# Knowing how much you don't know: a neural organization of uncertainty estimates

*Dominik R. Bach[1,2], Raymond J. Dolan[1,2]*

Abstract | How we estimate uncertainty is important in decision neuroscience and has wide-ranging implications in basic and clinical neuroscience, from computational models of optimality to ideas on psychopathological disorders including anxiety, depression and schizophrenia. Empirical research in neuroscience, which has been based on divergent theoretical assumptions, has focused on the fundamental question of how uncertainty is encoded in the brain and how it influences behaviour. Here, we integrate several theoretical concepts about uncertainty into a decision-making framework. We conclude that the currently available evidence indicates that distinct neural encoding (including summary statistic-type representations) of uncertainty occurs in distinct neural systems.

*"Uncertainty and expectation are the joys of life. Security is an insipid thing."* William Congreve, 1670–1729, playwright.

Nothing is perfect and everything we sense, think and do is imbued with imprecision. Imprecision impedes us reaching our goals, but if we can estimate this imprecision we can use it to improve the success of our actions: we open our arms wider to pick up a faltering toddler than one who is sitting still, we look more carefully for other cars when driving in fog, and when securing our pensions we avoid shares that have a history of excessively large gains and losses.

Another term for imprecision is uncertainty. We can be uncertain about many different things and, as a consequence, uncertainty has been studied from many perspectives in the neurosciences, often utilizing divergent theoretical assumptions and empirical approaches. This Review unravels various strands of thinking relating to the concept of uncertainty and integrates them within an organized terminological and theoretical framework, in which our aim is to highlight possible common mechanisms of neural encoding and reveal areas that are in need of greater theoretical or empirical refinement.

We discuss studies of uncertainty in a framework that proposes four different processing levels that pertain to decision making and action planning (FIG. 1): sensory processing, state evaluation, rule identification

and outcome prediction. This framework reflects different types of variables about which we can be uncertain (BOX 1) rather than different sources of uncertainty (for example, environmental or internal). This provides a simplifying heuristic, although we acknowledge that these four levels do not necessarily form exclusive or coherent processing levels. To add to the complexity, within each of these levels uncertainty is often conceptualized in several distinct ways.

For each of these levels, we address three questions. The first is whether uncertainty guides behaviour. In many areas of behaviour and decision making, normative models prescribe the utilization of uncertainty estimates[1–6]. But behaviour of biological agents is not always optimal, and we can construe several situations in which optimal behaviour might mandate that uncertainty be ignored[7]. Thus, although there is good reason to expect that uncertainty influences behaviour, this needs to be empirically demonstrated in detail.

Second, if uncertainty guides behaviour, is there a distinct neural encoding of uncertainty? Again, the answer is not obvious from theoretical considerations alone. For example, behavioural phenomena that seemingly relate to uncertainty (for example, risk preferences) can be explained by divergent mechanisms[8,9]. Any investigation into neural representations of uncertainty has to take into account that noise is inherent in complex systems and can be greatly amplified by suboptimal inference[10]. Noise is not necessarily used

[1]Wellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London WC1N 3BG, UK.
[2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Luisenstraße 56, 10099 Berlin, Germany.
Correspondence to D.R.B.
e-mail: d.bach@ucl.ac.uk

to guide behaviour and hence a key aim is to identify neural representations of uncertainty over and above incidental noise that influence further neural computations and, indeed, behaviour. Although there are several theoretical proposals on how uncertainty might be encoded at a neuronal level[2,11], empirical investigations often assume that uncertainty is encoded in the firing rate of specific neuronal populations. We call this a summary-statistic encoding of uncertainty, as it summarizes uncertainty into a single number. The focus on

summary-statistic encoding is not driven by theoretical considerations alone (indeed, it is perhaps not the most plausible option from a theoretical viewpoint), it also reflects an assumption that such an encoding is probably not due to incidental noise.

The third question we address concerns the topographical distribution of uncertainty representations in the brain. For example, is there a distinct uncertainty representation about line orientation in the visual system and a distinct uncertainty representation about
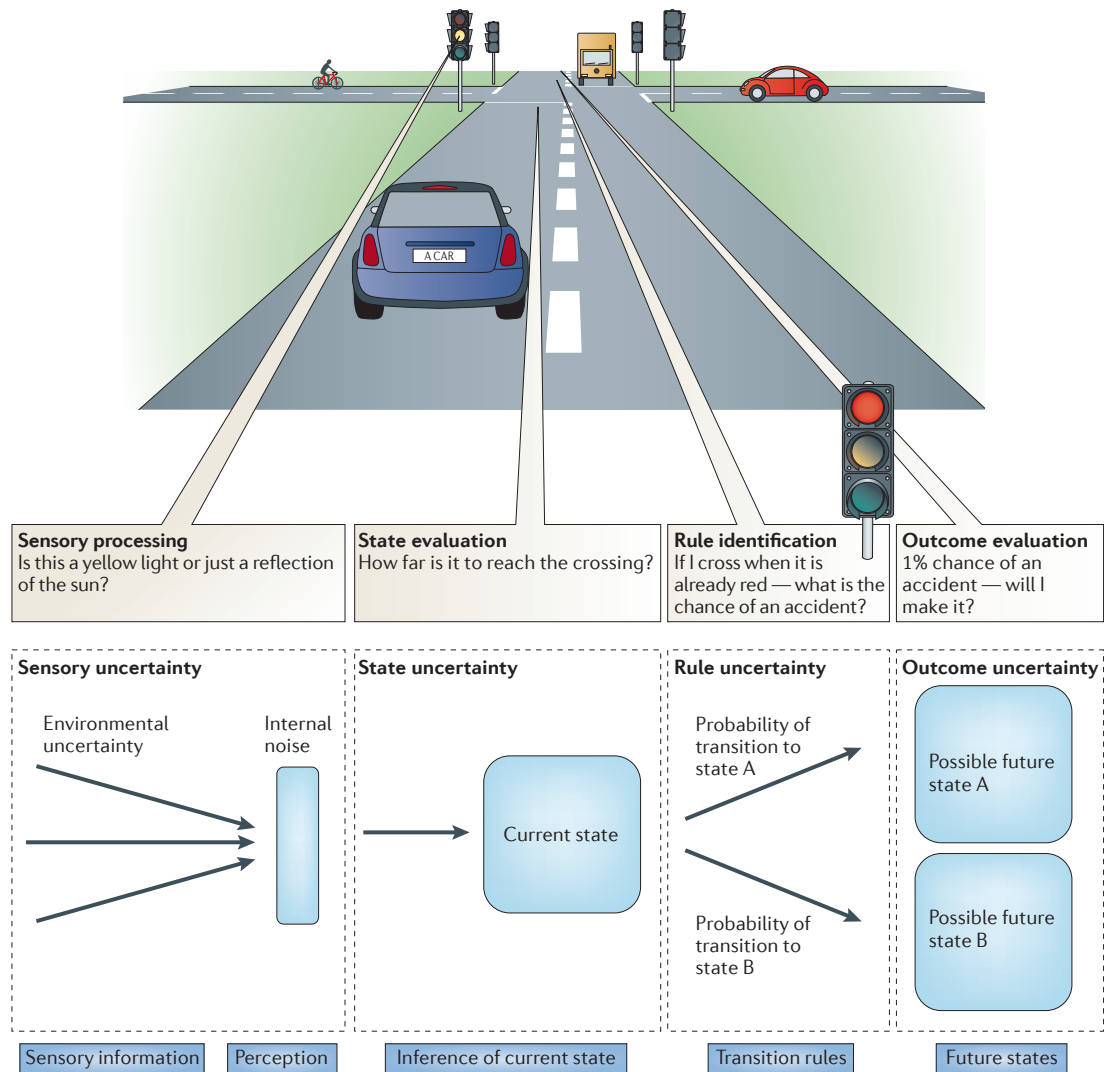


| Sensory processing | State evaluation | Rule identification | Outcome evaluation |
|---|---|---|---|
| Is this a yellow light or just a reflection of the sun? | How far is it to reach the crossing? | If I cross when it is already red — what is the chance of an accident? | 1% chance of an accident — will I make it? |

**Sensory uncertainty** · Environmental uncertainty · Internal noise

**State uncertainty** · Current state

**Rule uncertainty** · Probability of transition to state A · Probability of transition to state B

**Outcome uncertainty** · Possible future state A · Possible future state B

Sensory information · Perception · Inference of current state · Transition rules · Future states

Figure 1 | **Processing levels in decision making and action planning.** A real-life situation is used here to illustrate four processing levels in an action episode: a driver sees a yellow light while approaching a crossing. The first level concerns sensory processing: incoming information needs to be quantified or categorized. In this example, it needs to be determined whether the yellow light comes from the traffic light or is just a reflection of the sun. Sensory uncertainty refers to the fact that the driver knows only his noisy sensory input and not the actual status of the traffic light. The second level concerns state evaluation. In this example, the state of the environment (that is, the distance from the crossing) determines the chance of making it over the crossing before the light turns red. State uncertainty refers to the fact that inferring the distance to the crossing from a moving car is imprecise. The third level concerns rule identification. In this case, the driver will not make it over the crossing in time and plans to accelerate anyway, and the chances of having an accident constitute an example of a rule. Rule uncertainty in this example could arise from the driver having little experience of such a situation and receiving conflicting advice from two passengers ('don't worry' versus 'you'll kill all of us'). The fourth level concerns outcome prediction. Even if the driver knows the precise odds of having an accident, it is uncertain whether this possible accident will happen or not.

Summary statistic
A concise way of describing a set of observations without having to refer to each individual observation. Hence, the set of observations can be described with just a few values. For example, one for the location (for example, mean) and another for the dispersion, that is, uncertainty, (for example, variance).

decision outcomes in a decision-making network? Or is there a unified representation for different forms of uncertainty — a *de facto* canonical uncertainty representation? It has been suggested that uncertainty on a number of distinct decision-making variables is processed by a single brain area and is employed to elicit a common behavioural response[12]. Indeed, within computational models of decision making one can envisage collapsing different sources of uncertainty into a single quantity, and this might in principle occur in an

'uncertainty area'. However, theories that posit a hierarchically organized brain[1,2,6] usually assume that uncertainty about the value of a particular variable is bound to a representation of the value of that variable, thereby arguing against such a canonical uncertainty representation. A less strong version of a canonical uncertainty proposition is that common principles underlie the neural encoding of different forms of uncertainty, albeit organized in different locations within a neural hierarchy[11].

---

## Box 1 | Measures of uncertainty
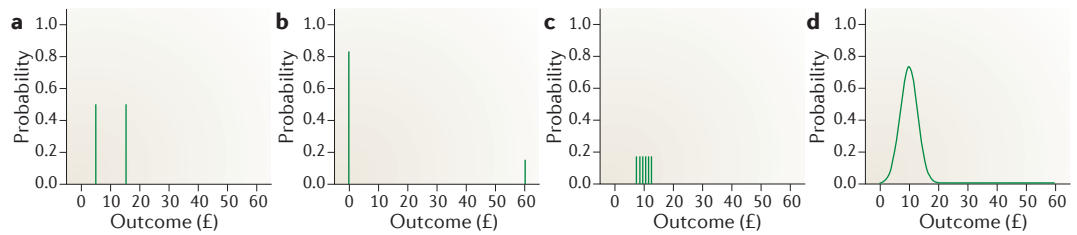


Uncertainty about a variable means that its true state or, for a quantitative variable, its magnitude is unknown. In other words, the variable in question can express one of several possible values. For each of these possible values, we can assign a probability to our degree of belief that it is the true value (see the figure). Loosely speaking, the most likely value of this variable forms our expectation of the value (for example, a value of 0 in the figure, panel **b**), and the dispersion in the distribution of possible values is one way of expressing our uncertainty on this expectation. Probabilistic computations can be performed using the complete distribution of possibilities, without an explicit measure of uncertainty. However, given a situation in which there are many possible values, a measure of uncertainty can greatly simplify such computations.

How can we quantify, or measure, uncertainty? In neuroscience, two principal approaches are used. The first approach is used when the variable in question takes several discrete states that are treated as nominal. That is, there is no scalar quantity associated with the states (for example, the colour of an object that can be red, blue or green) or when the scalar quantity is disregarded. Here, one can quantify uncertainty as the Shannon entropy[113] of a discrete probability distribution (that is, probability mass function), which quantifies how much information is gained if the true state of the variable is revealed (that is, the more uncertain the true value of the variable is, the more information is gained by knowing it).

The other principal approach, which is mainly used in economic theory, is to take a scalar quantity into account. For example, when we want to express that our uncertainty about winning either £40 or £60 from a toss of a coin is much lower than winning £0 or £100, even though the Shannon entropies of these lotteries are the same. In these cases, we can define uncertainty as dispersion of the probability distribution; for example, as variance or coefficient of variation[114]. This approach extends to continuous random variables that are governed by probability density functions.

A third, and more principled, approach is to measure uncertainty in continuous variables as their differential or relative entropy (Kullback–Leibler divergence) with respect to a reference distribution. Similar to dispersion measures, these entropy measures naturally take the scalar value of the variable into account. However, they have rarely been used in experimental research on the neural representation of uncertainty.

The figure illustrates several examples of probability functions and how to measure their associated uncertainty. The first example (**a**) shows a flip of a fair coin with probabilities of 0.5 and a win of £5 or £15 from heads or tails, respectively (Shannon entropy: 1 bit; variance 25 £²). The second example (**b**) shows a toss of a fair dice with a win of £60 from a six and nothing for the numbers one to five (Shannon entropy: 0.65 bit; variance 500 £²). Shannon entropy, which does not take into account outcome magnitude, is larger in panel **a**, but the variance of possible outcomes is larger in panel **b**. The third example (**c**) shows a toss of a fair dice, winning the amount of money shown by the number on the dice in £, plus £6.50 (Shannon entropy: 2.59 bit; variance 2.9 £²). This example has high Shannon entropy because the actual outcome is more unpredictable than in panels **a** and **b**, but there is little variance because possible outcomes are more similar than in the other examples. The final example (**d**) shows a probability density function for an unspecified example of continuous outcome possibilities. Here, uncertainty can, for example, be quantified as dispersion of the distribution.

A measure of uncertainty provides a basis for probing its behavioural and neural correlates. This is different from using a categorical approach, in which uncertainty is contrasted with certainty, to investigate behavioural and neural correlates of uncertainty. Such a categorical approach is related to theories which posit that uncertainty and certainty induce particular states of mind[115–117]. Hence, contrasting uncertainty with (near) certainty might highlight neural activity changes that are unrelated to uncertainty coding *per se*[118]. Attempts to disambiguate this confounding feature from a quantifiable encoding of uncertainty encoding are rare[119]. This Review therefore focuses on studies that continuously varied uncertainty on some variable.

---

## Uncertainty about sensory information

Imagine trying to negotiate an unfamiliar street in dark and misty weather. Incoming visual information needs to be quantitatively measured (for example, to gauge distances) and categorized (for example, to decipher letters on a street sign). Such sensory information is imbued with imprecision, and access to this imprecision might be beneficial (for example, equipped with this knowledge one might drive more carefully than usual). Studies have investigated sensory uncertainty mainly in three dominant, partially overlapping, domains: multisensory integration, sensorimotor control and unimodal sensory decision-making. To quantify sensory uncertainty, one approach is to infer the overall uncertainty due to stimulus uncertainty and internal noise from overt behaviour. Another is to use a measure of stimulus uncertainty alone; this is often implemented as a semi-quantitative measure termed task difficulty.

*Sensory uncertainty guides behaviour.* Multisensory integration experiments build on the idea that when we combine conflicting information (or cues), more uncertain information deserves less weight[13,14]. Indeed, humans and monkeys weight information sources according to their individual overall uncertainty in a near-optimal manner (that is, close to minimizing the error of the combined estimate). This has been demonstrated for a wide range of cue combinations from various modalities and under different response conditions, both when cue uncertainty is stationary[15-21] or experimentally varied[22-29]. Previous experience is an additional source of information (often termed prior information) that can be weighted and integrated with current sensory cues according to Bayes' theorem. Indeed, this is what humans appear to do, both under stationary conditions[30], and when sensory uncertainty[31,32] or prior uncertainty[33,34] are experimentally varied. Note that in these and other experiments[35-37], information integration is sometimes suboptimal, but, even in these cases, most of the time uncertainty does influence behaviour.

Behavioural sensitivity to sensory uncertainty is also evident in studies of sensorimotor control[13]. Optimal motor planning takes into account uncertainty in sensory information. For example, grip aperture should be wider when we are more uncertain about the position of an object to grip. This has been shown to occur for objects imbued with visual uncertainty[38] and for objects for which its position is uncertain owing to an imprecise coordinate transform between body and eye reference[39]. Furthermore, when humans are asked to point to a target, they are quicker to adjust their movement upon target change when the initial target has a higher position uncertainty, and they are also quicker when the final target has a lower position uncertainty[40]. In addition, a motor task such as catching a ball involves setting the optimal time point for starting the movement: if we do not observe the ball for long enough, sensory information is more uncertain; but if we start moving too late, the movement will be imbued with greater motor uncertainty (see below). Across different tasks, humans integrate visual and motor uncertainty in a near-optimal

manner, therefore showing that both visual and motor uncertainty guide behaviour under stationary conditions[41,42]. In another paradigm, human subjects combine estimates of their hand position based on previous proprioceptive and visual feedback with current visual feedback to adapt their hand movements. When current visual feedback is more uncertain, adaptation takes longer[43], and if previous feedback is more uncertain, adaptation is faster[43], as predicted by optimal integration.

A third strand of evidence for an effect of sensory uncertainty on behaviour comes from experiments on unimodal sensory decisions. Take an experiment in which monkeys perform a random-dot task (BOX 2) and saccade to indicate the direction of net dot motion, and they receive a reward for a correct response[44]. For half the trials, the animals can opt out of the 'bet' and instead saccade to a third target that yields a smaller, fixed reward. Opting out is advantageous when the fixed reward exceeds the expected reward from the random-dot task. This enables one to measure a monkey's estimate of expected reward — an estimate corresponding to the monkey's confidence that a decision would be correct. Crucially, monkeys learn to opt out more often when there is more uncertainty about net dot motion, indicating that the level of uncertainty guides sensory decision-making. Note that the stimulus configuration depends on the number of consistently moving dots, and the monkey brain might simply associate this configuration with a level of reward, rather than encode the uncertainty. To shed light on this point, recording of neural activity can be utilized (see below). Similar to findings from this study[44], when humans make sensory decisions regarding two noisy stimuli — one with high and one with low uncertainty — the difference in uncertainty between these two stimuli correlates with the probability of choosing the less noisy stimulus[45].

In summary, there is compelling evidence from a range of experimental contexts and tasks to suggest that sensory uncertainty, which is inherent in the environment or due to internal noise, influences behaviour, thus often leading to a more optimal response.

*Encoding of sensory uncertainty.* If sensory uncertainty guides behaviour, how is this uncertainty neurally encoded? This question has been explored mainly within experiments on unimodal sensory decisions. One study in rats[46] varied the ratio of two odours in an odour mixture and the animals had to correctly categorize the resulting odour in order to receive a reward. An increased certainty (that is, at a high ratio of one odour over another) increased the animals' willingness to wait for the reward and was associated with increased/decreased firing in two distinct populations of orbitofrontal cortex (OFC) neurons, respectively. Note that the expected reward in this task depended on stimulus uncertainty because, on average, animals received less reward for decisions imbued with more uncertainty. Hence, changes in OFC firing might reflect expected reward rather than uncertainty. However, on a trial-by-trial basis, neural firing was not well-explained

by learning history as would be required by this alternative explanation.

In another example, the coherence of motion displays (BOX 2) is encoded in firing rates of neurons in the macaque visual system (for example, the middle temporal[47] and medial superior temporal[48] visual cortex), and this could represent an uncertainty signal. However, it could also simply reflect an encoding of the number of consistently moving dots — unless a precise relationship with behaviour is demonstrated. To establish such a relationship, many studies rely on temporal integrator models (BOX 2), which posit a neural mechanism that accumulates sensory evidence over time. The speed of evidence accumulation in these models is inversely correlated to the degree of uncertainty. In the aforementioned study in monkeys[44], neurons in the lateral intraparietal cortex (area LIP) encoded evidence accumulation in a manner that was predicted by a temporal integrator model, with neuronal activity in this brain region drifting more slowly when stimulus uncertainty was high. Importantly, variations in drift rate over and above the impact of stimulus uncertainty were most likely to be caused by internal noise in the sensors or in the neural circuits. These variations in drift-rate predicted whether or not the monkey opted out of making a decision. This suggests that opting out reflects the resulting uncertainty in a neural system that accumulates evidence for a decision, and is not due to a simple association of the stimulus configuration with lower reward (see above). It also suggests that a change over time in the pattern of LIP firing rates is not merely an incidental reflection of environmental noise but is crucial for using uncertainty to guide behaviour. However, it leaves unanswered the question of how precisely a change of neuronal firing rates is translated into behaviour.

In principle, electroencephalography recordings in human subjects provide a good method to assess uncertainty encoding in temporal integrator models, but its application has proven difficult. Temporal integrator models predict that stimulus uncertainty, response timing and response accuracy are related variables. However, an event-related potential (ERP) study in humans in which blended images of cars and faces had to be categorized (as either car or face) showed separate ERP components that reflected these supposedly related variables[49]. Furthermore, in another ERP study using morphed images of males and females that were noise-free, the uncertainty component did not increase when the male:female ratio was more ambiguous (which is one form of stimulus uncertainty); it only increased when noise (another form of stimulus uncertainty) was added[50]. This result suggests that these ERP components cannot be taken to be informative about sensory uncertainty in a broad sense, but are specific to processes that are induced by sensory uncertainty.

BOLD fMRI (blood oxygenation-level dependent functional MRI) responses to uncertainty can be ambiguous too. Temporal integrator models predict that when there is enough time to make a decision, neural firing should increase over time up to a threshold, and then cease. BOLD responses reflect integrated neuronal firing over time, and because the decision threshold is constant, this integral will be larger when the integration takes longer[51]. BOLD responses might therefore reflect the rate of change in firing of neurons in an integrator.

However, if there are distinct neural populations that fire more when there is more uncertainty (that

## Box 2 | Temporal integrator models in sensory decision-making
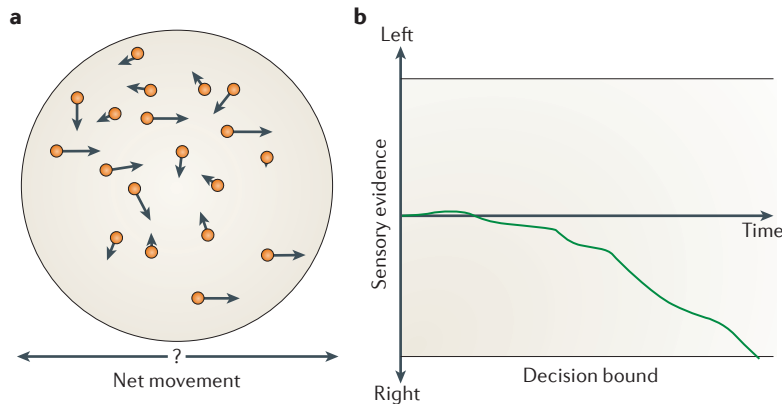


Sensory decision-making is often modelled by temporal integrator models (also termed sequential sampling, drift diffusion or bounded integrator models), which assume that sensory information is accumulated over time[120–122] and usually in the context of a two-alternative forced choice task. For example, imagine an array of orange dots in which 80% are moving randomly, and the remaining 20% are consistently moving to the right (see the figure, panel **a**). A subject's task is to decide whether the net movement of the dots is to the left or to the right. At each moment, the majority of dots could be — just by chance — moving to the right or to the left.

Temporal integrator models assume that we gather evidence over time: at each moment, the number of dots we see moving to the right side causes a signal. This could be a neuronal signal that results from summing up input from all neurons that are sensitive to a dot moving to the left, minus the signal from all neurons that are sensitive to a dot moving to the right. The random dot motion causes this signal to fluctuate, but it will slowly increase over time owing to the majority of dots moving consistently in one direction (see the figure, panel **b**). When this integrated signal reaches a certain bound (criterion), the subject is able to decide that the true movement direction is to the right.

The more random dots are added in this task, the shallower the slope of the evidence curve (that is, the lower the so-called drift rate), and the longer it takes to make a decision. At the same time, if the subject is forced to make a quick decision he/she will be less accurate as there is more noise[122,123]. Hence, within a single experimental context and up to a constant, the drift rate can be estimated from the reaction times and the accuracy of the response. Note that this constant can differ between experimental contexts owing to unknown properties of the putative neural integrator (for example, the bound), thereby rendering a comparison between the drift rate in two experiments impossible.

One might wonder why one would estimate the drift rate and not use response time or response accuracy as a proxy for uncertainty. Subjects tend to respond quickly both when there is near-zero sensory evidence and when there is near-complete sensory evidence[124]. Also, neural activity that is associated with response times, accuracy and task difficulty can dissociate[49,52]. On this basis, we suggest that it is problematic to use reaction times or accuracy alone as indicators of uncertainty. Estimates of the drift rate that are derived from both measures are more precise and can therefore serve as a useful means of inferring uncertainty, especially in cases of spontaneous fluctuations of uncertainty due to internal noise.

Of course, estimation of drift rates from reaction times and response accuracy crucially depends on the validity of the temporal integrator model, and the possible presence of a recently proposed urgency gating signal needs to be taken into account[120]. Furthermore, recent evidence from a study of decisions on taste stimuli in rats points to an integration of evidence over inputs rather than over time[107], and, in such cases, a temporal integrator model would be uninformative.

**Temporal integrator models**
Models that describe the accumulation of sensory evidence over time; for example, when viewing a noisy stimulus and having to decide on its identity.

**Environmental noise**
Random fluctuations in a measured signal that arises from the outside world. In the context of a sensory decision-making experiment, for example, this could be from noise in the stimulus.

is, as a summary-statistic encoding of uncertainty), these would also exhibit a greater BOLD signal. This distinction is important because, as mentioned before, a change over time in firing rate within a neural integrator is a reflection of environmental noise, but it is not necessarily used for further neural computations. Hence, it is important to investigate whether there is a relationship between BOLD responses and behaviour — or, indeed, a relationship with other neural computations.

BOLD correlates of sensory uncertainty have been reported in intraparietal sulcus, middle temporal visual cortex, frontal eye field, middle frontal gyrus, anterior cingulate cortex (ACC), supplementary motor areas and anterior insula[51]. Responses in most of these areas correlate with reaction times[51], which possibly reflects trial-by-trial fluctuations in sensory evidence accumulation, as in the aforementioned study in monkeys[44]. In a different study, BOLD responses were reported in supplementary eye fields, frontal eye fields, dorsolateral prefrontal cortex (Brodmann areas 9 and 46) and anterior insula[52], but they did not correlate with behaviour. Another study used a perceptual category learning task in which subjects had to assign distorted patterns of dots into one of two categories, each defined by a prototype[53]. The degree of distortion, which induces sensory uncertainty, correlated with BOLD responses in parietal cortex, cerebellum, dorsal ACC and supplementary motor cortex, inferior occipital/temporal gyrus, anterior insula and inferior frontal gyrus, frontal eye fields and precentral gyrus.

Another type of sensory uncertainty is uncertainty in perceptual categorization. This occurs not only if the sensory evidence is uncertain, but also if category boundaries are diffuse. This type of uncertainty has been investigated in a human neuroimaging experiment[54]. Here, categorization uncertainty was linked to enhanced BOLD responses in medial frontal gyrus, anterior insula, ventral striatum and dorsomedial thalamus[54].

Finally, two recent experiments have addressed the neural encoding of sensory uncertainty in multisensory integration. During the combination of tactile and visual information, increasing uncertainty in one modality suppresses primary sensory BOLD responses in the other modality. This suggests that sensory uncertainty influences neural computations, but does not reveal how uncertainty estimates are represented or used[55]. However, in the dorsal medial superior temporal cortex — a brain area that encodes the integration of vestibular and visual cues — uncertainty of cues from individual modalities and uncertainty from combined cues is represented in the slope of a neural population tuning curve[36]. This study[36] addressed cue integration and varied sensory uncertainty only on two levels, and so extending this approach to continuously varied uncertainty would be of great interest.

To summarize, neural representations of different forms of sensory uncertainty have been shown in several brain areas but, as yet, we lack a precise picture as to how these neural representations emerge. In part, this is due to methodological issues. Nevertheless,

experiments that demonstrate a relationship between putative neural uncertainty encoding and trial-to-trial behaviour have established important findings. First, for visual decisions, the slope of a neural evidence accumulator in parietal cortex seems to encode sensory uncertainty and is translated into behaviour[44]. Second, BOLD responses relating to sensory uncertainty and to behaviour have been demonstrated in frontal eye fields, primary motor cortex, middle frontal gyrus, anterior insula and anterior cingulate/ supplementary eye fields[51]. These observations could reflect the slope of neural evidence accumulation or a summary-statistic encoding of sensory uncertainty. In multisensory cue combination, it is possible that the slope of neural tuning curves for individual stimuli encodes individual uncertainty and is used for cue integration, which is a possibility that needs further investigation. The evidence indicates a specific neuronal encoding of sensory uncertainty in several brain areas; however, it remains to be shown how such coding is translated into behaviour.

## Uncertainty about state

Navigating an unknown street looking for a friend's house, we might wonder whether we have arrived in the correct borough. This is an example of uncertainty about the present state of the environment, over and above uncertainty in relation to sensory information. Such state uncertainty arises in a number of situations. First, in perception, in multisensory integration and in sensorimotor control, uncertainty regarding the physical state of our limbs or of objects in the environment often arises from previous or current sensory uncertainty (this is discussed in the section on sensory uncertainty).

Second, uncertainty on inferring the present context arises during reinforcement learning, which is often dependent on context or state[56]. This type of state uncertainty has not been theoretically or experimentally investigated. A closely related example of state uncertainty is provided in the framework of hierarchical defence systems[57–59]. An uncertain threat (for example, a predator odour) induces specific defensive behaviours in animals (for example, risk assessment and inhibition of ongoing behaviour). Uncertain threat can be re-conceptualized as uncertainty about the state of the environment (for example, is this an environment in which to expect a looming predator, or not?). In this context, state uncertainty can be thought of as inducing particular defensive behaviours in order to resolve this very uncertainty. As yet, uncertain threat in an ethological sense has not been quantitatively defined.

Third, computational neuroscience often models the environment as a series of discrete states. In many applications it is assumed that the present state is known, but one can construe situations in which it has to be explicitly inferred from incomplete information. This has been investigated in maze navigation, whereby the precise position is not known to the subject[60]. In addition, there is often uncertainty about the transition between the present and the next state, which may be governed by transition rules. Uncertainty on transition rules is

conceptualized in the section on rule uncertainty, and uncertainty on an upcoming state is discussed in the section on outcome uncertainty.

We are not aware of studies investigating whether uncertainty about the present state of the environment (over and above sensory uncertainty) guides behaviour. Neural representations have been investigated in a neuroimaging study that continuously varied state uncertainty[60]. This study addressed the problem of estimating one's position in a maze (that is, the current state of the environment) in order to reach a target position. BOLD responses correlating with this type of state uncertainty were observed in anterior prefrontal cortex (bilateral medial and superior frontal gyrus, and Brodmann areas 9 and 10).

In summary, it remains unclear how uncertainty on state inference is represented in the brain, and whether and how it informs behaviour. In particular, studies have so far mainly addressed state uncertainty resulting directly from sensory uncertainty. Few theoretical or empirical studies have focused on estimating uncertainty on discrete states of the world under conditions of precise (but incomplete) sensory information.

## Uncertainty about rules

Let us return to the scenario in which we are looking for our friend's house. We have concluded that we are in the correct street, but we do not know the house numbering system and, therefore, our chances of finding the correct house if we turn left. We can simplify and formalize this situation by defining the left turn as present state, and finding or not finding the house on this side of the street as the two possible successive states. The transition rule between the present and next state (that is, our chances of finding the house on this side of the street) depends on the house numbering system and is therefore uncertain. Such rule uncertainty is usually distinguished from uncertainty about which of the possible next states will actually occur (which is referred to as outcome uncertainty in this Review, see below).

Transition rules come in different flavours. Rules about the transition to a state with explicit value might be described as stimulus–outcome rules in classical (Pavlovian) conditioning, as action–outcome rules in instrumental (operant) conditioning or as lotteries (gambles or bets) in economic decision-making. There are also transition rules between discrete environmental states in many formal computational models and in other forms of associative learning. In addition to transition rules, co-occurrence rules between two events in associative learning can be uncertain.

Associative learning involves uncertainty about rules (until we have learned them), but there is little formalism to quantify rule uncertainty in this context. The concepts of unexpected uncertainty or risk prediction error (BOX 3) describe situations in which surprise leads to rule uncertainty, but these concepts do not include a quantification of the ensuing uncertainty. In a Bayesian framework, rule uncertainty can be quantified as the posterior uncertainty of rule prediction[61,62], but we are not aware of empirical studies that have examined the behavioural consequences, or representation, of such uncertainty.

An economic perspective deals with propositional situations that embody uncertainty about transition rules (BOX 4), namely the transition to a state with higher economic value. In such situations, a probability can be assigned to the degree of belief that a certain rule will be correct. The entropy (BOX 1) over these probabilities quantifies rule uncertainty in this case[63]. This is analogous to the posterior uncertainty of rule predictions in the aforementioned Bayesian framework, and constitutes a useful approach to experimentally control rule uncertainty.

*Rule uncertainty guides behaviour.* Several economic investigations have contrasted rule uncertainty with rule certainty, and this contrast probably confounds rule uncertainty with other features (BOX 4). In an attempt to continuously vary rule uncertainty, a combined behavioural and neuroimaging experiment used an economic decision-making paradigm[63] in which subjects made economic gambles on the likely occurrence of electric shocks. In each trial there were two possible action–outcome rules, and the probability that either of the two rules would be realized was varied continuously. This design embodies varying rule uncertainty: uncertainty is high when both rules are equally probable and is lower when it is likely that one of the two rules will be

---

## Box 3 | Uncertainty and prediction errors in reinforcement learning

Rule uncertainty and outcome uncertainty are important in reinforcement learning. Classical reinforcement learning theories (for example, the Rescorla–Wagner[125] theory or the Pearce–Hall[126] theory) assume that an association between two consecutive events can be altered by the difference between our prediction and the occurrence of the second event (that is, by the prediction error). Thus, if we predict a low probability of electric shock after a tone and do receive one, the tone–shock association strengthens. Conversely, if we predict a high probability of electric shock after a tone and do not receive one, the association weakens[125–127].

This assumption can explain that when a shock always follows a tone, the association will, after some time, precisely describe the transition rule so that there are no longer any prediction errors. By contrast, in the case of probabilistic reinforcement (for example, receiving a shock only 50% of the time), there will be a prediction error after each event, even if our predictions are correct on average and the 50% schedule never changes. In such cases, a prediction error does not mean that the rule has changed, and it would not be useful to update the association upon every prediction error. To account for this situation, one can quantify the expectation for the prediction error, and assume that there is a neural representation of this expected prediction error. Rule change should only be inferred if the actual prediction error is larger than expected.

This idea has been formalized as 'expected and unexpected uncertainty'[61], which represent outcome uncertainty and an unexpectedly large surprise that gives rise to rule uncertainty, respectively. A heuristic that could explain how an estimate of outcome uncertainty (that is, expected uncertainty) is acquired during reinforcement learning builds on prediction errors: estimates of outcome variance are updated by comparing the variance estimate and the instantaneous squared prediction error, normalized by the mean outcome[128]. This has been termed risk prediction error.

A re-analysis of data from an earlier study[89] on outcome uncertainty suggested that there was a neural representation of risk prediction error in the anterior insula[97]. However, the task design in this study did not provide an incentive for learning, such that risk predictions need not be updated. A recent study that provided an incentive for learning found that the risk prediction error was represented in inferior frontal gyrus, insula, anterior cingulate cortex, ventral striatum and amygdala[129]. Demonstrating such risk prediction errors argues for a reinforcement learning-type acquisition of outcome uncertainty estimates. How unexpected uncertainty, that is, a larger-than-expected prediction error, translates into rule uncertainty, still awaits formalization.

---

Box 4 | **Rule uncertainty and ambiguity in economic gambles**

Urn A              Urn B
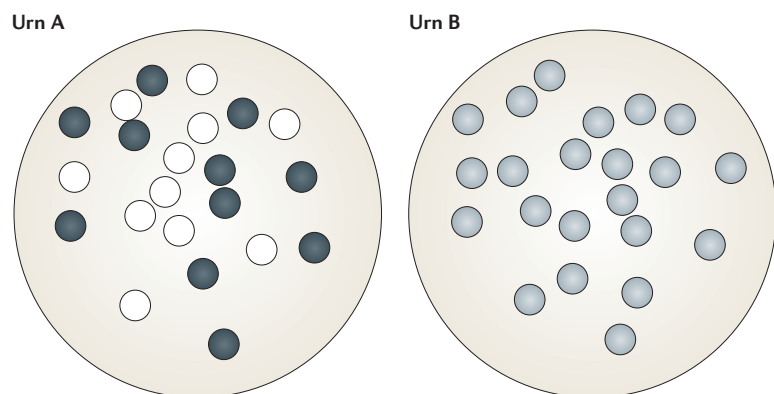


Imagine the following lottery: there is an urn (urn A) with 22 balls and half of these are black and the other half white, and so the winning chances for each colour are 50% (see the figure). Another urn (urn B) has the same number of balls, but the proportion of black and white balls is unknown (represented as grey balls in the figure). Would you rather bet on white from urn A or urn B? If your answer is urn A, classical economic theories would deduce that you think that the winning chances of white in urn B are lower than white in urn A (that is, lower than 50%).

Second question: would you rather bet on black in urn A or urn B? You might again answer urn A, and classical economic theories would conclude that you also think that black in urn B has winning chances lower than 50%. But that means that the probabilities in urn B add up to less than 100% — the famous Ellsberg paradox[7]. This thought experiment, building on earlier ideas[130], was among the first to illustrate the distinction between classical economic theory, which describes normative models of optimal behaviour, and findings from studies of real economic behaviour (later termed behavioural economics). Urn A constitutes a single-stage lottery, in which there is uncertainty about the two individual outcomes but the rules, that is, the probabilities of the two outcomes, are known. By contrast, urn B can be understood as a two-stage lottery, whereby the probabilities of the individual outcomes are governed by a distribution of probabilities, that is, by second-order probabilities.

There are 23 possibilities for the distribution of black and white balls in urn B. Classical economic theory posits that in the absence of further information, one should assume a uniform second-order probability distribution (that is, all 23 possible ball distributions in urn B have probability $P = 1/23$). For each possible ball distribution, one should multiply the ensuing first-order probabilities of a black or white draw, multiply them with its second-order probability and finally add up all probabilities for black and white across all ball distributions in order to derive expected first-order probabilities and make a decision. It turns out that for both urns, the expected outcome probabilities are the same (50% black, 50% white).

The situation in urn B has been termed ambiguity[7]. Empirical investigations have demonstrated that most people do not simply collapse the two stages of this bet (as prescribed by classical economic theory), but instead avoid ambiguity, even when this does not lead to (mathematically) optimal outcomes (for example, if one has to pay extra money to bet on urn A)[131–137].

Ambiguous and non-ambiguous gambles differ with respect to rule uncertainty, and restricting the range of possible outcome probabilities (that is, reducing rule uncertainty but keeping the outcome probabilities of the collapsed lottery constant) reduces ambiguity aversion[135,138]. However, they also differ across other factors, and these can explain ambiguity aversion under some circumstances. For example, ambiguity avoidance is stronger when decisions are made in public as opposed to in private[133,139], and this cannot be explained by aversion to rule uncertainty alone. Furthermore, both ambiguity aversion[140] and neural responses to ambiguity[141] are greatest when missing information about outcome probabilities is potentially knowable (that is, known to the experimenter, and usable in order to infer the conditional first-order probabilities), and not when rule uncertainty is highest. Consequently, BOLD (blood oxygen-level dependent) functional MRI responses to ambiguity and to continuously varied rule uncertainty differ markedly[63]. Hence, a categorical contrast of ambiguity versus no ambiguity is probably not a good model for rule uncertainty.

realized. Subjects over-weighted the probability of the worse possible rule when rule uncertainty was high and under-weighted it when rule uncertainty was low, and this biased their choices[63]. Thus, rule uncertainty can influence behaviour, even though under these particular (laboratory) circumstances this did not lead to minimization of electric shocks (that is, it did not elicit optimal behaviour). More research is needed in this area to specify under which circumstances rule uncertainty affects behaviour, and whether or not this leads to optimal outcomes.

*Encoding of rule uncertainty.* In the aforementioned neuroimaging study[63], BOLD responses to rule uncertainty were expressed in posterior midline areas extending into posterior cingulate, parietal and occipital cortices. It turns out that high rule uncertainty in this task might also imply lower subjective value for an expected outcome as well as higher processing demands, both of which could possibly explain the activation seen in the posterior cingulate cortex. In a reinforcement learning context, a correlate of the belief about rule change per time unit (termed volatility in this study) has been reported in the ACC[64]. Note that volatility is not equal to rule uncertainty, such as in the case of an unknown but stationary rule, in which rule uncertainty is high in the face of zero volatility.

A study using a visual discrimination paradigm showed that the speed of learning of rule changes can be improved with the noradrenaline agonist idazoxan[65]. This has been taken to suggest that noradrenaline has a role in the detection of unexpected uncertainty[61] and, indirectly, in the estimation of rule uncertainty. However, it could also suggest that noradrenaline has a role in learning new rules after a rule change has been detected, or in the abolition of old learning rules, and, so far, there are no empirical data that can distinguish between these possibilities.

Overall, few studies have investigated the effect of rule uncertainty on behaviour and its neural representation, although a number of interesting approaches to characterize and quantify rule uncertainty exist in reinforcement learning and economic decision-making. To complicate things, rule uncertainty naturally arises in learning situations. If one assumes that there are different learning systems for different problems[66], then it is plausible that rule uncertainty might be represented in different brain areas, depending on the learning system that is optimal to the task at hand[62].

## Uncertainty about outcomes

Imagine having arrived at our friend's house. We then recall that he might be going out for dinner tonight. Even if we know the precise odds that he is at home, we will be uncertain about whether he really is at home or not. For our purposes, uncertainty in relation to what is going to happen in the future is termed outcome uncertainty and this can be quantified. If we think about the world as being composed of discrete states, then outcome uncertainty describes the uncertainty about which state will be realized next. This uncertainty can be quantified if we (almost) precisely know the rule that governs the transition to an upcoming state. Experiments on outcome uncertainty described so far mainly deal with states that

are imbued with some economic, reinforcing or punishing value. In a more general perspective, any kind of uncertainty about the future can be thought of as outcome uncertainty, and four distinct literatures dominate in this field.

First, economic theory refers to outcome uncertainty as 'risk' in situations in which all possible outcomes and their associated probabilities are explicitly known. Accounts of outcome uncertainty in classical economic theories (most prominently, expected utility theory[9], subjective utility theory[67] and prospect theory[68]) posit that the internal utility of an outcome is a nonlinear function of the value of that outcome (BOX 5). Such nonlinearity can explain aversion to, or preference for, variable outcomes on the basis of utility maximization, without invoking a specific role for outcome uncertainty[8]. An alternative perspective, which is based on risk–return models in finance theory[69], posits that choice options can be represented in summary-statistic measures (that is, statistical moments) of their possible return with outcome uncertainty quantified as variance. A neural system implementing such an algorithm would require a specific representation of outcome uncertainty[70,71]. Below, we review studies that have attempted to determine which of the two theories can better capture human behaviour, as well as those that seek to identify neural encoding of outcome uncertainty in an economic context.

Second, recent developments in reinforcement-learning theory posit that a representation of outcome uncertainty (also termed expected uncertainty (BOX 3)) is needed to guide learning. Although there is little empirical data on the behavioural consequences of outcome uncertainty in this context, reinforcement-learning paradigms are sometimes used to investigate neural representation of outcome uncertainty[72–75].

Third, any framework that models discrete states of the world assumes that future states are uncertain to a variable degree. Some empirical work has drawn on such models without explicit value[75].

Finally, in motor control it is recognized that we can implement motor commands only with a degree of imprecision[4,13]. Such motor uncertainty constitutes outcome uncertainty and outcome uncertainty estimates in this context are key concepts in the field of sensorimotor control[13].
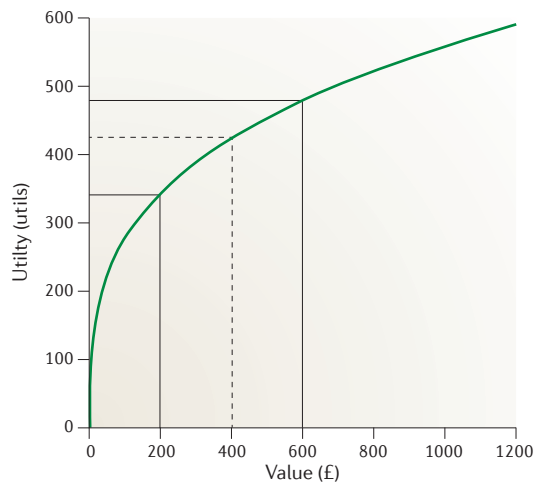
In the theoretical and empirical body of literature, different quantifications of outcome uncertainty have been proposed (BOX 1). We are not aware of studies seeking to determine which of these quantifications is best supported by neural or behavioural evidence.

***Outcome uncertainty guides behaviour.*** A literature search on behavioural economics studies shows that, in humans and other animals, variability in gamble outcomes influences choice behaviour[76]. This phenomenon can be explained by utility maximization[9,67,68]. Under the assumption that the algorithms posited by these theories are indeed implemented, then no neural representation of uncertainty is needed even if it appears to guide behaviour (BOX 5). However, these algorithms imply, as a corollary, that every individual possible outcome is neurally encoded, and this would be inefficient when there are many outcomes.

In a repeated lottery, participants were explicitly told the magnitudes of all possible outcomes[77]. By paying some money, they could sample from the lottery to experience the probabilities of these outcomes. Then, they were given a chance to bet on that lottery in order to gain money. In some trials, both outcome probabilities and magnitudes changed. In other trials, participants were informed that outcome magnitudes had changed from the previous trial, but that outcome probabilities were held constant. Had participants encoded all the outcome probabilities from the previous trial, there would have been nothing to gain from sampling the new lottery (because the magnitudes, the only thing that changed, were stated explicitly). However, in half of these trials, participants did sample the probabilities, which suggests that they did not encode all individual outcome probabilities, thus contradicting the idea of utility maximization accounts of risk preference[77]. In fact, in these trials, the participants' behaviour suggests that they encoded a summary-statistic measure of outcome uncertainty without having to encode individual outcome probabilities and magnitudes.

---

### Box 5 | Outcome uncertainty and expected utility theory

In the economic literature, utility is often conceptualized as a measure of satisfaction by a good. Utility is assumed not to be linearly related to the value of that good. For example, the utility of three cars might not be three times the utility of one car (that is, receiving three cars might not be three times as satisfying as receiving one car). The figure shows an example for a nonlinear utility function as prescribed by several mainstream economic theories: with increasing value, the increment in utility in this particular example becomes smaller. To make utility measurable and to make the use of utility functions tractable, expected utility theory[9] relies on a number of assumptions that are not reviewed here.



We briefly illustrate how this kind of utility function might explain a behavioural sensitivity to uncertainty (see the figure). In a lottery that can yield £200 or £600, as a result of a flip of a coin, the expected value is £400; hence a value-maximizer should be indifferent between this lottery and receiving £400 for sure. However, the expected utilities map nonlinearly, so that in the above example, the expected utility of the lottery is lower than the expected utility of winning £400. Hence, a utility-maximizer should avoid the lottery, regardless of whether there is an explicit uncertainty representation in the brain.

Now, imagine a choice between two lotteries, both of which have three outcomes with equal probabilities. The first lottery could yield £400, £500 or £600. The second lottery could yield £100, £200 or £1,200. Both lotteries have the same expected value of £500. However, most people would find £1,200 less than two times as useful as winning £600. Hence, the expected utility of the second lottery is lower than that of the first, and most people choose the first. This lottery also turns out to have lower outcome uncertainty. Hence, this framework can explain sensitivity to outcome uncertainty without having to explicitly describe uncertainty.

Outcome uncertainty also occurs in Posner-type tasks, which involve state transitions without economic value. Here, a correct action depends on a target position that was cued; however, the cues are not always correct. The validity of the cue and therefore the uncertainty about the target (the upcoming state), has been shown to be inversely correlated to a reaction time difference between correctly and incorrectly cued actions[61]. In other words, when outcome uncertainty was low, subjects needed more time to shift their attention after a change in target position. Also, corticospinal excitability, which indicates motor preparation, is high when there is little uncertainty about the state transition[78].

Motor uncertainty is the estimated uncertainty of an upcoming motor action and has been investigated from the perspective of sensorimotor control. In motor planning, sensory and motor uncertainty are often integrated, and humans do this in a near-optimal manner[41,42] (see section on sensory uncertainty). Furthermore, the cost of an imprecise motor action can be spatially asymmetric (for example, when walking close to a cliff edge, the cost of staggering to the cliff edge is higher than staggering to the other side). In such cases, motor uncertainty influences average motion in humans[79] (in our example, one would walk further away from the cliff edge). Furthermore, when controlling a ball, humans apply different control movements at the same state of the ball depending on uncertainty of the future trajectory of the ball[80].

In summary, there is strong evidence to suggest that outcome uncertainty influences behaviour, both when future states of the world are imbued with economic value, and when discrete states have no explicit value or when future states are defined in continuous time as in sensorimotor control. Some economic theories argue against an explicit notion of uncertainty to explain such phenomena[9], but there is evidence to suggest that at the very least some assumptions of these theories are not met[77]. This indicates that we should expect to find an explicit neural encoding of outcome uncertainty to explain its impact on behaviour.

*Encoding of outcome uncertainty.* The neural representation of uncertainty has been widely investigated in neuroeconomic studies, in which outcome uncertainty is usually termed risk — a terminology that is sometimes used in reinforcement-learning paradigms. In reinforcement learning, single-unit recordings in monkeys suggest that prediction error signals are scaled by the variability in possible outcomes[81]. That is, phasic dopamine neuron responses are of the same magnitude for large deviations from an expected outcome in high-uncertainty contexts and for small deviations in low-uncertainty contexts[81], a phenomenon that requires a neural representation of uncertainty.

There is also direct evidence for a specific representation of outcome uncertainty in monkeys who passively view reward cues. Here, it has been suggested that the rate of change in a tonic ramping of midbrain dopamine neurons before reward delivery encodes reward variance[82] (but see the debate in REFS 83,84). In addition, a summary-statistic encoding of outcome uncertainty has been

found in OFC neurons in a study in which outcome uncertainty was defined as variance in reward magnitudes when two possible rewards had equal probability and different magnitude[85]. Crucially, the mean reward magnitude was encoded by a (largely) different set of neurons in the OFC in this experiment.

This is an important result, because according to some economic theories, situations with highly variable outcomes have lower internal utility than those with less variable outcomes. Hence, neural activity that correlates with outcome uncertainty might reflect utility encoding if these theories apply to how the brain works. However, in this experiment, there is evidence showing that distinct neurons respond to uncertainty versus utility, which therefore discounts this possibility.

Several human neuroimaging experiments have addressed encoding of outcome uncertainty. In some of these experiments, outcome certainty was defined as the probability of receiving reward, and findings of these studies are therefore more informative about expected reward anticipation than uncertainty[86–88]. In other studies, encoding of outcome uncertainty versus expected reward were distinguished. These studies involved paradigms in which participants performed an action to receive an outcome or just passively awaited a signalled outcome, whereby outcome possibilities could either be learnt or were explicitly signalled as in economic gambles. Such studies investigated both immediate and sustained (over a few seconds) BOLD responses to a cue that signalled outcome uncertainty. Sustained responses possibly reflect (slowly) ramping activity in dopaminergic neurons, which correlate with uncertainty[82]. Immediate responses to outcome uncertainty during passive anticipation of explicitly signalled outcomes were seen within parahippocampal gyrus, transverse temporal gyrus, anterior insula, midbrain areas, ACC and localized parietal and frontal regions[89], and during passive anticipation of previously learned outcomes in the lateral orbitofrontal gyrus[72]. After choosing between certain and a variable option with previously learned outcomes, immediate responses to outcome uncertainty of the variable option were observed in anterior insula[73]. Upon choosing between an explicitly signalled option and one with previously learned outcomes, outcome uncertainty of the explicit option was represented in anterior insula, in contrast to that seen for a previously learned option in ACC[74].

Sustained BOLD responses during passive anticipation of explicitly signalled outcomes were found in anterior insula, ventral striatum, thalamus, midbrain and small parietal and frontal regions in one study[89], in medial prefrontal/orbitofrontal regions and hippocampus in another study[90], in putamen in a third study[91], in putamen, anterior insula and ventromedial orbitofrontal cortex in a fourth study[92], and in parietal cortex in a fifth study[93]. Sustained responses to outcome uncertainty during choice involving explicitly signalled outcomes have been observed in anterior insula and orbitofrontal cortex[94], and in a task involving choices on previously learned outcomes in the midbrain[75]. Furthermore, activation of dorsal ACC has been reported during choice for explicitly signalled outcomes, although it is not clear whether this referred to

immediate or sustained responses[95]. In a paradigm unrelated to decision making, during a long stream of visual events with an associated incidental task, the entropy of outcome predictions (that is, the average uncertainty) for each stimulus covaried with immediate BOLD responses in the hippocampus[96].

Note that all the learning tasks mentioned above involved, but did not explicitly investigate responses to, rule uncertainty, which is sometimes even confounded with outcome uncertainty[75]. In addition, for all sustained responses, outcome uncertainty is moderately confounded with the resolution of uncertainty and the associated prediction errors[97]. Furthermore, the reported studies have different standards in terms of the thresholds of statistical significance.

Outcome uncertainty is a crucial feature of Posner tasks, as discussed above. When cue validity (and therefore state uncertainty) is kept constant, the difference in reaction times between non-cued and cued targets increases after a lesion to basal forebrain acetylcholine areas[98] and decreases after nicotine application[99,100]. This has been interpreted as indicating an involvement of acetylcholine in the representation of outcome uncertainty[61]. However, as uncertainty was not varied in these experiments, acetylcholine might simply be involved in shifting attention after an invalid cue. Thus, the notion that acetylcholine signals uncertainty about outcomes requires further support.

In summary, a few firm conclusions can be drawn regarding outcome uncertainty. First, phasic prediction error signals in dopaminergic neurons are scaled by outcome uncertainty. A possible, but debated, mechanism for encoding outcome uncertainty is sustained firing of these same neurons. Second, a neuronal population in OFC carries outcome uncertainty signals in their firing rate, independent of value signals. Third, BOLD fMRI studies have implicated more than ten distinct brain regions as representing outcome uncertainty in different situations. These findings await replication in comparable experimental designs and with conservative thresholds of statistical significance. A parsimonious conclusion is that outcome uncertainty is not a unitary concept in the sense that its representation depends on context. Taken together, it remains unclear how findings in dopaminergic midbrain regions and OFC relate to each other and to overt behaviour, and to what extent outcome uncertainty representations depend on the actual outcomes that are uncertain.

## Conclusions and future directions

In this Review, we considered uncertainty within a hierarchical processing model (FIG. 2), which allowed us to compare empirical findings from widely differing studies and experimental contexts. We first considered whether uncertainty on different variables guides behaviour, and conclude that there is strong evidence from different species to suggest that sensory and outcome uncertainty do influence behaviour, with sensory uncertainty (most of the time) being utilized in a near-optimal manner.

Economic theory explains the behavioural influence of outcome uncertainty by invoking algorithms not involving uncertainty, but empirical evidence now points to a an algorithm that explicitly uses outcome uncertainty. There is little theoretical or empirical work to suggest that state uncertainty influences behaviour (over and above a direct influence of sensory uncertainty on state estimation), and there is only weak evidence to suggest a behavioural impact of rule uncertainty, so far only in situations in which this impact is not mathematically optimal. In addition to a need for additional empirical work, recent theoretical advances on rule uncertainty in reinforcement learning needs to be integrated with economic theory and with studies on uncertainty of state transition rules in non-economic paradigms. It seems parsimonious to conclude that in most cases, uncertainty guides behaviour, often in an optimal fashion.

A second question we addressed was evidence for a neural representation of uncertainty. We conclude that neuronal correlates of uncertainty have been demonstrated across all the posited levels of uncertainty (sensory, state, rule and outcome). There is strong evidence for specific representations of sensory and outcome uncertainty that influence other neural computations. However, only a handful of well-controlled animal studies have addressed how precisely this uncertainty is encoded, and it is still
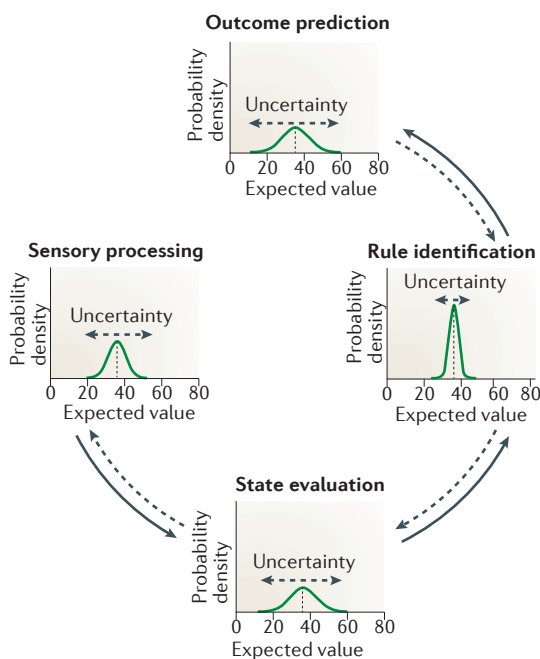


Figure 2 | **The brain as a hierarchical probabilistic machine.** Schematic example of hierarchical computations in an action episode. Each graph reflects a neural variable that is represented by a probability distribution of possible values. Possible values can be discrete states or continuous values; in this example we depict magnitude values. The expected value is the most likely value, and the uncertainty of its estimation can be inferred from the dispersion of the probability distribution (see BOX 1 for quantification of the dispersion). Both the expected value and its associated uncertainty influence computations performed at other processing levels through forward (solid arrows) and backward (dashed arrow) projections.

not understood how it is translated into behaviour. A proliferation of BOLD fMRI studies in humans examining uncertainty across all levels of the proposed hierarchy are beset with interpretative problems. Primarily, this is because of methodological reasons, which are due both to weaknesses of individual experiments and inherent difficulties in linking BOLD responses to precise neural mechanisms.

Third, we considered the functional anatomy of uncertainty representation: is uncertainty encoded in a distributed manner or encoded in a single area that serves as an uncertainty module, which calculates all levels of imprecision? We have highlighted the fact that studies report different brain areas that carry uncertainty signals, even between studies using similar experimental manipulations. This might reflect experimental variations and confounding factors that we discussed in this article. However, we think it is unlikely that such heterogeneous results could arise if one brain area is responsible for calculating uncertainty. In particular, when distinct forms of uncertainty are directly contrasted, a distributed neural representation emerges. As a conclusion, we propose a spatially distributed encoding of uncertainty that relates to the actual variable we are uncertain about.

How is a putative distributed coding scheme organized? One might assume that coding of uncertainty about a variable is associated with encoding value. Although this has seldom been investigated in a single experiment, the literature considered in this article suggests that this assumption is not supported. One methodological reason for this lack of support is that within hierarchical models of brain function[2], uncertainty at one level of neuronal computation necessarily influences the value of variables at higher (bottom-up) or lower (top-down) levels (FIG. 2). This means that when one manipulates uncertainty and measures its correlates on one variable, one implicitly manipulates (and measures) the value of these other variables. As a simple example, increasing sensory uncertainty in a rewarded sensory decision-making task decreases expected reward; that is, the outcome value. The present state of research does not allow firm conclusions to be made about the spatial distribution of uncertainty representations in the brain, and a clear specification is required of the likely consequences from a specific manipulation of uncertainty. The kind of hierarchical model used in this Review is a heuristic example of specifying processing levels. Specifying process levels are necessary to be able to make a statement of the precise variables related to uncertainty that are influenced in any particular experiment.

Assuming that there are specific neural representations of sensory, state, rule and outcome uncertainty, what type of neural code is involved? This is an issue that has been barely addressed experimentally. Several models based on theoretical analyses have been proposed[11]. Among these are population codes that require neuronal tissue and fixed-form codes that require only a limited number of neurons[101]. Beyond sensory decision-making (in which the rate of change in individual neurons' firing rate are assumed to carry uncertainty),

a summary-statistic population code for uncertainty is implicitly assumed in most human fMRI studies, as these use mass-univariate analysis (an approach that cannot detect representations that rest on very few neurons).

The finding that uncertainty manipulations result in widespread brain activation argues for population codes, and, specifically, for a summary-statistic representation. However, other potential codes have not been rigorously investigated. In addition, it is difficult to unambiguously detect specific uncertainty correlates in individual neurons or neuron populations in animal experiments. In human neuroimaging research, it seems appropriate to use multivariate fMRI methods and compare their sensitivity with the sensitivity of mass-univariate methods in determining the nature of the coding. This can open the door to an unexploited field that might bridge the gap between single-neuron recordings and neuronal tissue-based methods such as human neuroimaging. Note that this Review focused on instantaneous representations of uncertainty. It is conceivable that uncertainty in, for example, a sensory modality is reflected in enduring, hard-wired characteristics of a neural network[11], which was possibly acquired during evolution or ontogenesis, with a representation at the level of brain structure rather than function.

By arranging the literature on uncertainty into distinct levels of a hierarchical process, we highlight that an abundance of experiments address sensory uncertainty or uncertainty in (particularly economic) outcomes, whereas state uncertainty and rule uncertainty have not been widely studied. In addition, uncertainty at the same hierarchical level can occur in different circumstances, and there is a heavy bias here as well. Economic outcome uncertainty is investigated much more than motor outcome uncertainty, and economic rule uncertainty has received a lot more attention than reinforcement rule uncertainty. Additionally, it appears that there is sometimes little connection between distinct fields of research, which is apparent in the idiosyncratic terminology that is used in different fields for the same concepts.

In this article, we have not taken an a priori theoretical stance on whether or how uncertainty estimates are utilized in the brain, but have focused on empirical literature. Nevertheless, the empirical data are in alignment with a class of brain theories that assume a hierarchical brain architecture that is suited to solve probabilistic computations so as to infer the causal structure of the world[1–3,6,101]; for example, the Bayesian brain and the Helmholtz machine. The representation and use of uncertainty estimates are important concepts (among other concepts) in such theories. This is evident in sensory and motor uncertainty research, which is often guided by such theories. We think that it will be fruitful to integrate other approaches to uncertainty — in particular from the field of neuroeconomics — into such theories.

Many computational models assume that human behaviour is optimal and prescribe algorithms that could lead to such optimality. Humans indeed often perform at a near-optimal level and use uncertainty estimates in doing so. However, there are many examples of

# REVIEWS

behaviour that is not optimal. Such suboptimal behaviour might arise from physical characteristics of neural computations (that is, neural algorithms might be approximations that require less computational power than an optimal algorithm)[10] or from prior expectation that would render the neural computation optimal in a larger context outside the laboratory. This leads to the question of how uncertainty processing relates to psychiatric disorders: neural development and learning history might influence prior expectations on certain variables and in this way account for seemingly irrational and maladaptive behaviours. Indeed, it can be argued that a number of psychiatric illnesses may be linked to aberrant processing of environmental uncertainty (that is, in generalized anxiety disorder and depression)[102] or to increased uncertainty in neural computations (for example, as in schizophrenia)[103]. Hence, research into

uncertainty coding has implications not only for a basic understanding of brain functions, but also for informing clinical contexts. Furthermore, humans often develop an explicit, metacognitive notion of uncertainty[104] (which we have not discussed in this Review), and this is likely to be of particular interest in the context of psychiatric conditions.

Although uncertainty is discomforting, it is equally true that monotony is an even greater affliction[105,106]. Humans and other animals have evolved to cope with external and internal noise, and hence uncertainty in the environment may benefit neural computations[107–112]. Perhaps this is why *Homo sapiens* continuously strives to cross frontiers, reach higher ground, is addicted to exploration and invests so much store in scientific research — activities that abolish old uncertainties and endlessly create new ones.

1. Dayan, P., Hinton, G. E., Neal, R. M. & Zemel, R. S. The Helmholtz machine. *Neural Comput.* **7**, 889–904 (1995).
2. Friston, K. The free-energy principle: a unified brain theory? *Nature Rev. Neurosci.* **11**, 127–138 (2010).
3. Knill, D. C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
4. Kording, K. P. & Wolpert, D. M. Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* **10**, 319–326 (2006).
5. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* **5**, 39 (2011).
6. Hinton, G. E. & Dayan, P. Varieties of Helmholtz machine. *Neural Netw.* **9**, 1385–1403 (1996).
7. Ellsberg, D. Risk, ambiguity, and the Savage axioms. *Q. J. Econ.* **75**, 643–669 (1961).
8. Bernoulli, D. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae* **5**, 175–192 (1738) (in Latin).
9. von Neumann, J. & Morgenstern, O. *Theory of Games and Economic Behavior* (Princeton Univ. Press, 1944).
10. Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E. & Pouget, A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* **74**, 30–39 (2012).
11. Vilares, I. & Kording, K. Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Ann. NY Acad. Sci.* **1224**, 22–39 (2011).
12. Singer, T., Critchley, H. D. & Preuschoff, K. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* **13**, 334–340 (2009).
13. Orban, G. & Wolpert, D. M. Representations of uncertainty in sensorimotor control. *Curr. Opin. Neurobiol.* **21**, 629–635 (2011).
14. Trommershauser, J., Kording, K. & Landy, M. S. *Sensory Cue Integration* (Oxford Univ. Press, 2011).
15. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
    **One of the first experimental papers to show optimal multisensory (haptic and visual) integration when visual uncertainty was dynamically varied.**
16. Knill, D. C. & Saunders, J. A. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res.* **43**, 2539–2558 (2003).
17. Hillis, J. M., Watt, S. J., Landy, M. S. & Banks, M. S. Slant from texture and disparity cues: optimal cue combination. *J. Vis.* **4**, 967–992 (2004).
18. van Beers, R. J., Sittig, A. C. & Gon, J. J. Integration of proprioceptive and visual position-information: An experimentally supported model. *J. Neurophysiol.* **81**, 1355–1364 (1999).
19. Battaglia, P. W., Jacobs, R. A. & Aslin, R. N. Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **20**, 1391–1397 (2003).
20. Wozny, D. R., Beierholm, U. R. & Shams, L. Human trimodal perception follows optimal statistical inference. *J. Vis.* **8**, 24–11 (2008).

21. Serwe, S., Kording, K. P. & Trommershauser, J. Visual-haptic cue integration with spatial and temporal disparity during pointing movements. *Exp. Brain Res.* **210**, 67–80 (2011).
22. Alais, D. & Burr, D. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* **14**, 257–262 (2004).
23. Jacobs, R. A. Optimal integration of texture and motion cues to depth. *Vision Res.* **39**, 3621–3629 (1999).
24. Young, M. J., Landy, M. S. & Maloney, L. T. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res.* **33**, 2685–2696 (1993).
25. Landy, M. S. & Kojima, H. Ideal cue combination for localizing texture-defined edges. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **18**, 2307–2320 (2001).
26. van Beers, R. J., Wolpert, D. M. & Haggard, P. When feeling is more important than seeing in sensorimotor adaptation. *Curr. Biol.* **12**, 834–837 (2002).
27. Fetsch, C. R., Turner, A. H., Deangelis, G. C. & Angelaki, D. E. Dynamic reweighting of visual and vestibular cues during self-motion perception. *J. Neurosci.* **29**, 15601–15612 (2009).
28. Helbig, H. B. & Ernst, M. O. Optimal integration of shape information from vision and touch. *Exp. Brain Res.* **179**, 595–606 (2007).
29. Helbig, H. B. & Ernst, M. O. Visual-haptic cue weighting is independent of modality-specific attention. *J. Vis.* **8**, 21–16 (2008).
30. Brouwer, A. M. & Knill, D. C. Humans use visual and remembered information about object location to plan pointing movements. *J. Vis.* **9**, 24–19 (2009).
31. Kording, K. P. & Wolpert, D. M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247 (2004).
32. Tassinari, H., Hudson, T. E. & Landy, M. S. Combining priors and noisy visual cues in a rapid pointing task. *J. Neurosci.* **26**, 10154–10163 (2006).
33. Kording, K. P., Ku, S. P. & Wolpert, D. M. Bayesian integration in force estimation. *J. Neurophysiol.* **92**, 3161–3165 (2004).
34. Miyazaki, M., Nozaki, D. & Nakajima, Y. Testing Bayesian models of human coincidence timing. *J. Neurophysiol.* **94**, 395–399 (2005).
35. Rosas, P., Wagemans, J., Ernst, M. O. & Wichmann, F. A. Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **22**, 801–809 (2005).
36. Fetsch, C. R., Pouget, A., Deangelis, G. C. & Angelaki, D. E. Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neurosci.* **15**, 146–154 (2011).
37. Serwe, S., Drewing, K. & Trommershauser, J. Combination of noisy directional visual and proprioceptive information. *J. Vis.* **9**, 28–14 (2009).
38. Schlicht, E. J. & Schrater, P. R. Effects of visual uncertainty on grasping movements. *Exp. Brain Res.* **182**, 47–57 (2007).
39. Schlicht, E. J. & Schrater, P. R. Impact of coordinate transformation uncertainty on human sensorimotor control. *J. Neurophysiol.* **97**, 4203–4214 (2007).

40. Izawa, J. & Shadmehr, R. On-line processing of uncertain information in visuomotor control. *J. Neurosci.* **28**, 11360–11368 (2008).
41. Faisal, A. A. & Wolpert, D. M. Near optimal combination of sensory and motor uncertainty in time during a naturalistic perception-action task. *J. Neurophysiol.* **101**, 1901–1912 (2009).
42. Battaglia, P. W. & Schrater, P. R. Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *J. Neurosci.* **27**, 6984–6994 (2007).
43. Wei, K. & Kording, K. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front. Comput. Neurosci.* **4**, 11 (2010).
44. Kiani, R. & Shadlen, M. N. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* **324**, 759–764 (2009).
    **Demonstration of LIP neurons that fire in accordance with a temporal integrator model in such a way that their firing reflects sensory uncertainty and predicts behaviour.**
45. Barthelme, S. & Mamassian, P. Evaluation of objective uncertainty in the visual system. *PLoS. Comput. Biol.* **5**, e1000504 (2009).
46. Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008).
47. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
48. Heuer, H. W. & Britten, K. H. Linear responses to stochastic motion signals in area MST. *J. Neurophysiol.* **98**, 1115–1124 (2007).
49. Philiastides, M. G., Ratcliff, R. & Sajda, P. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J. Neurosci.* **26**, 8965–8975 (2006).
50. Banko, E. M., Gal, V., Kortvelyes, J., Kovacs, G. & Vidnyanszky, Z. Dissociating the effect of noise on sensory processing and overall decision difficulty. *J. Neurosci.* **31**, 2663–2674 (2011).
51. Kayser, A. S., Buchsbaum, B. R., Erickson, D. T. & D'Esposito, M. The functional anatomy of a perceptual decision in the human brain. *J. Neurophysiol.* **103**, 1179–1194 (2010).
52. Philiastides, M. G. & Sajda, P. EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J. Neurosci.* **27**, 13082–13091 (2007).
53. Daniel, R. *et al.* Assessing the neural basis of uncertainty in perceptual category learning through varying levels of distortion. *J. Cogn Neurosci.* **23**, 1781–1793 (2011).
54. Grinband, J., Hirsch, J. & Ferrera, V. P. A neural representation of categorization uncertainty in the human brain. *Neuron* **49**, 757–763 (2006).
55. Helbig, H. B. *et al.* The neural mechanisms of reliability weighted integration of shape information from vision and touch. *Neuroimage* **60**, 1063–1072 (2012).

56. Mackintosh, N. J. *Conditioning and Associative Learning* (Oxford Univ. Press, 1983).
57. Blanchard, D. C. & Blanchard, R. J. Ethoexperimental approaches to the biology of emotion. *Annu. Rev. Psychol.* **39**, 43–68 (1988).
58. Graeff, F. G. Neuroanatomy and neurotransmitter regulation of defensive behaviors and related emotions in mammals. *Braz. J. Med. Biol. Res.* **27**, 811–829 (1994).
59. Gray, J. A. & McNaughton, N. *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System* (Oxford Univ. Press, 2000).
60. Yoshida, W. & Ishii, S. Resolution of uncertainty in prefrontal cortex. *Neuron* **50**, 781–789 (2006).
   **One of the few neuroimaging investigations of state uncertainty, realized in a maze task with unknown position.**
61. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
62. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neurosci.* **8**, 1704–1711 (2005).
63. Bach, D. R., Hulme, O., Penny, W. D. & Dolan, R. J. The known unknowns: neural representation of second-order uncertainty, and ambiguity. *J. Neurosci.* **31**, 4811–4820 (2011).
   **One of the first studies to continuously manipulate rule uncertainty. It demonstrated both an effect on behaviour and a neural encoding that were different from that of the categorical concept ambiguity.**
64. Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nature Neurosci.* **10**, 1214–1221 (2007).
65. Devauges, V. & Sara, S. J. Activation of the noradrenergic system facilitates an attentional shift in the rat. *Behav. Brain Res.* **39**, 19–28 (1990).
66. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nature Rev. Neurosci.* **9**, 545–556 (2008).
67. Savage, L. J. *The Foundations of Statistics* (Wiley, 1954).
68. Kahneman, D. & Tversky, A. Prospect theory — analysis of decision under risk. *Econometrica* **47**, 263–291 (1979).
69. Markowitz, H. Portfolio selection. *J. Finance* **7**, 77–91 (1952).
70. Rushworth, M. F. & Behrens, T. E. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neurosci.* **11**, 389–397 (2008).
71. Platt, M. L. & Huettel, S. A. Risky business: the neuroeconomics of decision making under uncertainty. *Nature Neurosci.* **11**, 398–403 (2008).
72. Tobler, P. N., O'Doherty, J. P., Dolan, R. J. & Schultz, W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* **97**, 1621–1632 (2007).
73. Rolls, E. T., McCabe, C. & Redoute, J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb. Cortex* **18**, 652–663 (2008).
74. Fitzgerald, T. H., Seymour, B., Bach, D. R. & Dolan, R. J. Differentiable neural substrates for learned and described value and risk. *Curr. Biol.* **20**, 1823–1829 (2010).
75. Aron, A. R. *et al.* Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J. Neurophysiol.* **92**, 1144–1152 (2004).
76. Kagel, J. H. & Roth, A. E. *Handbook of Experimental Economics* (Princeton Univ. Press, 1995).
77. d'Acremont, M. & Bossaerts, P. Neurobiological studies of risk assessment: a comparison of expected utility and mean-variance approaches. *Cogn. Affect. Behav. Neurosci.* **8**, 363–374 (2008).
   **Elegant economic experiment to refute an assumption that some algorithms prescribed by expected utility theory are implemented in a nervous system.**
78. Bestmann, S. *et al.* Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Curr. Biol.* **18**, 775–780 (2008).
79. Trommershauser, J., Maloney, L. T. & Landy, M. S. Statistical decision theory and trade-offs in the control of motor response. *Spat. Vis.* **16**, 255–275 (2003).

80. Nagengast, A. J., Braun, D. A. & Wolpert, D. M. Risk-sensitive optimal feedback control accounts for sensorimotor behavior under uncertainty. *PLoS Comput. Biol.* **6**, e1000857 (2010).
81. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
   **Shows adaption of dopaminergic midbrain neurons. Among other issues, it is demonstrated that prediction error signals are scaled by outcome uncertainty.**
82. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
   **Report of slowly ramping activity in dopaminergic midbrain neurons correlating with outcome uncertainty.**
83. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behav. Brain Funct.* **1**, 7 (2005).
84. Niv, Y., Duff, M. O. & Dayan, P. Dopamine, uncertainty and TD learning. *Behav. Brain Funct.* **1**, 6 (2005).
85. O'Neill, M. & Schultz, W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* **68**, 789–800 (2010).
   **Unambiguous demonstration that OFC neurons separately code outcome uncertainty and expected value.**
86. Stern, E. R., Gonzalez, R., Welsh, R. C. & Taylor, S. F. Updating beliefs for a decision: neural correlates of uncertainty and underconfidence. *J. Neurosci.* **30**, 8032–8041 (2010).
87. Critchley, H. D., Mathias, C. J. & Dolan, R. J. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* **29**, 537–545 (2001).
88. Huettel, S. A., Song, A. W. & McCarthy, G. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* **25**, 3304–3311 (2005).
89. Preuschoff, K., Bossaerts, P. & Quartz, S. R. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**, 381–390 (2006).
90. Abler, B., Herrnberger, B., Gron, G. & Spitzer, M. From uncertainty to reward: BOLD characteristics differentiate signaling pathways. *BMC. Neurosci.* **10**, 154 (2009).
91. Dreher, J. C., Kohn, P. & Berman, K. F. Neural coding of distinct statistical properties of reward information in humans. *Cereb. Cortex* **16**, 561–573 (2006).
92. Symmonds, M., Bossaerts, P. & Dolan, R. J. A behavioral and neural evaluation of prospective decision-making under risk. *J. Neurosci.* **30**, 14380–14389 (2010).
93. Symmonds, M., Wright, N. D., Bach, D. R. & Dolan, R. J. Deconstructing risk: separable encoding of variance and skewness in the brain. *Neuroimage* **58**, 1139–1149 (2011).
94. Mohr, P. N., Biele, G., Krugel, L. K., Li, S. C. & Heekeren, H. R. Neural foundations of risk-return trade-off in investment decisions. *Neuroimage* **49**, 2556–2563 (2010).
95. Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J. & Schultz, W. Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *J. Neurosci.* **29**, 12574–12583 (2009).
96. Strange, B. A., Duggins, A., Penny, W., Dolan, R. J. & Friston, K. J. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw.* **18**, 225–230 (2005).
97. Preuschoff, K., Quartz, S. R. & Bossaerts, P. Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* **28**, 2745–2752 (2008).
98. Voytko, M. L. *et al.* Basal forebrain lesions in monkeys disrupt attention but not learning and memory. *J. Neurosci.* **14**, 167–186 (1994).
99. Witte, E. A., Davidson, M. C. & Marrocco, R. T. Effects of altering brain cholinergic activity on covert orienting of attention: comparison of monkey and human performance. *Psychopharmacol. (Berl.)* **132**, 324–334 (1997).
100. Phillips, J. M., McAlonan, K., Robb, W. G. & Brown, V. J. Cholinergic neurotransmission influences covert orientation of visuospatial attention in the rat. *Psychopharmacol. (Berl.)* **150**, 112–116 (2000).
101. Friston, K. The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* **13**, 293–301 (2009).

102. Hazlett-Stevens, H. & Borkovec, T. D. Interpretive cues and ambiguity in generalized anxiety disorder. *Behav. Res. Ther.* **42**, 881–892 (2004).
103. Loh, M., Rolls, E. T. & Deco, G. A dynamical systems hypothesis of schizophrenia. *PLoS. Comput. Biol.* **3**, 228 (2007).
104. Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J. & Rees, G. Relating introspective accuracy to individual differences in brain structure. *Science* **329**, 1541–1543 (2010).
105. Hill, A. B. & Perkins, R. E. Towards a model of boredom. *Br. J. Psychol.* **76**, 235–240 (1985).
106. Perkins, R. E. & Hill, A. B. Cognitive and affective aspects of boredom. *Br. J. Psychol.* **76**, 221–234 (1985).
107. Miller, P. & Katz, D. B. Stochastic transitions between neural states in taste processing and decision-making. *J. Neurosci.* **30**, 2559–2570 (2010).
108. Kitajo, K., Nozaki, D., Ward, L. M. & Yamamoto, Y. Behavioral stochastic resonance within the human brain. *Phys. Rev. Lett.* **90**, 218103 (2003).
109. Sasaki, H. *et al.* Suprathreshold stochastic resonance in visual signal detection. *Behav. Brain Res.* **193**, 152–155 (2008).
110. Sasaki, H. *et al.* Subthreshold noise facilitates the detection and discrimination of visual signals. *Neurosci. Lett.* **436**, 255–258 (2008).
111. Simonotto, E., Riani, M., Roberts, M., Twitty, J. & Moss, F. Visual perception of stochastic resonance. *Phys. Rev. Lett.* **78**, 1186–1189 (1997).
112. Lugo, E., Doti, R. & Faubert, J. Ubiquitous crossmodal stochastic resonance in humans: auditory noise facilitates tactile, visual and proprioceptive sensations. *PLoS ONE* **3**, e2860 (2008).
113. Shannon, C. E. A mathematical theory of communication. *Bell System Techn. J.* **27**, 379–423, 623–656 (1948).
114. Weber, E. U., Shafir, S. & Blais, A. R. Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychol. Rev.* **111**, 430–445 (2004).
115. Mineka, S. & Kihlstrom, J. F. Unpredictable and uncontrollable events: a new perspective on experimental neurosis. *J. Abnorm. Psychol.* **87**, 256–271 (1978).
116. Herry, C. *et al.* Processing of temporal unpredictability in human and animal amygdala. *J. Neurosci.* **27**, 5958–5966 (2007).
117. Anselme, P. The uncertainty processing theory of motivation. *Behav. Brain Res.* **208**, 291–310 (2010).
118. O'Neill, M. & Kobayashi, S. Risky business: disambiguating ambiguity-related responses in the brain. *J. Neurophysiol.* **102**, 645–647 (2009).
119. Vickery, T. J. & Jiang, Y. V. Inferior parietal lobule supports decision making under uncertainty in humans. *Cereb. Cortex* **19**, 916–925 (2009).
120. Cisek, P., Puskas, G. A. & El-Murr, S. Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* **29**, 11560–11571 (2009).
121. Bogacz, R. Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn. Sci.* **11**, 118–125 (2007).
122. Heekeren, H. R., Marrett, S. & Ungerleider, L. G. The neural systems that mediate human perceptual decision making. *Nature Rev. Neurosci.* **9**, 467–479 (2008).
123. Petrusic, W. M. & Baranski, J. V. Judging confidence influences decision processing in comparative judgments. *Psychonom. Bull. Rev.* **10**, 177–183 (2003).
124. Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A. & Ward, B. D. Neural correlates of sensory and decision processes in auditory object identification. *Nature Neurosci.* **7**, 295–301 (2004).
125. Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current research and Theory* (eds Black, A.H. & Prokasy, W.F.) 64–99 (Appleton-Century-Crofts, 1972).
126. Pearce, J. M. & Hall, G. A. Model for Pavlovian learning — variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).
127. Mackintosh, N. J. Theory of attention — variations in associability of stimuli with reinforcement. *Psychol. Rev.* **82**, 276–298 (1975).
128. Preuschoff, K. & Bossaerts, P. Adding prediction risk to the theory of reward learning. *Ann. NY Acad. Sci.* **1104**, 135–146 (2007).
129. d'Acremont, M., Lu, Z. L., Li, X., Van der, L. M. & Bechara, A. Neural correlates of risk prediction error during reinforcement learning in humans. *Neuroimage* **47**, 1929–1939 (2009).

130. Knight, F. H. *Risk, Uncertainty and Profit* (Houghton Mifflin, 1921).
131. Becker, S. W. & Brownson, F. O. What price ambiguity? Or the role of ambiguity in decision making. *J. Polit. Econ.* **72**, 62–73 (1964).
132. Yates, J. F. & Zukowski, L. G. Characterization of ambiguity in decision-making. *Behav. Sci.* **21**, 19–25 (1976).
133. Curley, S. P., Yates, F. & Abrams, R. A. Psychological sources of ambiguity avoidance. *Organ. Behav. Hum. Decis. Process.* **38**, 230–256 (1986).
134. MacCrimmon, K. R. & Larson, S. in *Expected Utility Hypotheses and the Allais Paradox* (eds Allais, M. & Hagen, O.) 333–410 (D. Reidel, 1979).
135. Keren, G. & Gerritsen, L. E. M. On the robustness and possible accounts of ambiguity aversion. *Acta Psychol. (Amst.).* **103**, 149–172 (1999).
136. Slovic, P. & Tversky, A. Who accepts Savage's axiom? *Behav. Sci.* **19**, 368–373 (1974).
137. Pulford, B. D. & Colman, A. M. Size doesn't really matter. Ambiguity aversion in Ellsberg urns with few balls. *Exp. Psychol.* **55**, 31–37 (2008).
138. Larson, J. R. Exploring the external validity of a subjectively weighted utility model of decision making. *Organ. Behav. Hum. Perform.* **26**, 293–304 (1980).
139. Trautmann, S. T., Vieider, F. M. & Wakker, P. P. Causes of ambiguity aversion: known versus unknown preferences. *J. Risk Uncertainty* **36**, 225–243 (2008).
140. Chow, C. C. & Sarin, R. K. Known, unknown, and unknowable certainties. *Theory Decision* **52**, 127–138 (2002).
141. Bach, D. R., Seymour, B. & Dolan, R. J. Neural activity associated with the passive prediction of ambiguity and risk for aversive events. *J. Neurosci.* **29**, 1648–1656 (2009).

**FURTHER INFORMATION**
Dominik R. Bach's homepage: http://www.bachlab.org

ALL LINKS ARE ACTIVE IN THE ONLINE PDF