, a recent line of research has attempted to show that non-human animals are also capable of confidence judgements, with mixed and sometimes contentious results [5,6]. How could animals possibly think about their thoughts and report their confidence? And even if they did, how could one even attempt to establish this without an explicit self-report?

While such philosophically charged debates persist, an alternative approach, rooted in computational theory, is taking hold. According to this view,

assessment of the certainty of beliefs can be considered to be at the heart of statistical inference. Formulated in this way, assigning a confidence value to a belief can often be accomplished using relatively simple algorithms that summarize the consistency and reliability of the supporting evidence [7]. Therefore, it should come as no surprise that in the course of building neurocomputational theories to account for psychophysical phenomena, many researchers came to the view that probabilistic reasoning is something that nervous systems do as a matter of their construction [8–12].

In this light, confidence reports might reflect a readout of neural dynamics that is available to practically any organism with a nervous system. Hence, representations of confidence might not be explicit or anatomically segregated [13]. Although statistical notions can account for the behavioural observations used to index metacognition, it remains to be seen whether there are aspects of metacognition that will require expanding this framework. For instance, it may be that while choice and confidence are computed together, confidence is then relayed to a brain region serving as a clearinghouse for confidence information from different sources. Confidence representations in such a region can be viewed as metacognitive but nevertheless may still require only simple computations to generate.

Here, we argue that progress in understanding confidence judgements requires placing the study of confidence on a solid computational foundation. Based on the emerging computational framework, we discuss a range of confidence-reporting behaviours, some suitable for animals, and consider what constraints these data provide about the underlying computational processes. Our review also comes from the vantage point of two neuroscientists: we advocate opening up the brain's 'black box' and searching for neural representations mediating metacognition. We conclude by presenting a case study for an approach to understand the neurobiological basis of decision confidence in rats.

2. BEHAVIOURAL REPORTS OF CONFIDENCE IN HUMANS AND OTHER ANIMALS

The topic of behavioural studies of confidence is broad and therefore we will largely limit our discussion to confidence about simple psychophysical decisions and focus mostly on animal behaviour. Behavioural reports of confidence in humans can be explicit, usually verbal or numerical self-reports, and these are usually taken at face value. In contrast, in non-human animals, only implicit behavioural reports are available. This has led to interpretational difficulties, a topic we address next.

(a) *Explicit reports of confidence in humans*
The most straightforward behavioural paradigm for testing confidence is to ask subjects to assign a numerical rating to how sure they are in their answer [14–18]. Indeed, humans performing psychophysical discrimination tasks can readily assign appropriate confidence ratings to their answers [19]. By appropriate rating, we mean that performance accuracy in humans is well
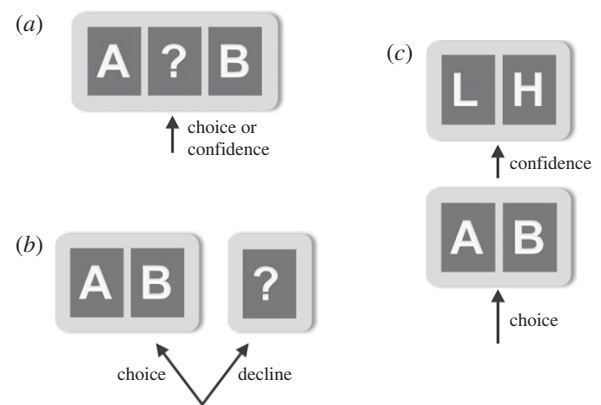
Figure 1. Behavioural tasks for studying confidence in animals. (*a*) In uncertain option tasks, there are three choices, the two categories, A and B, and the uncertain option. (*b*) In decline option or opt-out tasks, there is first a choice between taking the test or declining it, then taking the test and answering A or B. In a fraction of trials, the option to decline is omitted. (*c*) In post-decision wagering, for every trial there is first a two-category discrimination, A or B, and then a confidence report, such as low or high options.

correlated with the self-reported confidence measures. Self-reported confidence also correlates with choice reaction times [14,15]. It should be noted that confidence reports are not always perfectly calibrated; there are systematic deviations found such as overconfidence when decisions are difficult and underconfidence when they are easy [19–21]. Clearly, it is not possible to ask animals to provide explicit confidence reports, therefore animal studies need to employ more sophisticated tasks designed to elicit implicit reports of confidence.

(b) *'Uncertain option' task*
A widely used class of tasks extends the two-choice categorization paradigm by adding a third choice, the 'uncertain option', to the available responses. Interestingly, the scientific study of this paradigm originated in experimental psychology along with quantitative studies of perception [22–25]; however, these early attempts were found to be too subjective [23,24,26] and were soon abandoned in favour of 'forced choice' tasks to quantify percepts based on binary choices. After nearly a century of neglect, a series of studies by Smith and co-workers [5] reinvigorated the field of confidence judgements using these paradigms in both human and non-human subjects with the goal of placing the notion of subjective confidence on a scientific footing.

In the first of the modern studies using this paradigm, Smith *et al.* [27] used a perceptual categorization paradigm (figure 1*a*). A sensory stimulus is presented along a continuum (e.g. frequency of an auditory tone) that needs to be categorized into two classes based on an arbitrarily chosen boundary (above or below 2.1 kHz). Subjects are then given three response options: left category, right category or uncertain response. As might be expected, subjects tend to choose the uncertain option most frequently near the category boundary. Post-experimental questionnaires indicate that humans choose the

uncertain option when they report having low confidence in the answer.

Importantly, the design of the task allowed Smith and colleagues to ask the same questions to non-human animals and hence to reignite an age-old debate about the cognitive sophistication of different animal species. In order to test animals in this task, the different options were linked with different reward contingencies: correct choices were rewarded with one unit of food reward, uncertain responses were rewarded with a smaller amount of food while incorrect choices lead to omission of reward and a time-out punishment. Smith and colleagues studied monkeys, dolphins and rats and compared their performance with humans [27–31]. Under these conditions, monkeys and dolphins showed a qualitative similarity in response strategies as well as a quantitative agreement in the response distributions of animals and humans. The striking similarities of dolphin and monkey behaviour to that of humans suggested that these animals possess more sophisticated cognitive architectures than previously appreciated. Interestingly, their studies also failed to show that evolutionarily 'simpler' animals such as rats could perform confidence judgements [5]. The authors concluded that this might reflect a failure to find a suitable task for accessing the appropriate abilities in these species. Indeed, pigeons can respond similarly to primates in such tasks [32], although they fail to exhibit other uncertainty-monitoring behaviours (see below).

A weakness in the design of the confidence-reporting tasks used by the early studies by Smith and colleagues is the possibility that uncertain responses could be simply associated with those stimuli intermediate to the extreme category exemplars. In other words, the two-alternative plus uncertain option task can be alternatively viewed as a three-choice decision task (left/right/uncertain, figure 1*a*), which can be solved simply by learning appropriate stimulus-response categories without necessitating confidence estimates. This criticism was first addressed with task variants that require same-different discrimination [28]. In this version of the task, there are no external stimuli that can be associated with the uncertain option. Nevertheless, from a computational perspective, if the difference between the two stimuli is represented in the brain, then again the uncertain response can simply be associated with a class of such difference representations. To argue against such associative mechanisms, Smith *et al.* [33] also reported on a task version in which monkeys were only rewarded at the end of a block of trials, so that individual responses were not reinforced. However, while this manipulation does rule out the simplest forms of reward learning, it is still compatible with more sophisticated forms [34].

(c) 'Decline option' tasks

Hampton introduced a memory test that included a 'decline' option rather than three choices [35] (figure 1*b*). In this test, monkeys performed a delayed-matching-to-sample task using visual stimuli.

At the end of the delay, subjects were presented with the option of accepting or declining the discrimination test. Once a subject accepted the test, it received either a food reward for correct choices or a timeout punishment for error choices. After declining the test, the subject received a non-preferred food reward without a timeout. Therefore, the optimal strategy was to decline when less certain, and indeed, monkeys tended to decline the discrimination more often for longer delays. However, since the difficulty was determined solely by the delay, the decline option could be learned simply by associating it with longer delays for which performance was poorer. To circumvent this problem, Hampton introduced forced-choice trials in which the subject had no choice but to perform the discrimination test. If longer delays were simply associated with the choice to decline the test, then decline responses would be equally likely regardless of performance. However, he found a systematic increase in performance in freely chosen trials compared with forced-choice trials, consistent with the idea that monkeys are monitoring the likelihood of being correct, rather than associating delays with decline responses.

A similar decline option test was used by Kiani & Shadlen [36] in macaque monkeys, who made binary decisions about the direction of visual motion. On some trials, after stimulus presentation, a third 'opt out' choice was presented for which monkeys received a smaller but guaranteed reward. They found that the frequency of choosing the 'uncertain' option increased with increasing stimulus difficulty and with shorter stimulus sampling. Moreover, as observed by Hampton, monkeys' performance on trials in which they declined to opt out was better than when they were forced to perform the discrimination.

Inman & Shettleworth [37] and Teller [38] tested pigeons using similarly designed 'decline' tests with a delayed-matching-to-sample task. They observed that the rates of choosing the decline option slightly increased with delay duration. Because performance decreased with delay, decline choices also increased as performance decreased. However, there was no difference in the performance on forced-choice versus free-choice (i.e. non-declined) trials. These results were interpreted as arguing against the metacognitive abilities of pigeons [39].

After these negative results with pigeons and rats in uncertain and decline option tasks, it came as a surprise that Foote & Crystal [40] reported that rats have metacognitive capacity. Similar to the design of Hampton, they used a task with freely chosen choice trials with a decline option interleaved with forced-choice trials. But unlike Hampton, Inman and Teller they used an auditory discrimination task without an explicit memory component. They found that decline option choices increased in frequency with decision difficulty and that free-choice performance was better than forced-choice performance on the most difficult stimuli, arguing against associative learning of decline choices.

The argument that this class of task tests confidence-reporting abilities chiefly rests on the decrease in performance on forced-choice trials compared with freely chosen choice trials. The results for rats showed

a small change in performance (less than 10%) and only for the most difficult discrimination type. An alternative explanation is that attention or motivation waxes and wanes, and animals' choices and performance are impacted by their general 'vigilance' state. When their vigilance is high, animals would be expected to choose to take the test and perform well compared with a low vigilance state when animals would tend to decline and accept the safer, low-value option. Although being aware of one's vigilance state may be considered as a form of metacognition, it is distinct from mechanisms of confidence judgements.

(d) *Post-decision wagering*

The 'uncertain option' and 'decline option' tasks have the weakness that either a choice report or a confidence report is collected in each trial but not both. The 'post-decision wager' paradigms improve on this by obtaining both choice and confidence on every trial [41,42]. The central feature of this class of paradigms is that after the choice is made, confidence is assessed by asking a subject to place a bet on her reward (figure 1*c*). The probability of betting high (or the amount wagered) on a particular decision serves as the index for confidence. Persaud and colleagues used this paradigm to test a subject with blindsight, who had lost nearly his entire left visual cortex yet could make visual discriminations in his blind field despite having no awareness. Using a post-decision wagering paradigm, they found that wagers were better correlated with the subjects' explicit self-reported visibility of the stimulus than with actual task performance [41]. Hence, the authors argued that post-decision wagers not only provided an index of confidence, but also served as an objective assessment of 'awareness', independent of perceptual performance.

Leaving aside the thorny issue of whether post-decision wagers can be used to study awareness [43,44], for the purposes of studying confidence judgements, the wagering paradigm has many attractive features. Wagers provide a means to make confidence reports valuable and hence by providing appropriate reward incentives animals could be trained to perform post-decision wagering.

One caveat with post-decision wagering paradigms is that because the pay-off matrix interacts with the level of confidence to determine the final payoff, care must be taken with the design of the matrix. It has been observed that in the study of Persaud *et al.* [41], the optimal strategy for the pay-off matrix was to always bet high regardless of the degree of confidence [45,46]. Although subjects were in fact found to vary their wager with uncertainty, it would be difficult to disambiguate a suboptimal wagering strategy from the lack of appropriate estimations of confidence [45,47]. Therefore, the design of the pay-off matrix as well as an independent evaluation of the wagering strategy are important considerations. Using a continuum of wagers instead of a binary bet (certain/uncertain) somewhat mitigates the difficulty of finding an optimal pay-off matrix. A second concern about the post-decision wager task is that bets might be placed by associating the optimal wager with each stimulus

using reinforcement learning. In particular, each two-choice plus wager test could be transformed into a four-choice test where distinct stimuli could be associated with distinct responses. For instance, in a motion-direction categorization, weak versus strong motion to one direction could be associated with low versus high wagers. Importantly, however, this concern can be alleviated by appropriate analysis of behavioural data, as we will discuss below. Moreover, regardless of concerns over the optimality of wagering, animals could in principle be trained to perform this kind of task [48], providing a rich class of confidence-reporting behaviours.

(e) *Decision restart and leaving decision tasks*

Recently, Kepecs *et al.* [49] introduced a behavioural task similar to post-decision wagering that can be used in animals. They trained rats to perform a two-alternative forced-choice (2AFC) olfactory discrimination task. In this task, subjects initiated a trial by entering a central odour port. This triggered the delivery of a stimulus comprising a binary odour mixture. The subject was rewarded for responding to the left or to the right depending on the dominant mixture component. In order to vary the decision difficulty, the ratio of the two odours was systematically varied. Rats performed at near chance level for 50/50 odour mixtures and nearly perfectly for pure odours. To assess confidence, reward was delayed by several seconds while subjects were given the option to 'restart' trials by leaving the reward port and re-entering the odour sampling port. In other words, after the original decision about the stimulus, rats were given a new decision of whether to stay and risk no reward with timeout punishment, or leave and start a new, potentially easier trial. An important feature of this task, similar to the post-decision wagering and confidence-rating tasks, is that the choice and confidence reports are collected in the same trial, an issue we will further discuss below.

It was observed that the probability of aborting and restarting trials increased with stimulus difficulty and that discrimination performance was better on trials in which the subjects waited for the reward compared with trials in which they reinitiated. Moreover, as we will describe in more detail below, the probability to reinitiate choices matched the pattern expected for confidence judgements. In particular, even for trials of the same stimulus, Kepecs *et al.* [49], observed that restarts were systematically less frequent for correct than for error trials. This correct/error analysis permitted the authors to circumvent the criticism that reinforcement learning on stimuli could explain the pattern of behaviour (see below).

One difficulty with this task is that its parameters (e.g. reward delay) need to be carefully tuned in order to have a reasonable balance between restarted and non-restarted trials. This difficulty, in part, can be traced to the issue of choosing an appropriate pay-off matrix in post-decision wagering paradigms. The cost of waiting and the value of restarts must be chosen just right and are difficult to infer *a priori*. For instance, in some parameter regimes rats never restarted trials, while in others they always restarted until an easier stimulus was
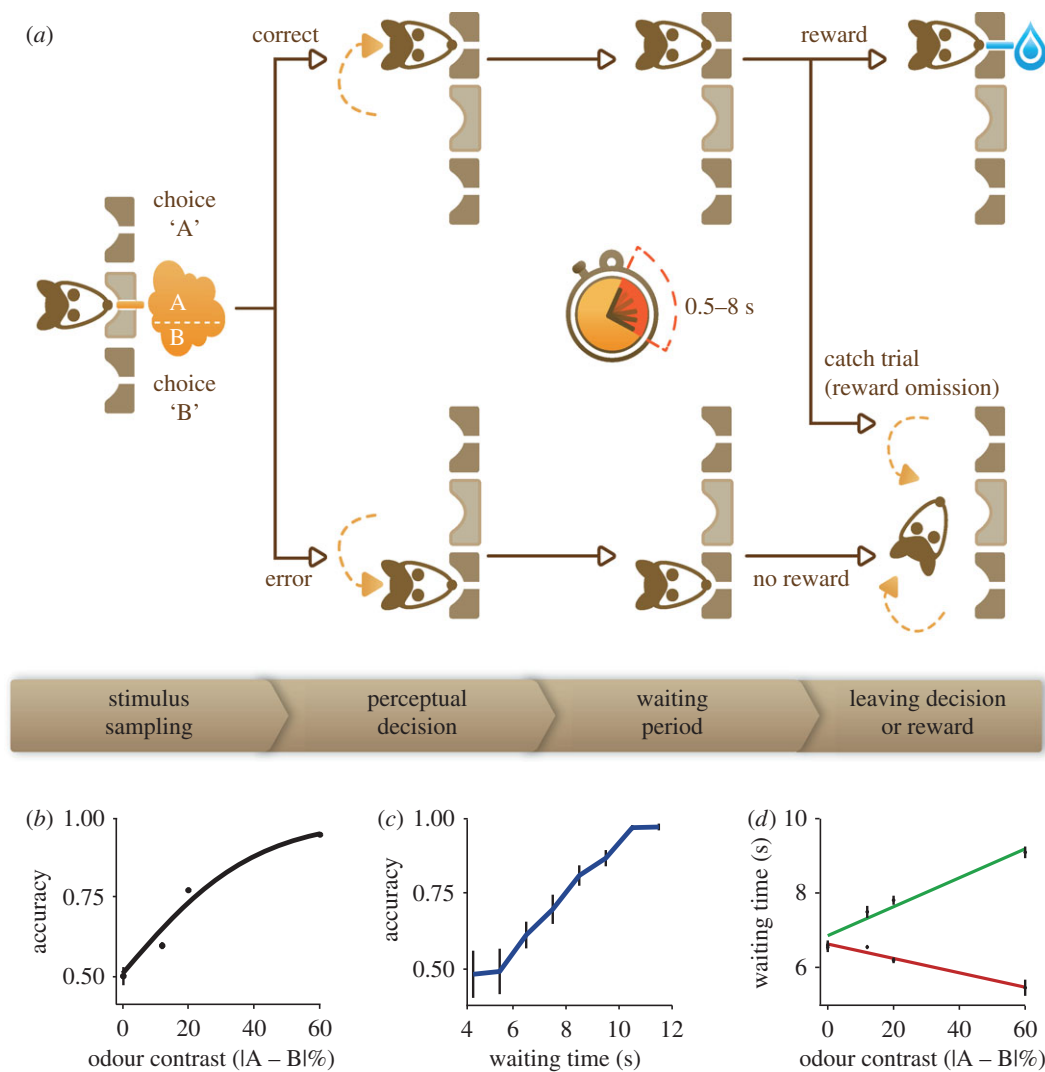
Figure 2. Leaving decision tasks for studying confidence in animals. (*a*) Schematic of the behavioural paradigm. To start a trial, the rat enters the central odour port and after a pseudorandom delay of 0.2–0.5 s an unequal mixture of two odours is delivered. Rats respond by moving to the A or B choice port, where a drop of water is delivered after a 0.5–8 s (exponentially distributed) waiting period for correct decisions. In catch trials (approx. 10% of correct choices), the rat is not rewarded and no feedback is provided. Therefore, the waiting time can be measured (from entry into choice ports until withdrawal) for all error and a subset of correct choices. (*b*) Psychometric function for an example rat. (*c*) Choice accuracy as a function of waiting time. For this plot, we assumed that the distribution of waiting times for correct catch trials is a representative sample for the entire correct waiting time distribution. (*d*) Mean waiting time as a function of odour mixture contrast and trial outcome (correct/error) for an example rat.

provided (A. Kepecs, H. Zariwala & Z. F. Mainen 2008, unpublished observations). A related issue with binary wagers is that in each trial only a single bit of information is gained about decision confidence.

Both of these issues can be mitigated using task versions that provide graded reports of decision confidence (figure 2*a*). Rats were trained in a task variant we call a 'leaving decision' task. In this version, we delay reward delivery using an exponential distribution of delays (to keep reward expectancy relatively constant) and measure the time an animal is willing to wait at the choice ports (figure 2*a*). Incorrect choices are not explicitly signalled and hence rats eventually leave the choice ports to initiate a new trial. In order to measure confidence for correct choices, we introduce a small fraction (approx. 10–15%) of catch trials for which rewards are omitted. The waiting time at the reward port before the leaving decision (obtained for all incorrect and a fraction of correct trials) provides a graded measure reflecting decision confidence (figure 2*c*). Waiting time, naturally, also depends on when the animal is expecting the reward delivery. However, we found that the relative patterns of waiting times were systematically related to decision parameters (figure 2*c*,*d*, see below) for a range of reward-delay distributions providing a behaviourally robust proxy for decision confidence. Indeed, an accurate estimate of decision confidence modulating the waiting time will help to maximize reward rate, while also minimizing effort and opportunity costs incurrent by waiting. Because the animals' cost functions are difficult to infer, we cannot make quantitative predictions about the optimal waiting time. Nevertheless, based on reasonable assumptions, we expect accuracy to be monotonically related to waiting time, in agreement with our observations (figure 2*c*). Note that although waiting time can also be measured in the 'decision restart' task variant, for all restarted trials,

we did not always observe a simple relationship between waiting time and confidence.

(f) *Looking tests*

In addition to tests based on psychophysical methodologies, the metacognitive abilities of animals have also been addressed using more ethologically minded behaviours. So-called 'looking' paradigms take advantage of the fact that during foraging animals may naturally seek information about where foods might be located [50,51]. Such information-seeking can be considered an assessment of the animal's state of knowledge: the less certain they are about their given state of knowledge the more likely they will seek new information [52]. Indeed, chimpanzees, orangutans, macaques, capuchins and human children all show a tendency to seek new information when specifically faced with uncertainty [51,53–55]. By looking more frequently in the appropriate situations, these species can demonstrate knowledge about their own belief states. In contrast, dogs have so far failed to show such information-seeking behaviour [56,57]. However, in the case of such a failure, it remains possible that the set-up was not ecologically relevant for the species in question.

It may be useful to consider looking tests as an instance of a more general class of behaviours in which confidence assessment can be useful to direct information seeking or exploration. Given the limitations of experimental control possible in ethological settings, it would be profitable to transform the looking tests into psychophysical paradigms where the confidence can be read out by choices to seek out more information [58].

(g) *Criticisms of confidence-reporting behaviours*

Many of these studies discussed above triggered controversies; some of the criticisms have been highlighted above. To summarize, critics have systematically attempted to come up with alternative explanations for the performance of non-humans animals that do not require uncertainty monitoring. The primary thrust of these critiques has been that some confidence-reporting tasks can be solved by learning appropriate stimulus-response categories without the need for true uncertainty monitoring [59]. A second important criticism is that a behavioural report can in some cases arise from reporting the level cognitive variable, such as 'motivation', 'attention' or 'vigilance' that impacts performance, rather than confidence *per se* [35]. Thus, a simpler mechanism might be sufficient to account for the observed behaviour without invoking confidence or metacognition. However, we have seen that these alternative classes of explanation, while very important to address, are being tackled through increasingly sophisticated task designs [49,60,61].

A third, somewhat different, line of criticism has questioned the similarity of various confidence tests to confidence-reporting tests that can be performed by humans [4,62,63]. We find this line of criticism much less compelling. For instance, it has been suggested that if long-term memory is not required then a task

cannot be considered metacognitive. Yet, from a neuroscientific (mechanistic) perspective, it is hard to see the relevant difference between a memory representation and a perceptual representation. A second argument has been that generalization across tasks is an important requirement [61,64,65]. This argument is akin to the claim that a speaker of a single language, such as English, does not demonstrate linguistic competence until she is also shown to generalize to another language, such as Hungarian. Clearly, cognitive flexibility and knowledge of Hungarian are advantageous skill, but not necessary to demonstrate linguistic competence. Likewise, it may be expected that confidence reporting, like other sophisticated cognitive capabilities, may not be solved in a fully general form by most animals or even humans. Finally, some studies have been criticized on the basis of the number of subjects (e.g. 'two monkeys alive have metacognition') [66]. We find this criticism somewhat out of line, especially considering that it has long been routine in monkey psychophysical and neurophysiological studies to use only two subjects. The legitimacy of extrapolating from few subjects is based in part on the argument that individuals of a species share a common neural, and hence cognitive, architecture.

We conclude this section by noting with some puzzlement that it has rarely, if ever, been suggested that human behavioural reports of confidence themselves might be suspect. Why should it be taken for granted that self-reported confidence judgements in humans require an instance of metacognition and uncertainty-monitoring processes? Ultimately, whether applied to human or to other animals, we are stuck with observable behaviour. That human behavioural reports can have a linguistic component while animal reports cannot does not justify two entirely distinct sets of rules for human versus animal experiments. Regardless of the species of the subject, we ought to determine whether a particular behavioural report can be implemented through a simpler mechanism, such as associative learning. In order to best make this case, it is critical to be very careful about how confidence behaviour is defined. To do so, we will argue that semantic definitions need to be dropped in favour of formal (mathematical) ones. It is to this topic that we turn next.

3. COMPUTATIONAL PERSPECTIVE ON CONFIDENCE JUDGEMENTS

The study of decision-making provides important insights and useful departure points for a computational approach to uncertainty monitoring and confidence judgements. Smith and colleagues in their groundbreaking review advocated and initiated a formal approach to study confidence judgements [5]. We argue that this approach can be taken further to provide a mathematically formal and quantitative foundation. That is because formal definitions can yield concrete, testable predictions without resorting to semantic arguments about abstract terms [67,68]. To seek a formal basis for confidence judgements, we will first consider computational models of simpler forms of decision-making [69].

From a statistical perspective, a two-choice decision process can be viewed as a hypothesis test. In statistics,
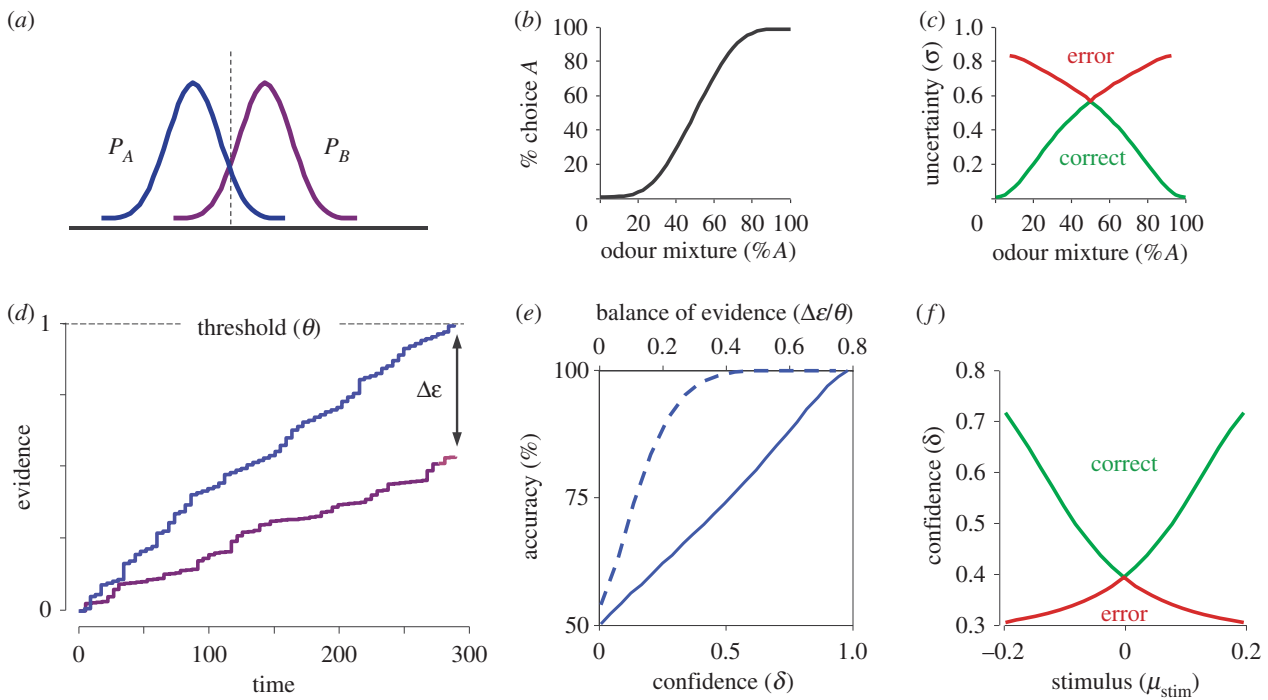
Figure 3. Computational models for choice and confidence. (*a*–*c*) Bayesian signal detection theory model. (*a*) The two stimulus distributions, *A* and *B*. (*b*) Psychometric function. (*c*) Predictions for confidence estimate as a function of observables. (*d,e*) Dual integrator, race model. (*d*) Schematic showing the accumulation of evidence in two integrators up to a threshold. Blue line denotes evidence for *A* and purple line denotes evidence for *B*. (*e*) Calibration of 'balance of evidence' yields veridical confidence estimate. (*f*) Prediction for confidence estimate as a function of observables.

each hypothesis test can be paired with an interval estimation problem to compute the degree of confidence in the hypothesis [7]. Perhaps, the most familiar quantitative measure of confidence is the *p*-value that can be computed for a hypothesis test. Indeed, the notion of confidence is truly at the heart of statistics, and similarly it should be at the centre of attention for decision-making as well. Moreover, statistical analysis provides a solid departure point for any attempt to seek psychophysical or neural evidence for confidence.

We begin with the core idea that confidence in a decision can be mechanistically computed and formalized in appropriate extensions of decision models. First, we will define confidence and then discuss how to derive (compute) it.

(a) *Defining confidence*

Confidence can be generally defined as the degree of belief in the truth of a proposition or the reliability of a piece of information (memory, observation and prediction). Confidence is also a form uncertainty, and, previously, several classifications of uncertainty have been discussed. In psychology, external and internal uncertainties have been referred to as 'Thurstonian' and 'Brunswikian' uncertainty, respectively [70,71]. In economics, there are somewhat parallel notions of 'risk' and 'ambiguity' [72–76]. Risk refers to probabilistic outcomes with fully known probabilities, while in the case of ambiguity, the probabilities are not known.

Here, we focus on decision confidence, an important instance of uncertainty, which summarizes the confidence associated with a decision. Decision confidence can be defined, from a theoretical perspective,

as an estimate by the decision-maker of the probability that a decision taken is correct. Note that we also use 'decision uncertainty' interchangeably, after a sign change, with 'decision confidence'.

(b) *Bayesian and signal detection models for decision confidence*

Signal detection theory (SDT) and Bayesian decision theory [7] provide quantitative tools to compare the quality of stimulus representation in neurons with variability in behavioural performance [77]. These quantitative approaches have provided a strong basis for probing the neural mechanisms that underlie perception [78].

We begin with a model for a two-alternative decision, such as the olfactory mixture discrimination paradigm discussed above [7]. For the purposes of our argument, we will consider a simplified case of discriminating two stimuli, *A* and *B*, in Gaussian noise (figure 3*a*). Stimulus *A* is distributed as $P_1(s|A) \sim N(s_A, \sigma_A)$, with a mean of s_A and a variance of σ_A^2, and similarly, stimulus *B* is drawn from $P_2(s|B) \sim N(s_B, \sigma_B)$. For simplicity, we assume that their variances are equal. Bayesian decision theory proposes that subjects should maximize their expected reward based on both prior information and current evidence. This can be achieved by choosing the larger posterior

$$\text{choice} = (A \text{ if } P(A|s) > P(B|s), \quad \text{otherwise } B),$$

where

$$P(A|s) = \frac{P(A)P(s|A)}{P(s)}$$

is computed using Bayes' rule. From this, we can compute confidence as

$$P(\text{correct}|s) = \max\left(P(A|s), P(B|s)\right).$$

An alternative is to compute an intermediate decision variable

$$d(s) = \log\frac{P(s|A)}{P(s|B)} + \log\frac{P(A)}{P(B)},$$

the log posterior ratio, which, assuming a uniform prior distribution over the stimuli, simplifies to the log-likelihood ratio, $\log P(s|A)/P(s|B)$. In this formulation, the decision rule is to choose A if $d(s) > 0$ and $|d(s)|$, the decision distance, is a measure of confidence.

This model yields specific predictions about how the representation of confidence relates to other variables. First, by definition confidence predicts the probability of a choice being correct, in keeping with the intuitive notion of 'confidence'. Second, when computed as a function of the stimulus difficulty and choice outcome (the observables in a 2AFC task), the model predicts a distinctive and somewhat counterintuitive 'X' pattern (figure 3c), in which confidence increases with signal-to-noise for correct choices while decreasing for error choices.

Note that these patterns are not only robust to different stimulus distributions, but can also be derived from other choice model frameworks based on integration of evidence [49,79,80], attractor models [81,82] and even support vector machine (SVM) classifiers [83] as we discuss next.

(c) Other models of decision confidence: integrators, attractors and classifiers

We have seen how natural it is to introduce a notion of confidence in SDT. Integrator or drift-diffusion models of decision-making can be seen as adding a temporal dimension to SDT [84,85], and we can extend these models in similar ways. This is most natural to examine for the 'race' variant of the integrator model where evidence for and against a proposition accumulates in separate decision variables (figure 3d). Vickers proposed that confidence in a decision may be computed as the 'balance of evidence', the difference between the two decision variables at decision time [86]. This distance can be transformed into a veridical estimate of confidence with qualitatively the same properties as the estimate from SDT models (figure 3e,f).

Another class of models where similar notions of confidence can be applied is classifier models from machine learning theory. For instance, SVMs are a class of algorithms that learn by example to assign labels to objects. In a probabilistic interpretation of SVM classifiers [83], the size of the margin for a sample (distance of the separating hyperplane to the sample) is proportional to the probability of that point belonging to a class given the classifier (separating hyperplane). This yields what is known technically as a measure of the 'posterior variance of the belief state given the current model', and, in other words, an estimate of confidence about the

category [83,87]. Beyond providing a good prediction of classification accuracy, this confidence measure also yields the same qualitative 'X' patterns when plotted as function of stimulus and outcome [88].

These normative models can account for how confidence may be computed algorithmically but not for how confidence is used to make choices, such as the confidence-guided restart decisions discussed above. Insabato and colleagues introduced a two-layer attractor network based on integrate-and-fire neurons that can accomplish this [82,89]. The first network is a competitive decision-making module whose dynamics compute a categorical choice based on a noisy (uncertain) stimulus. This decision network then feeds into a second attractor network in which a 'low confidence' and a 'high confidence' neuron pool compete with the winning population representing the decision to stay or restart, respectively. The attractor networks are not handcrafted and tuned for this purpose but rather based on generic decision networks [90,91] that have been used to account for other decision processes. Interestingly, the design of this model suggests a generic architecture in which one network monitors the confidence of another, similar to cognitive ideas about uncertainty monitoring [92].

Some of the specific models presented above can be interpreted as normative models, prescribing how the computation of confidence ought to be done based on some assumptions. In this sense, they are useful for describing what a representation of confidence or its behavioural report should look like. Some of these models are also generative and can be taken literally as an algorithm that neural circuits might use to compute confidence. These models may also be useful in considering criticisms, such as the argument that some metacognitive judgements can be performed simply by stimulus-response learning based on internal stimulus representations. At least for certain behavioural tasks, as we have shown, this is not the case since mapping confidence requires a specific, highly nonlinear form of the subjective stimulus beyond simple forms of associative learning. Although more sophisticated algorithms might compute confidence in different ways, ultimately what characterizes confidence is that it concerns a judgement about the quality of a subjective internal variable, notwithstanding how that response is learned. In this sense, the models presented here can serve as normative guideposts.

Taken together these computational modelling results establish, first, that computing confidence is not difficult and can be done using simple operations, second, that the results are nearly independent of the specific model used to derive it, and finally that it requires computations that are distinct from the computation of other decision variables such as value or evidence.

(d) Veridical confidence estimates and calibration by reinforcement learning

Thus far, we have discussed how simple models can be used to compute decision confidence on a trial-by-trial basis. However, we dodged the question of how to find the appropriate transform function that will result in a veridical confidence estimate. First, what do we mean by an estimate being veridical? A veridical estimate is

the one that correctly predicts the probability of its object. For instance, in Bayesian statistics, the posterior probability of an event is a veridical estimate of confidence. In our examples above, we call estimates veridical if they are linearly related to accuracy. Of course, most confidence judgements are not entirely veridical; in fact, people tend to systematically overestimate or underestimate their confidence. While such systematic deviations have been extensively studied, they are likely to involve a host of emotional and social factors that we wish to leave aside for now. Rather we focus on the basic computational question of how can naive confidence estimates be tuned at all so they roughly correspond to reality [93]?

To obtain veridical confidence estimates, it is necessary to calibrate the transfer function (e.g. figure 3e). We can assume that the calibration transform changes on a much slower timescale compared with variations in confidence, and hence this computation boils down to a function-learning problem. Therefore, a subject can use reinforcement learning, based on the difference between the received and predicted outcome (derived from confidence), to learn the appropriate calibration function. Interestingly, consistent with this proposal, experiments show that confidence ratings in humans become more veridical with appropriate feedback [16]. Note that this use of reinforcement learning to calibrate confidence still relies on a trial-to-trial computation of a confidence estimate.

(e) Applying predictions of confidence models to confidence-reporting tasks

These computational foundations for confidence emphasize the separation between *how* a particular representation is computed and *what* function it ultimately serves. But in order to study confidence, nearly all behavioural tasks exploit the fact that animals try to maximize their reward and therefore incentives are set up so that maximizing reward requires the use of confidence information. There is a multitude of possible approaches, for instance, using the idea that confidence can also be used to drive information-seeking behaviour [58,94,95], which is exploited in the more ethologically configured 'looking' paradigms discussed above. Clearly, confidence signals can have many functions, and correspondingly many psychological labels. Therefore, our first goal is not to study how confidence is functionally used ('reward maximization' or 'information seeking') but rather its algorithmic origin: how it was computed. To accomplish this, we can use the computational models introduced above to formally link the unobservable internal variable, confidence, to observable variables, such as stimuli and outcomes. This general strategy is beginning to be used to infer various decision variables such as subjective value representations [94,96–102].

In order to argue that rats reported their confidence by restarting, we showed that the probability of restarting was not only dependent on stimulus difficulty but also the correctness of the choice. Figure 2d shows the observed (folded) 'X' pattern for the 'leaving decision' task variant. This pattern of data is critical in that it can rule out the two main criticisms discussed above

with reference to confidence tests. First, this pattern cannot be explained by assuming that reinforcement learning assigned a particular degree of confidence to each stimulus. That is because correct and error choices for the same stimulus are associated with different confidence measures. Note that we do not exclude the possibility that reinforcement learning processes may be used to calibrate confidence on a slower timescale as discussed above. In this respect, fitting confidence reports to reinforcement learning models is useful to rule out the contribution of such process to correct/error differences in confidence reports. Second, since this pattern is 'non-factorizable', it cannot be reproduced by independently combining stimulus difficulty effects, manipulated by the experimenter, with a waxing and waning internal factor, such as vigilance or attention. This rules out the alternative explanation used for the 'decline option' tests [35,36] according to which decline choices follow a stimulus-difficulty factor times an attention or memory-dependent factor. The leaving decision version of this task enables even stronger inferences, because waiting time is a graded variable. Indeed, as expected for a proxy for confidence, waiting time predicts decision accuracy (figure 2c). Moreover, these trial-to-trial confidence reports can be directly fitted to alternative models, such as those based on reinforcement learning in an attempt to exclude them [49].

It is interesting to note that the same method of separating correct and error choices could be applied to the 'post-decision wagering' test [41]. While appropriate data from these tasks, i.e. sorted by both correct and error as well as by difficulty, may already be available, to our knowledge they have not been reported in this way.

Also note that some other tasks, such as the 'decline option' test or 'uncertain option' task, do not admit this possibility because one obtains either an answer or a confidence judgement, but not both, in any given trial. If the animal declines to take the test, there is no choice report, hence no error trials to look at. As a result, only weak inferences are possible leaving us with a plethora of alternative explanations for the observed data [3,4,42,62,63].

To summarize the past two sections, the lesson we take from these studies and the related debates is threefold. First, confidence-reporting tasks should collect data about the choice and the confidence associated with it *in the same trial* and for as large a fraction of trials as possible. Lacking this, it is difficult if not impossible to rule out alternative mechanisms. Second, the confidence readout should ideally be a *graded* variable. Finally, we believe that to call a particular behaviour a 'confidence report' we need to drop semantic definitions and focus on formal accounts of confidence by fitting appropriate models to the behavioural data.

4. A CASE STUDY OF DECISION CONFIDENCE IN RAT ORBITOFRONTAL CORTEX

Although computational models can be used to rule out and to some degree infer certain computational strategies, behavioural reports alone provide fundamentally limited evidence about the mechanisms generating that report. Therefore, ultimately we need
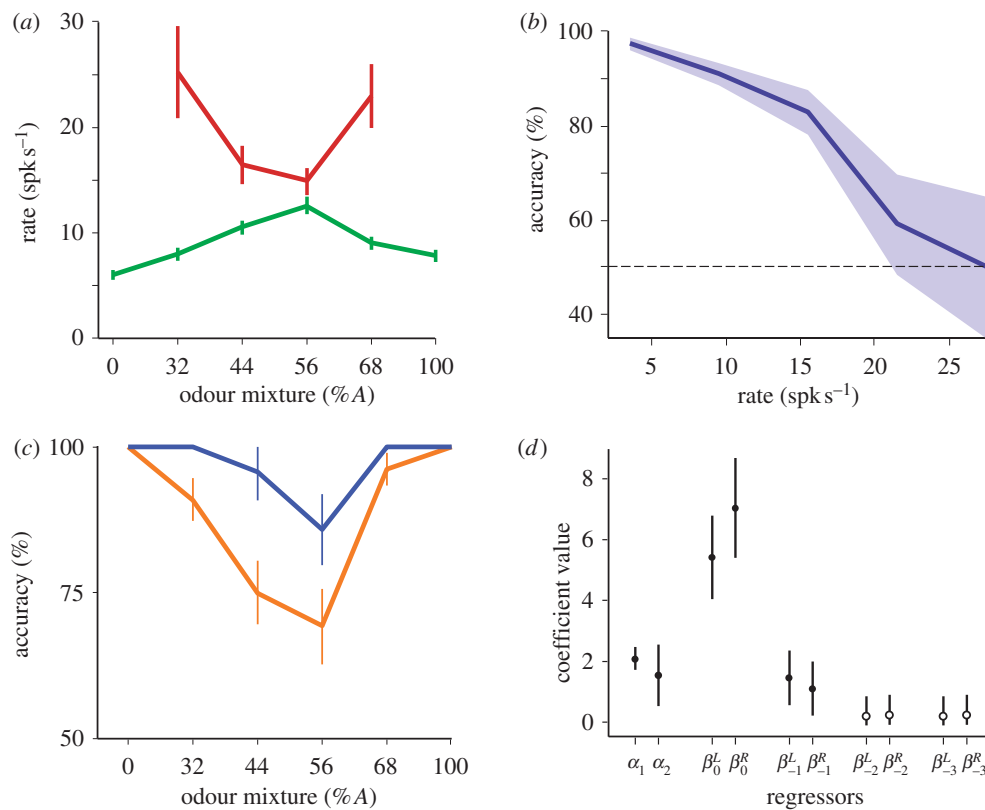
Figure 4. Neural correlates of decision confidence in rat orbitofrontal cortex. (a–d) Analyses of firing rate for a single neuron rat in OFC after the animal made its choice and before it received feedback. (a) Tuning curve as a function of stimulus and outcome (red, error; green, correct). (b) Firing rate conditional accuracy function. (c) Psychometric function conditioned of firing rate (blue, low rates; orange, high rates). (d) Regression analysis of firing rate based on reward history. Coefficient α_1 is an offset term, α_2 is stimulus difficulty and β coefficients represent outcomes (correct/error) divided by left/right choice side and a function of recent trial history (current trial = 0). Note that the largest coefficients are for the current trial and beyond the past trial the coefficients are not significantly different from zero (unfilled circles).

to look into the brain and attempt to identify the necessary neural representations and processes underlying the assumed computations.

(a) *Representation of decision confidence in orbitofrontal cortex*

We wondered if orbitofrontal cortex (OFC), an area involved in representing and predicting decision outcomes, carries neural signals related to confidence [103–105]. Our principal strategy was to look for neural correlates of confidence and try to understand their origin in a mechanistic framework [49]. We recorded neural firing in OFC while rats performed the olfactory mixture categorization task described earlier. We focused our analysis on the reward-anticipation period, after the choice was made and while rats were waiting at the reward port, but before they received any feedback about their answer. Figure 4 shows an example neuron whose firing rate during this anticipation period signals decision uncertainty.

How did we establish that this is confidence and not some other variable? Similar to our approach to analysing the confidence-reporting behaviour, we first plotted firing rates as a function of the stimulus and outcome (two observables). We noticed the same non-factorizable 'X' pattern as a function of stimulus and outcome that is a key prediction of confidence models (figure 4a). This is also the pattern of behavioural responses we observed during leaving decisions. Second, the firing rates predict

accuracy, being highest on average for trials associated with chance performance and lowest on average for trials with near perfect performance (figure 4b). This is essentially a definition of confidence (or rather its inverse in this case, decision uncertainty). Third, to show that these neurons predict accuracy beyond what is knowable from the stimulus, we plotted behavioural accuracy as a function of stimuli conditioned on low and high firing rates (figure 4c). This shows that the correlation between firing rates and performance is not solely owing to stimulus information, because knowing even a single bit about firing rates (high or low) can significantly improve behavioural predictions. Fourth, the firing rates could not be explained by recent reward history as determined by a regression analysis, showing that reinforcement learning based on past experiences with outcomes did not produce this pattern. Note that some forms of performance fluctuations coupled to reinforcement learning could result in different firing rates for correct and error choices. Therefore, while the 'X' pattern is suggestive, it is crucial to explicitly rule out these history-based mechanisms. Although, in principle, this analysis could be applied to binary choices as well, it is more powerful for continuous variables like firing rate. This result implies that firing rates were produced by a process that uses information mostly from the current trial. Taken together, the most parsimonious explanation of these data is that neurons have access to a measure of decision confidence.

We found over 20 per cent of neurons in OFC had correlates like the neuron described earlier that can be called 'decision uncertainty', while about 10 per cent carried a signal of the opposite sign called 'decision confidence'. This confidence signal is a scalar quantity, which is why it is surprising that so many OFC neurons encoded this one variable. OFC supports a broad range of functions and we expect that it will encode other variables as well. The current understanding of OFC suggests that it is mainly involved in outcome prediction [103,104,106]. Predicting outcomes is difficult and different situations call for different computational mechanisms. In a psychophysical decision task like the one we used, the only source of stochasticity is the decision of the animal. Therefore, an estimate of confidence provides the appropriate prediction of trial outcome. At the same time, we expect that OFC neurons will incorporate different signals as needed to make outcome predictions, as we discuss elsewhere [107].

At this point, it is important to return to semantics. First, we use the word confidence to refer to the formal notion of decision confidence, which happens to overlap to a large degree with our intuitive notion of confidence. Second, we wish to emphasize strongly that we are not labelling these neurons as representing 'confidence' or anything else; the claim being made is that they must have had *access* to mechanisms that computed an estimate of confidence. In this sense, what we established is not a neural correlate of a behaviour but rather a computational basis for a neural signal. This is both a strength and a shortcoming of our previous study. We did not show that the neurons' firing correlates with the confidence behaviour on a trial-by-trial basis, and in this sense, we did not establish a neural correlate of an observed behaviour. On the other hand, there is a long and troubled history of labelling neurophysiological signals with psychological concepts based on a behavioural correlate alone [108]. Indeed, our neurons may also be correlated with 'anxiety', 'arousal' or 'exploration'. And in fact, these concepts can be related to different uses of uncertainty, and may all turn out to be neural correlates in some behavioural conditions. Rather our interpretation hinges on the only class of computational mechanisms that we found could successfully explain the observed firing patterns. This also implies that our claims can be disproved, either by showing that alternative models, without computing confidence, can also account for our data, or make predictions based on these models that are inconsistent with our data.

Although we use decision confidence in the formally defined sense, as the probability that a choice was correct based on the available evidence, this definition overlaps to some extent with our intuitive notion about confidence. Nevertheless, it will be important to directly assess how formal definitions of confidence and implicit confidence reports by animals correspond to the human notion of subjective confidence [109].

5. BEHAVIOURAL USES OF CONFIDENCE ESTIMATES

Until now, we discussed confidence in a limited context focusing on explicit or implicit reports, and

argued that formalized notions based on statistics provide a useful way forward. Placing this topic in a broader context, there is a vast literature in neuroscience, psychology, economics and related fields showing that uncertainty and confidence are critical factors in understanding behaviour [74,76,110–115]. Most of these fields use formal notions of confidence, and while impacting behaviour in lawful ways this need not always correlate with a conscious sense of confidence. To highlight this dissociation, we briefly point to a set of examples where confidence signals are used to guide behaviour in the apparent absence of awareness. Interestingly, in these cases, humans and non-human animals seem to be on par in the uses of uncertainty. Note, however, that in several of the examples although the requisite monitoring processes might use uncertainty, an explicit report may not be available. It will be valuable to examine how the uses of uncertainty and confidence in these behavioural situations are related to the metacognitive notion of confidence [116].

(a) *Foraging and leaving decisions*
As animals search for food they must continually assess the quality of their current location and the uncertainty about future possibilities to find new food items [117]. In other words, they must continually decide whether to stay or to go, depending on their level of confidence in the current location or 'patch'. Behavioural observations suggest that the time an animal spends at a particular patch depends not only the mean amount of food but also the variability [118,119]. Optimal foraging theory can account for these features by incorporating information about variability and other costs [117,120]. The patch allocation, i.e. the time animals spend at a particular location, should depend on the uncertainty of the estimate of its value [117]. Similarly, our 'decision restart' and 'leaving decision' tasks provide a psychophysical instantiation of a foraging decision: whether to stay at the reward port or leave and to start a new trial. The optimal solution here also depends on uncertainty concerning the immediately preceding perceptual decision, which determines the outcome. Therefore, foraging decisions are an example where uncertainty estimates are directly turned into actions, an online use of uncertainty as a decision variable.

(b) *Active learning and driving knowledge acquisition*
When you are not confident about something, it is a good time to learn. A subfield of metacognition refers to this as 'judgements of learning' [121,122]. The notion is that in order to figure out how much and what to learn, one needs to have meta-representations [123,124]. Interestingly, the field of machine learning in computer science uses a very similar but quantitative version of this insight. Statistical learning theory proposes that 'active learners' use not only reinforcements but also their current estimates of uncertainty to set the size of updates, i.e. learn more when uncertain and less when certain [125]. For instance, the Kalman-filter captures the insight that learning rate ought to

vary with uncertainty [126]. Simplified versions of the Kalman-filter have been used to account for a range of findings in animal learning theory about how stimulus salience enhances learning [127–130].

One of the key uses of estimating uncertainty or knowing your confidence is to drive information seeking behaviour so as to reduce the level of uncertainty. This is related both to foraging decisions as well as the active-learning examples given immediately above. The basic idea here is that the value of information is related to the uncertainty of the agent [95]. When the agent is very confident about the state of the world, then the value of information is low [131–134]. When the agent is less confident, then the value of information is high. Thus, when faced with a decision of how much time to allocate to information gathering or a decision between exploiting current information versus acquiring a new piece of information, we would expect that a representation of uncertainty might be particularly useful [135,136]. In other words, we expect the value of exploration to decrease proportionately with the current confidence in that piece of information.

(c) *Statistical inference and multi-sensory integration*

Perhaps the most ubiquitous and important use of uncertainty is in the process of statistical inference: using pieces of partly unreliable evidence to infer things about the world [7]. In principle, probability theory, or more specifically Bayesian inference tells us how one ought to reason in the face of uncertainty [9]. What this theory says in a nutshell is that evidence must be weighted according to its confidence (or inversely according to its uncertainty). There are a growing number of examples of statistically optimal behavioural performance (in the sense of correctly combined evidence), mainly in humans [137–139]. Since these examples involve primarily 'low-level' sensorimotor tasks to which humans often have no explicit access, it is not clear whether explicit ('metacognitive') access to confidence estimates are relevant. For example, a tennis player is unlikely to be able to report his relative confidence in the prior, his expectation about where the ball tends to fall, compared to his confidence in the present likelihood, the perceptual evidence about the ball's current trajectory. Nevertheless, his swing is accurate. It has been suggested that such problems reflect a probabilistic style of neural computation, one that would implicitly rather than explicitly represent confidence [10–12,140,141].

6. SUMMARY

Confidence judgements appear to us as personal, subjective reports about beliefs, generated by a process of apparent self-reflection. If so, then could animals also experience a similar sense of certainty? And is it even possible to ask this question as a pragmatic neuroscientist with the hope of finding an answer?

Undeniably, the concepts of 'confidence' and 'uncertainty' have established meanings in the context of human subjective experience. The importance of these concepts greatly motivates our research and therefore

it is important to assess the relationship between formally defined measures and the subjective entity in humans. Of course, it is impossible for us to know whether animals, in any of the tasks discussed above, ever feel the same sense of uncertainty that we humans do. In fact, some philosophers argue that it is impossible for us to know whether anyone else experiences the same sense of uncertainty—the problem of other minds. But in practice, it appears that verbal sharing of confidence information in humans can achieve the same goal of metacognitive alignment [142].

Before we can approach these vexing questions from a scientific perspective, it is important to establish that there is no justification in having distinct rules for interpreting human and animal experiments. Behavioural experiments, both in humans and in animals, need to be interpreted based on observables and not subjective experiences only accessible via introspection. This demands a behaviourist perspective but also new tools to go beyond old-fashioned, denialist behaviourism so that we are able to study a variable that is not directly accessible to measurement. This is possible using model-based approaches that enable one to link hidden, internal variables driving behaviour to external, observable variables in a quantitative manner. Such formal approaches not only enable us to drop semantic definitions, but also to go beyond fruitless debates. Models are concrete: they can be tested, disproved and iteratively improved, moving the scientific debate forward.

Here, we outlined an approach to studying confidence predicated on two pillars: an appropriately designed behavioural task to elicit implicit reports of confidence, and a computational framework to interpret the behavioural and neuronal data. Establishing a confidence-reporting behaviour requires us to incentivize animals to use confidence, for instance, by enabling animals to collect more reward or seek out valuable information based on confidence. We saw that in order to interpret behaviour and rule out alternative explanations, it is crucial to use tasks where data about the choice and the confidence associated with it are collected simultaneously in the same trial. Moreover, it is advantageous, although not required, that the confidence report is a graded variable and the task provides a large number of trials for quantitative analysis. To begin to infer behavioural algorithms for how confidence may have been computed (or whether it was), we presented a normative theoretical framework and several computational models.

In so doing, we have tried to lift the veil of this murky, semantically thorny subject. By showing that confidence judgements need not involve mysterious acts of self-awareness but something more humble like computing the distance between two neural representations, we hope to have taken a step towards reducing the act of measuring the quality of knowledge to something amenable to neuroscience, just as the notion of *subjective value* and its ilk have been [143–146]. Indeed, recent studies on the neural basis of confidence have brought a neurobiological dawn to this old subject [17,18,36,49,89]. We also believe that as a consequence of this demystification, animals may be put on a more even footing with

humans, at least with respect to the confidence-reporting variety of metacognition. Yet, this may reflect as much a humbling of our human abilities as a glorification of the animal kingdom.

REFERENCES

1 Flavell, J. H. 1979 Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *Am. Psychol.* **34**, 906–911. (doi:10.1037/0003-066X.34.10.906)

2 Metcalfe, J. & Shimamura, A. P. (eds) 1994 *Metacognition: knowing about knowing*. Cambridge, MA: MIT Press.

3 Terrace, H. S. & Metcalfe, J. (eds) 2005 *The missing link in cognition: origins of self-reflective consciousness*. New York, NY: Oxford University Press.

4 Metcalfe, J. 2008 Evolution of metacognition. In *Handbook of metamemory and memory* (eds J. Dunlosky & R. J. Bjork), pp. 29–46. Hove, UK: Psychology Press.

5 Smith, J. D., Shields, W. E. & Washburn, D. A. 2003 The comparative psychology of uncertainty monitoring and metacognition. *Behav. Brain Sci.* **26**, 317–339; discussion 340–373.

6 Smith, J. D., Couchman, J. J. & Beren, J. 2012 The highs and lows of theoretical interpretation in animal-metacognition research. *Phil. Trans. R. Soc. B* **367**, 1297–1309. (doi:10.1098/rstb.2011.0366)

7 Cox, D. R. 2006 *Principles of statistical inference*. Cambridge, UK: Cambridge University Press.

8 Zemel, R. S., Dayan, P. & Pouget, A. 1998 Probabilistic interpretation of population codes. *Neural Comput.* **10**, 403–430. (doi:10.1162/089976698300017818)

9 Rao, R. P. N., Olshausen, B. A. & Lewicki, M. S. 2002 *Probabilistic models of the brain: perception and neural function*. Cambridge, MA: MIT Press.

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

11 Ma, W. J., Beck, J. M., Latham, P. E. & Pouget, A. 2006 Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438. (doi:10.1038/nn1790)

12 Rao, R. P. 2004 Bayesian computation in recurrent neural circuits. *Neural Comput.* **16**, 1–38. (doi:10.1162/08997660460733976)

13 Higham, P. A. 2007 No special K! A signal detection framework for the strategic regulation of memory accuracy. *J. Exp. Psychol. Gen.* **136**, 1–22. (doi:10.1037/0096-3445.136.1.1)

14 Johnson, D. M. 1939 Confidence and speed in the two-category judgment. Thesis, Columbia University, NY.

15 Festinger, L. 1943 Studies in decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *J. Exp. Psychol.* **32**, 291–306. (doi:10.1037/h0056685)

16 Baranski, J. V. & Petrusic, W. M. 1994 The calibration and resolution of confidence in perceptual judgments. *Percept. Psychophys.* **55**, 412–428. (doi:10.3758/BF03205299)

17 Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J. & Rees, G. 2010 Relating introspective accuracy to individual differences in brain structure. *Science* **329**, 1541–1543. (doi:10.1126/science.1191883)

18 Yokoyama, O. *et al.* 2010 Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neurosci. Res.* **68**, 199–206. (doi:10.1016/j.neures.2010.07.2041)

19 Gigerenzer, G., Hoffrage, U. & Kleinbolting, H. 1991 Probabilistic mental models: a Brunswikian theory of confidence. *Psychol. Rev.* **98**, 506–528. (doi:10.1037/0033-295X.98.4.506)

20 Klayman, J., Soll, J. B., González-Vallejo, C. & Barlas, S. 1999 Overconfidence: it depends on how, what, and whom you ask. *Organ. Behav. Hum. Decis. Process.* **79**, 216–247. (doi:10.1006/obhd.1999.2847)

21 Finn, B. 2008 Framing effects on metacognitive monitoring and control. *Mem. Cognit.* **36**, 813–821. (doi:10.3758/MC.36.4.813)

22 Angell, F. 1907 On judgments of 'like' in discrimination experiments. *Am. J. Psychol.* **18**, 253–260. (doi:10.2307/1412416)

23 Watson, C. S., Kellogg, S. C., Kawanishi, D. T. & Lucas, P. A. 1973 The uncertain response in detection-oriented psychophysics. *J. Exp. Psychol.* **99**, 180–185. (doi:10.1037/h0034736)

24 Woodworth, R. S. 1938 *Experimental psychology*. New York, NY: Henry Holt and Company.

25 Peirce, C. S. & Jastrow, J. 1885 On small differences of sensation. *Mem. Natl Acad. Sci.* **3**, 73–83.

26 George, S. S. 1917 Attitude in relation to the psychophysical judgment. *Am. J. Psychol.* **28**, 1–37. (doi:10.2307/1412939)

27 Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R. & Erb, L. 1995 The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *J. Exp. Psychol. Gen.* **124**, 391–408. (doi:10.1037/0096-3445.124.4.391)

28 Shields, W. E., Smith, J. D. & Washburn, D. A. 1997 Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same–different task. *J. Exp. Psychol. Gen.* **126**, 147–164. (doi:10.1037/0096-3445.126.2.147)

29 Smith, J. D., Shields, W. E., Schull, J. & Washburn, D. A. 1997 The uncertain response in humans and animals. *Cognition* **62**, 75–97. (doi:10.1016/S0010-0277(96)00726-3)

30 Shields, W. E., Smith, J. D., Guttmannova, K. & Washburn, D. A. 2005 Confidence judgments by humans and rhesus monkeys. *J. Gen. Psychol.* **132**, 165–186.

31 Beran, M. J., Smith, J. D., Redford, J. S. & Washburn, D. A. 2006 Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *J. Exp. Psychol. Anim. Behav. Process.* **32**, 111–119. (doi:10.1037/0097-7403.32.2.111)

32 Sole, L. M., Shettleworth, S. J. & Bennett, P. J. 2003 Uncertainty in pigeons. *Psychon. Bull. Rev.* **10**, 738–745. (doi:10.3758/BF03196540)

33 Smith, J. D., Beran, M. J., Redford, J. S. & Washburn, D. A. 2006 Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring. *J. Exp. Psychol. Gen.* **135**, 282–297. (doi:10.1037/0096-3445.135.2.282)

34 Sutton, R. S. & Barto, A. G. 1998 *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.

35 Hampton, R. R. 2001 Rhesus monkeys know when they remember. *Proc. Natl Acad. Sci. USA* **98**, 5359–5362. (doi:10.1073/pnas.071600998)

36 Kiani, R. & Shadlen, M. N. 2009 Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* **324**, 759–764. (doi:10.1126/science.1169405)

37 Inman, A. & Shettleworth, S. J. 1999 Detecting metamemory in nonverbal subjects: a test with pigeon. *J. Exp. Psychol. Anim. Behav. Process.* **25**, 389–395. (doi:10.1037/0097-7403.25.3.389)

38 Teller, S. A. 1989 Metamemory in the pigeon: prediction of performance on a delayed matching to sample task. Thesis, Reed College, Portland, OR.

39 Sutton, J. E. & Shettleworth, S. J. 2008 Memory without awareness: pigeons do not show metamemory in delayed matching to sample. *J. Exp. Psychol. Anim. Behav. Process.* **34**, 266–282. (doi:10.1037/0097-7403. 34.2.266)

40 Foote, A. L. & Crystal, J. D. 2007 Metacognition in the rat. *Curr. Biol.* **17**, 551–555. (doi:10.1016/j.cub.2007. 01.061)

41 Persaud, N., McLeod, P. & Cowey, A. 2007 Post-decision wagering objectively measures awareness. *Nat. Neurosci.* **10**, 257–261. (doi:10.1038/nn1840)

42 Persaud, N. & McLeod, P. 2008 Wagering demonstrates subconscious processing in a binary exclusion task. *Conscious. Cogn.* **17**, 565–575. (doi:10.1016/j.concog.2007. 05.003)

43 Sahraie, A., Weiskrantz, L. & Barbur, J. L. 1998 Awareness and confidence ratings in motion perception without geniculo-striate projection. *Behav. Brain Res.* **96**, 71–77. (doi:10.1016/S0166-4328(97)00194-0)

44 Rajaram, S., Hamilton, M. & Bolton, A. 2002 Distinguishing states of awareness from confidence during retrieval: evidence from amnesia. *Cogn. Affect. Behav. Neurosci.* **2**, 227–235. (doi:10.3758/CABN.2.3.227)

45 Clifford, C. W., Arabzadeh, E. & Harris, J. A. 2008 Getting technical about awareness. *Trends Cogn. Sci.* **12**, 54–58. (doi:10.1016/j.tics.2007.11.009)

46 Fleming, S. M. & Dolan, R. J. 2010 Effects of loss aversion on post-decision wagering: implications for measures of awareness. *Conscious. Cogn.* **19**, 352–363. (doi:10.1016/j.concog.2009.11.002)

47 Schurger, A. & Sher, S. 2008 Awareness, loss aversion, and post-decision wagering. *Trends Cogn. Sci.* **12**, 209–210. (doi:10.1016/j.tics.2008.02.012)

48 Middlebrooks, P. G. & Sommer, M. A. 2011 Metacognition in monkeys during an oculomotor task. *J. Exp. Psychol. Learn Mem. Cogn.* **37**, 325–337. (doi:10. 1037/a0021611)

49 Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. 2008 Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231. (doi:10.1038/nature07200)

50 Call, J. & Carpenter, M. 2001 Do apes and children know what they have seen? *Anim. Cogn.* **3**, 207–220. (doi:10.1007/s100710100078)

51 Hampton, R. R., Zivin, A. & Murray, E. A. 2004 Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Anim. Cogn.* **7**, 239–246. (doi:10.1007/s10071-004-0215-1)

52 Radecki, C. M. & Jaccard, J. 1995 Perceptions of knowledge, actual knowledge, and information search behavior. *J. Exp. Soc. Psychol.* **31**, 107–138. (doi:10. 1006/jesp.1995.1006)

53 Basile, B. M., Hampton, R. R., Suomi, S. J. & Murray, E. A. 2009 An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). *Anim. Cogn.* **12**, 169–180. (doi:10.1007/s10071-008-0180-1)

54 Brauer, J., Call, J. & Tomasello, M. 2007 Chimpanzees really know what others can see in a competitive situation. *Anim. Cogn.* **10**, 439–448. (doi:10.1007/s10071-007-0088-1)

55 Suda-King, C. 2008 Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Anim. Cogn.* **11**, 21–42. (doi:10.1007/s10071-007-0082-7)

56 Brauer, J., Call, J. & Tomasello, M. 2004 Visual perspective taking in dogs (*Canis familiaris*) in the presence of barriers. *Appl. Anim. Behav. Sci.* **88**, 299–317. (doi:10.1016/j.applanim.2004.03.004)

57 Brauer, J., Kaminski, J., Riedel, J., Call, J. & Tomasello, M. 2006 Making inferences about the location of hidden food: social dog, causal ape. *J. Comp. Psychol.* **120**, 38–47. (doi:10.1037/0735-7036.120.1.38)

58 Bromberg-Martin, E. S. & Hikosaka, O. 2009 Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119–126. (doi:10.1016/j.neuron.2009.06.009)

59 Shettleworth, S. J. & Sutton, J. E. 2003 Metacognition in animals: it's all in the methods. *Behav. Brain Sci.* **23**, 353–354.

60 Smith, J. D., Beran, M. J., Couchman, J. J. & Coutinho, M. V. 2008 The comparative study of metacognition: sharper paradigms, safer inferences. *Psychon. Bull. Rev.* **15**, 679–691. (doi:10.3758/PBR.15.4.679)

61 Hampton, R. R. 2009 Multiple demonstrations of metacognition in nonhumans: converging evidence or multiple mechanisms? *Comp. Cogn. Behav. Rev.* **4**, 17–28.

62 Schwartz, B. L. & Metcalfe, J. 1994 Methodological problems and pitfalls in the study of human metacognition. In *Metacognition: knowing about knowing* (eds J. Metcalfe & A. Shimamura), pp. 93–113. Cambridge, MA: MIT Press.

63 Kornell, N. 2009 Metacognition in humans and animals. *Curr. Dir. Psychol. Sci.* **18**, 11–15. (doi:10.1111/j.1467-8721.2009.01597.x)

64 Washburn, D. A., Smith, J. D. & Shields, W. E. 2006 Rhesus monkeys (*Macaca mulatta*) immediately generalize the uncertain response. *J. Exp. Psychol. Anim. Behav. Process.* **32**, 185–189. (doi:10.1037/0097-7403.32.2.185)

65 Kornell, N., Son, L. K. & Terrace, H. S. 2007 Transfer of metacognitive skills and hint seeking in monkeys. *Psychol. Sci.* **18**, 64–71. (doi:10.1111/j.1467-9280. 2007.01850.x)

66 Metcalfe, J. 2004 Drawing the line on metacognition. *Behav. Brain Sci.* **26**, 350–351.

67 Glimcher, P. W. 2011 *Foundations of neuroeconomic analysis*. New York, NY: Oxford University Press.

68 Geeold, J. I. & Shadlen, M. N. 2001 Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16. (doi:10.1016/S1364-6613 (00)01567-9)

69 Dayan, P. & Daw, N. D. 2008 Decision theory, reinforcement learning, and the brain. *Cogn. Affect Behav. Neurosci.* **8**, 429–453. (doi:10.3758/CABN.8.4.429)

70 Juslin, P. & Olsson, H. 1997 Thurstonian and Brunswikian origins of uncertainty in judgment: a sampling model of confidence in sensory discrimination. *Psychol. Rev.* **104**, 344–366. (doi:10.1037/0033-295X.104.2.344)

71 Vickers, D. & Pietsch, A. 2001 Decision making and memory: a critique of Juslin and Olsson's (1997) sampling model of sensory discrimination. *Psychol. Rev.* **108**, 789–804. (doi:10.1037/0033-295X.108.4.789)

72 Knight, F. 1921 *Risk, ambiguity, and profit*. Boston, MA: Houghton, Mifflin.

73 Hsu, M., Bhatt, M., Adolphs, R., Tranel, D. & Camerer, C. F. 2005 Neural systems responding to degrees of uncertainty in human decision-making. *Science*, **310**, 1680–1683. (doi:10.1126/science.1115327)

74 Glimcher, P. W. 2003 *Decisions, uncertainty, and the brain: the science of neuroeconomics*. Cambridge, MA: MIT Press.

75 McCoy, A. N. & Platt, M. L. 2005 Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* **8**, 1220–1227. (doi:10.1038/nn1523)

76 Platt, M. L. & Huettel, S. A. 2008 Risky business: the neuroeconomics of decision making under uncertainty. *Nat. Neurosci.* **11**, 398–403. (doi:10.1038/nn2062)

77 Green, D. M. & Swets, J. A. 1966 *Signal detection theory and psychophysics*. New York, NY: Wiley.

78 Parker, A. J. & Newsome, W. T. 1998 Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* **21**, 227–277. (doi:10.1146/annurev.neuro.21.1.227)

79 Vickers, D. 1970 Evidence for an accumulator model of psychophysical discrimination. *Ergonomics* **13**, 37–58. (doi:10.1080/00140137008931117)

80 Moreno-Bote, R. 2010 Decision confidence and uncertainty in diffusion models with partially correlated neuronal integrators. *Neural Comput.* **22**, 1786–1811. (doi:10.1162/neco.2010.12-08-930)

81 Rolls, E. T., Grabenhorst, F. & Deco, G. 2010 Decision-making, errors, and confidence in the brain. *J. Neurophysiol.* **104**, 2359–2374. (doi:10.1152/jn.00571.2010)

82 Insabato, A., Pannunzi, M., Rolls, E. T. & Deco, G. 2010 Confidence-related decision making. *J. Neurophysiol.* **104**, 539–547. (doi:10.1152/jn.01068.2009)

83 Sollich, P. 2002 Bayesian methods for support vector machines: evidence and predictive class probabilities. *Mach. Learn.* **46**, 21–52. (doi:10.1023/A:1012489924661)

84 Mazurek, M. E., Roitman, J. D., Ditterich, J. & Shadlen, M. N. 2003 A role for neural integrators in perceptual decision making. *Cereb. Cortex* **13**, 1257–1269. (doi:10.1093/cercor/bhg097)

85 Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. 2006 The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765. (doi:10.1037/0033-295X.113.4.700)

86 Vickers, D. & Packer, J. 1982 Effects of alternating set for speed or accuracy on response time, accuracy and confidence in a unidimensional discrimination task. *Acta Psychol. (Amst).* **50**, 179–197. (doi:10.1016/0001-6918(82)90006-3)

87 Tong, S. & Koller, D. 2002 Support vector machine active learning with applications to text classification. *J. Mach. Learn. Res.* **2**, 45–66.

88 Shenoy, P., Kepecs, A. & Rao, R. 2009 Active updating of decision boundaries in rats can be explained using Bayesian classifiers. In *Computational and systems neuroscience*. Salt Lake City, UT: Frontiers in Systems Neuroscience.

89 Rolls, E. T., Grabenhorst, F. & Deco, G. 2010 Choice, difficulty, and confidence in the brain. *Neuroimage* **53**, 694–706. (doi:10.1016/j.neuroimage.2010.06.073)

90 Rolls, E. T. & Deco, G. 2010 *The noisy brain: stochastic dynamics as a principle of brain function*. New York, NY: Oxford University Press.

91 Wang, X. J. 2002 Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968. (doi:10.1016/S0896-6273(02)01092-9)

92 Timmermans, B., Schilbach, L., Pasquali, A. & Cleeremans, A. 2012 Higher-order thoughts in action: consciousness as an unconscious re-description process. *Phil. Trans. R. Soc. B* **367**, 1412–1423. (doi:10.1098/rstb.2011.0421)

93 Juslin, P. & Olsson, H. 1999 Computational models of subjective probability calibration. In *Judgment and decision making: neo-Brunswikian and process-tracing approaches* (eds P. Juslin & H. Montgomery), pp. 67–95. Hove, UK: Psychology Press.

94 Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. 2006 Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879. (doi:10.1038/nature04766)

95 Cohen, J. D., McClure, S. M. & Yu, A. J. 2007 Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil. Trans. R. Soc. B* **362**, 933–942. (doi:10.1098/rstb.2007.2098)

96 Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. 2006 Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045. (doi:10.1038/nature05051)

97 Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J. & Schultz, W. 2009 Risk-dependent reward value signal in human prefrontal cortex. *Proc. Natl Acad. Sci. USA* **106**, 7185–7190. (doi:10.1073/pnas.0809599106)

98 Corrado, G. & Doya, K. 2007 Understanding neural coding through the model-based analysis of decision making. *J. Neurosci.* **27**, 8178. (doi:10.1523/JNEUROSCI.1590-07.2007)

99 Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J. & Frackowiak, R. S. 2004 Temporal difference models describe higher-order learning in humans. *Nature* **429**, 664–667. (doi:10.1038/nature02581)

100 Sugrue, L. P., Corrado, G. S. & Newsome, W. T. 2004 Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787. (doi:10.1126/science.1094765)

101 Barraclough, D. J., Conroy, M. L. & Lee, D. 2004 Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* **7**, 404–410. (doi:10.1038/nn1209)

102 Lau, B. & Glimcher, P. W. 2005 Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84**, 555–579. (doi:10.1901/jeab.2005.110-04)

103 Rolls, E. T. & Grabenhorst, F. 2008 The orbitofrontal cortex and beyond: from affect to decision-making. *Prog. Neurobiol.* **86**, 216–244. (doi:10.1016/j.pneurobio.2008.09.001)

104 Schoenbaum, G. & Roesch, M. 2005 Orbitofrontal cortex, associative learning, and expectancies. *Neuron* **47**, 633–636. (doi:10.1016/j.neuron.2005.07.018)

105 Elliott, R., Dolan, R. J. & Frith, C. D. 2000 Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb. Cortex* **10**, 308–317. (doi:10.1093/cercor/10.3.308)

106 Wallis, J. D. 2007 Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci.* **30**, 31–56. (doi:10.1146/annurev.neuro.30.051606.094334)

107 Mainen, Z. F. & Kepecs, A. 2009 Neural representation of behavioral outcomes in the orbitofrontal cortex. *Curr. Opin. Neurobiol.* **19**, 84–91. (doi:10.1016/j.conb.2009.03.010)

108 Maunsell, J. H. 2004 Neuronal representations of cognitive state: reward or attention? *Trends Cogn. Sci.* **8**, 261–265. (doi:10.1016/j.tics.2004.04.003)

109 Sanders, J. & Kepecs, A. 2010 How certain are you? Explicit and implicit measures of decision confidence. In *Proc. 6th Annu. Conf. on Neuroeconomics, Evanston*, IL.

110 Kahneman, D., Slovic, P. & Tversky, A. 1982 *Judgment under uncertainty: heuristics and biases*. New York, NY: Cambridge University Press.

111 Kaelbling, L. P., Littman, M. L. & Cassandra, A. R. 1998 Planning and acting in partially observable stochastic domains. *Artif. Intell.* **101**, 99–134. (doi:10.1016/S0004-3702(98)00023-X)

112 Volz, K. G., Schubotz, R. I. & von Cramon, D. Y. 2005 Variants of uncertainty in decision-making and their neural correlates. *Brain Res. Bull.* **67**, 403–412. (doi:10.1016/j.brainresbull.2005.06.011)

113 Yu, A. J. & Dayan, P. 2005 Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692. (doi:10.1016/j.neuron.2005.04.026)

114 Daw, N. D., Niv, Y. & Dayan, P. 2005 Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711. (doi:10.1038/nn1560)

115 Averbeck, B. B., Sohn, J. W. & Lee, D. 2006 Activity in prefrontal cortex during dynamic selection of action sequences. *Nat. Neurosci.* **9**, 276–282. (doi:10.1038/nn1634)

116 Fleming, S. M., Dolan, R. J. & Frith, C. D. 2012 Meta-cognition: computation, biology and function. *Phil. Trans. R. Soc. B* **367**, 1280–1286. (doi:10.1098/rstb.2012.0021)

117 Stephens, D. W. & Krebs, J. R. 1986 *Foraging theory.* Princeton, NJ: Princeton University Press.

118 Bernstein, C., Kacelnik, A. & Krebs, J. R. 1988 Individual decisions and the distribution of predators in a patchy environment. *J. Anim. Ecol.* **57**, 1007–1026. (doi:10.2307/5108)

119 Hayden, B. Y., Pearson, J. M. & Platt, M. L. 2011 Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939. (doi:10.1038/nn.2856)

120 Kamil, A. C., Misthal, R. L. & Stephens, D. W. 1993 Failure of simple optimal foraging models to predict residence time when patch quality is uncertain. *Behav. Ecol.* **4**, 350–363. (doi:10.1093/beheco/4.4.350)

121 Nelson, T. O. & Narens, L. 1990 Metamemory: a theoretical framework and new findings. *Psychol. Learn. Motiv.* **26**, 125–173. (doi:10.1016/S0079-7421(08)60053-5)

122 Son, L. K. & Metcalfe, J. 2000 Metacognitive and control strategies in study-time allocation. *J. Exp. Psychol. Learn Mem. Cogn.* **26**, 204–221. (doi:10.1037/0278-7393.26.1.204)

123 Son, L. K. 2004 Spacing one's study: evidence for a metacognitive control strategy. *J. Exp. Psychol. Learn. Mem. Cogn.* **30**, 601–604. (doi:10.1037/0278-7393.30.3.601)

124 Son, L. K. & Sethi, R. 2006 Metacognitive control and optimal learning. *Cogn. Sci. A Multidiscip. J.* **30**, 759–774. (doi:10.1207/s15516709cog0000_74)

125 Schohn, G. & Cohn, D. 2000 Less is more: active learning with support vector machines. In *Proc. 17th Intl Conf. Machine Learning, Stanford, CA, June 29–July 2, 2000*, pp. 839–846. San Francisco, CA: Morgan Kaufmann.

126 Sutton, R. S. 1992 Gain adaptation beats least squares. In *Proc. 7th Yale Workshop on Adaptive and Learning Systems, 20–22 May*, pp. 161–166. New Haven, CT: Yale University.

127 Dayan, P. & Kakade, S. 2001 Explaining away in weight space. In *Advances in Neural Information Processing Systems 13* (eds T. K. Leen, T. G. Dietterich & V. Tresp), pp. 451–457. Cambridge, MA: MIT Press.

128 Pearce, J. M. & Hall, G. 1980 A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552. (doi:10.1037/0033-295X.87.6.532)

129 Courville, A. C., Daw, N. D. & Touretzky, D. S. 2006 Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* **10**, 294–300. (doi:10.1016/j.tics.2006.05.004)

130 Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A. & Daw, N. D. 2011 Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.* **14**, 1250–1252. (doi:10.1038/nn.2904)

131 Dearden, R., Friedman, N. & Russell, S. 1998 Bayesian Q-learning. In *Proc. 15th National Conf. on Artificial Intelligence*, pp. 761–768. Menlo Park, CA: AAAI.

132 Strens, M. 2000 A Bayesian framework for reinforcement learning. In *Proc. 17th Intl Conf. on Machine Learning, Stanford, CA*, pp. 943–950.

133 Bromberg-Martin, E. S. & Hikosaka, O. 2011 Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* **14**, 1209–1216. (doi:10.1038/nn.2902)

134 Rushworth, M. F. & Behrens, T. E. 2008 Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* **11**, 389–397. (doi:10.1038/nn2066)

135 Sallet, J. & Rushworth, M. F. 2009 Should I stay or should I go: genetic bases for uncertainty-driven exploration. *Nat. Neurosci.* **12**, 963–965. (doi:10.1038/nn0809-963)

136 Frank, M. J., Doll, B. B., Oas-Terpstra, J. & Moreno, F. 2009 Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat. Neurosci.* **12**, 1062–1068. (doi:10.1038/nn.2342)

137 Ernst, M. O. & Bülthoff, H. H. 2004 Merging the senses into a robust percept. *Trends Cogn. Sci.* **8**, 162–169. (doi:10.1016/j.tics.2004.02.002)

138 Körding, K. P. & Wolpert, D. M. 2004 Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247. (doi:10.1038/nature02169)

139 Graf, E. W., Warren, P. A. & Maloney, L. T. 2005 Explicit estimation of visual uncertainty in human motion processing. *Vis. Res.* **45**, 3050–3059. (doi:10.1016/j.visres.2005.08.007)

140 Körding, K. P. & Wolpert, D. M. 2006 Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* **10**, 319–326. (doi:10.1016/j.tics.2006.05.003)

141 Deneve, S. 2008 Bayesian spiking neurons I: inference. *Neural Comput.* **20**, 91–117. (doi:10.1162/neco.2008.20.1.91)

142 Bahrami, B., Olsen, K., Bang, D., Roepstroff, A., Rees, G. & Frith, C. 2012 What failure in collective decision-making tells us about metacognition. *Phil. Trans. R. Soc. B* **367**, 1350–1365. (doi:10.1098/rstb.2011.0420)

143 Montague, P. R. & Berns, G. S. 2002 Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284. (doi:10.1016/S0896-6273(02)00974-1)

144 Glimcher, P. W., Dorris, M. C. & Bayer, H. M. 2005 Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.* **52**, 213–256. (doi:10.1016/j.geb.2004.06.011)

145 Sugrue, L. P., Corrado, G. S. & Newsome, W. T. 2005 Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**, 363–375. (doi:10.1038/nrn1666)

146 Padoa-Schioppa, C. & Assad, J. A. 2006 Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226. (doi:10.1038/nature04676)