

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *The Year in Cognitive Neuroscience***Bayesian models: the structure of the world, uncertainty, behavior, and the brain**Iris Vilares^{1,2} and Konrad Kording¹

¹Departments of Physical Medicine and Rehabilitation, Physiology, and Applied Mathematics, Northwestern University, Chicago, Illinois. Rehabilitation Institute of Chicago, Northwestern University, Chicago, Illinois. ²International Neuroscience Doctoral Programme, Champalimad Neuroscience Programme, Instituto Gulbenkian de Ciência, Oeiras, Portugal

Address for correspondence: Iris Vilares, Rehabilitation Institute of Chicago, Northwestern University, 345E Superior St. Oerter, Rm 931, Chicago, IL 60611. iris-vilares@northwestern.edu

Experiments on humans and other animals have shown that uncertainty due to unreliable or incomplete information affects behavior. Recent studies have formalized uncertainty and asked which behaviors would minimize its effect. This formalization results in a wide range of Bayesian models that derive from assumptions about the world, and it often seems unclear how these models relate to one another. In this review, we use the concept of graphical models to analyze differences and commonalities across Bayesian approaches to the modeling of behavioral and neural data. We review behavioral and neural data associated with each type of Bayesian model and explain how these models can be related. We finish with an overview of different theories that propose possible ways in which the brain can represent uncertainty.

Keywords: Bayesian models; uncertainty; graphical models; psychophysics; neural representations

Introduction

What is the purpose of the nervous system or its parts? To successfully study a system, it is crucial to understand its purpose (see “computational level” in Ref. 1). This insight drives a line of research called *normative models*, which starts with an idea of what the objective of a given system could be and then derives what would be the optimal solution to arrive to that objective.⁴ The model predictions are then usually compared with the way the system actually behaves or is organized.² In this way, normative models can test hypotheses about the potential purpose of parts of the nervous system.

Uncertainty is relevant in most situations in which humans need to make decisions and will thus affect the problems to be solved by the brain. For example, we only have noisy senses, and any information that we sense must be ambiguous because we only observe incomplete portions of the world

at any given time, or shadows of reality, as is beautifully illustrated by Plato’s allegory of the cave.³ Therefore, it may be argued that a central purpose of the nervous system is to estimate the state of the world from noisy and incomplete data.^{4,5} Bayesian statistics gives a systematic way of calculating optimal estimates based on noisy or uncertain data. Comparing such optimal behavior with actual behavior yields insights into the way the nervous system works.

Models in Bayesian statistics start with the idea that the nervous system needs to estimate variables in the world that are relevant (x) based on observed information (o), typically coming from our senses (e.g., audition, vision, olfaction). Bayes rule⁶ then allows calculating how likely each potential estimate x is given the observed information o :

$$p(x|o) = p(o|x)p(x)/p(o).$$

For example, consider the case of estimating if our cat is hungry (x = “hungry”) given that it meowed (o = “meow”). If the cat usually meows if hungry ($p(o|x) = 90\%$), is hungry a quarter of the time ($p(x) = 25\%$), and does not meow frequently

⁴ Normative models contrast with descriptive models, which only describe the solutions without evaluating how useful such a solution would be.

($p(o) = 30\%$), then it is quite probably hungry when it meows ($p(x|o) = 90\% \times 25\% / 30\% = 75\%$). If, on the other hand, the cat meows quite frequently ($p(o) = 70\%$), then the probability of being hungry if it meowed is much lower ($p(x|o) = 90\% \times 25\% / 70\% = 32\%$). The formula above can also be seen as a way of updating the previous belief about the world, or prior ($p(x)$) by the current sensory evidence, or likelihood ($p(o|x)$).^b

Models using Bayes rule have been used to explain many results in perception, action, neural coding, and cognition. Bayesian models that have been used in these contexts have many different forms. The differences between these models derive from distinct assumptions about the variables in the world and the way they relate to one another. Each model is then the unique consequence of one set of assumptions about the world. However, all these Bayesian models share the same basic principle that different pieces of information can be combined in order to estimate the relevant variables.

Bayesian statisticians have developed a way of depicting how random variables relate to one another—by using graphical models.^{7–9} We should mention here that several kinds of graphical models can be used.^{7–11} Types of graphical models include factor graphs, Markov random fields, ancestral graphs, and *Bayesian network* or *directed acyclic graphs*, which we discuss here. Bayesian networks/directed acyclic graphs are a type of graphical model that, besides indicating purely statistical relations, can be interpreted as a model of the causal structure in the world.⁹ Therefore, we will use the Bayesian network class of graphical models to depict the different types of existing Bayesian models and structure this review.

In Bayesian networks, there are two kinds of random variables: those that are observed and those that are not observed. For example, we can hear our cat meow (observed variable) but we can never directly observe that it is hungry, only infer that from its behavior. Therefore, hunger is fundamentally an unobserved variable—unless some new neurophysiologic procedure is invented that measures Qualia. Unobserved variables are called latent or, as we will call it here, hidden. These hidden variables

are typically estimated given the observed variables in Bayesian modeling.

There are many relevant reviews of and books about Bayesian methods. Some focus on cue combination,^{12–15} while others focus on general Bayesian estimation,^{8,16–18} and yet others focus on the information integration for making choices^{2,19,20} or continuous control²¹ and in possible representations of uncertainty.²² These reviews provide excellent introductions into the mathematical treatment of estimation in various settings. Here, we want to instead focus on the structure each Bayesian model assumes about the world and give a taxonomy of these Bayesian models by focusing on the underlying graphical models.

Our review is structured as follows. In each section we will cover progressively more difficult problems to which Bayesian models are frequently applied. Each section starts with a simple example of the problem being dealt with and a graphical model that represents it, continues by making a short review of behavioral and modeling work related to that issue, and ends with available neural data that indicate where and how Bayesian computation can be occurring. We will end this review with an overview of proposals of how the brain may represent uncertainty.

Cue combination

Imagine that our cat ran away to avoid having a bath and is now hiding in the garden. It made the bushes move, providing a visual cue about its position, and meows, which provides an auditory cue (Fig. 1A). This is a typical example where one variable, the hidden location of the cat, is reflected in observed variables, in this case audition and vision, which are called cues.

In the graphical model associated with this example (Fig. 1B), the hidden variable (e.g., the cat's position) that we want to estimate is assumed to independently cause the observed visual and auditory cues. The circle representing the cat's position variable has no incoming arrows, indicating that there will be only a prior belief ($P(\text{position})$, see next section). The vision circle (O_v) has an incoming arrow from position, indicating that the probability of having a specific visual observation depends on the value of the position variable ($P(\text{vision} | \text{position})$), i.e., the probability of seeing a specific bush moving depends on where the cat is hiding. Similarly, the

^b Divided by a normalizing constant ($p(o)$).

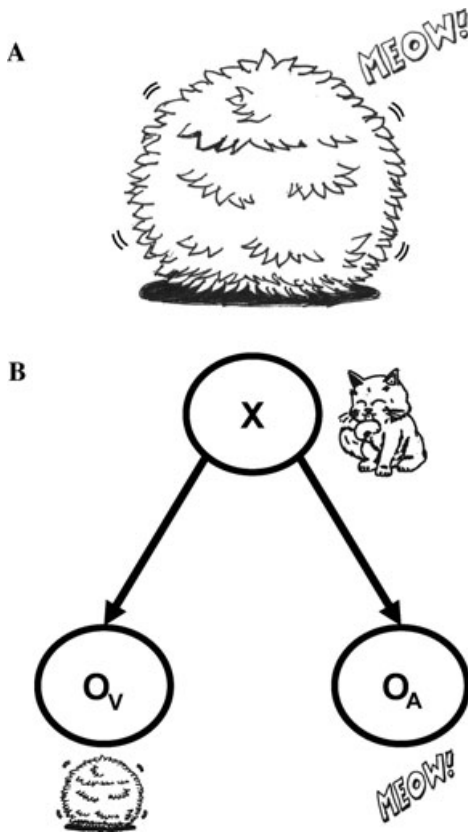


Figure 1. Cue combination. (A) Example of an indirect observation of a cat’s position. You can see the bush moving and hear a “meow” sound, but you cannot directly observe the cat. (B) Graphical model of what is seen in A. The variable X (position of the cat) is unobserved, but it produces two observed variables: the moving bush, which provides a visual cue (O_V), and the “meow” sound, which provides an auditory cue (O_A). Cartoons made by Hugo M. Martins.

audition circle (O_A) has an incoming arrow from position, and we thus have $P(\text{audition}|\text{position})$. Jointly, these pieces of information define the joint probability distribution of the three random variables:

$$\begin{aligned}
 &P(\text{position, vision, audition}) \\
 &= p(\text{position})p(\text{vision}|\text{position}) \\
 &\times p(\text{audition}|\text{position}).
 \end{aligned}$$

Importantly, the fact that there is no arrow directly connecting vision and audition indicates that, conditioned on the knowledge of position, vision and audition are independent. This can be seen as indicating that noise in vision is independent of noise in audition. This assumption about the way visual and

auditory cues are generated may be true or not, but it enables the construction of a simple normative model of behavior. The central assumptions about the world that give rise to each Bayesian model (in this case, the assumption of how cues are generated) can thus be effectively formalized in a graphical model.

In our example, we can say that the goal of our nervous system is to discover where the cat is, i.e., estimate the hidden variable, “position of the cat.” Assuming that this variable generated the observed visual and auditory cues, the nervous system has to invert this generative process and estimate the hidden variable position of the cat by combining the visual and auditory cues.

Bayesian statistics provides a way of calculating how to optimally combine the cues, i.e., a way of maximally reducing the final uncertainty about the cat’s position. The resulting estimate combines the cues, weighting them according to their reliability. If people combine cue information in a Bayesian way, the resulting estimate has generally lower uncertainty than estimates based on any of the cues alone.

This property of lower uncertainty in the final estimate is one of the crucial advantages of behavior that combines different pieces of knowledge in a way predicted by Bayesian models. Previously, it has been suggested that the nervous system could use a winner-take-all approach, taking only into account the most reliable cue (in this case, generally vision).²³ However, this would result in a final uncertainty generally higher than what would be obtained by employing Bayesian statistics.⁶ Furthermore, an approach in which only the most reliable cue is used cannot explain the cue interaction effects observed in daily life (e.g., the McGurk effect²⁴).

Many experiments have probed how humans estimate hidden variables given more than one cue, and the results are in accordance to what would be predicted by Bayes theory.^{25–31} The results of many of these studies have been framed in terms of Bayesian statistics, and these studies almost invariably assume the same graphical model (Fig. 1B). The typical experimental strategy is as follows. They first measure independently the uncertainties associated

⁶ Being only equal if one of the cues has no uncertainty associated (i.e., its variance is zero).

with each cue. Then, based on these estimates, they calculate what would be the weighting parameter (w) that would optimally combine both cues. They measure the way people actually combine cues when they have both cues available and compare the results with the model predictions.^d Importantly, the model predicts behavior in one condition (cue combination) based on the subject's behavior in different situations (with only one cue).

Using an experimental setup like the one just described, Ernst and Banks have discovered that subjects use information from both vision and the sense of touch in order to estimate sizes, and they combine these cues in a way close to the statistical optimal.³⁰ Other studies also showed that the variability in the estimates obtained when subjects combined proprioceptive and visual information was smaller than the variability obtained when they could only use one of the senses, a phenomenon also predicted by Bayes theory.²⁸ Furthermore, they tended to rely more on the most accurate cue.²⁷

This close-to-optimal combination of different sensory cues has also been demonstrated in many other experiments and sensory modalities. For example, it has been found that subjects can combine optimally visual and auditory cues.^{32,33} This combination is on the basis of the McGurk effect, in which when there is a discrepancy between what the lips of a person are saying and the actual sound, subjects hear a syllable that is a mix between the visual and the auditory syllable.²⁴ Cue combination has also been found within the visual system, with subjects combining cues of texture and motion,³⁴ texture and binocular disparity,^{35–37} or even two texture properties³¹ in order to estimate some position or slant. There is thus now ample evidence that in many if not most situations, cues are combined in a way close to the Bayesian optimal,^{31–37} which has been discussed in a good number of recent reviews.^{12–14,38–41}

The cue combination studies we mentioned so far assume continuous cues and the estimation of continuous hidden variables, like the position of the hand or the distance of an object. However, there are also cases in which we want to estimate

discrete variables—for example, how often someone has touched our hand or how often a light has flashed. In such cases, Bayesian models have also been shown to fit well human behavior.^{42–44} The finding of near-optimal cue combination thus applies not only to continuous variables, but also to situations where discrete numbers are estimated.

If people are Bayesian in their behavior, this means that the brain has to somehow represent and use uncertainty information for cue combination. How does the nervous system represent cues and their reliabilities to be able to combine them? We will discuss theoretical proposals of how the nervous system may represent uncertainty in the final section of this review, but focus now on available electrophysiological and imaging results specifically related to cue combination.

Many brain areas have been implied to participate in multisensory integration (for reviews, see Refs. 40, 45–46). For example, in the superior colliculus (a brain area that receives visual, auditory and, somatosensory inputs), neuronal responses to a given sensory stimulus were influenced by the existence or nonexistence of other sensory cues.⁴⁷ Multisensory integration has also been analyzed in the superior temporal sulcus (STS), where it was found that, in a fraction of neurons responsive to the sight of an action, the sound of that action significantly modulated their visual response.⁴⁸ Neurons in the STS thus appear to form multisensory representations of observed actions. Further evidence of multisensory integration comes from imaging studies where activity in higher visual areas (hMT1/V5 and lateral occipital complex) of human subjects suggested combination of both binocular disparity and perspective visual cues, potentially in order to arrive at a unified 3D visual perception.⁴⁹

Bayesian models have inspired research on neurons in the dorsal part of the medial superior temporal area (MSTd) in the monkey. These neurons have been shown not only to integrate visual and vestibular sensory information, but also to do so in a way that closely resembles a Bayesian integrator, i.e., these neurons sum inputs linearly and subadditively (meaning that each cue weights one or less), and, moreover, the weights that they give to each cue change with the relative reliability of each cue.^{50,51}

Cue combination literature usually analyzes cases where the nervous system is likely to assume that there is a single hidden variable and multiple cues

^d Although it would be possible for people to combine information from more than two cues, for simplicity studies generally just try to understand how people combine two cues.

that are indicative about that variable. Both this variable and the observed variables may be continuous or discrete, bounded or unbounded (i.e., a distribution may be between 0 and 1 or $-\infty$ and $+\infty$), but the same graphical model describes all these models. While the models are different, the solution strategy is essentially the same—the probability distributions associated with each cue simply get multiplied together using Bayes rule. These models have been applied to many situations and describe the results of many experiments.

Combining a cue with prior knowledge

Even if only one cue is present, we can combine it with previous knowledge (prior) that we have about the hidden variable in order to better estimate it.^e In our example, if you can hear the cat but you cannot see it (Fig. 2A), you might still recall previous times in which the cat ran away to the garden: you have prior knowledge of probable positions of the cat, and can use it to estimate the cat’s position. Again, a hidden variable (position) causes the observed cue (the “meow” sound), a relation that can be formulated by a graphical model (Fig. 2B). In this case, estimates should depend on what we have learned in the past about the hidden variable (prior) and on our current observation (likelihood).

Indeed, standard Bayesian prior-likelihood experiments and cue combination experiments explained above are essentially based on the same graphical model and the mathematical treatment is analogous. The only difference is that in cue combination, there are at least two cues and priors are usually ignored and, in prior-likelihood experiments, the prior is incorporated but other potential cues are usually ignored. Priors can be considered a simple summary of the past information subjects have had in a particular task.⁵² Moreover, they are independent of the current feedback, i.e., the prior is the

^e It should be mentioned here that in cases where cues are combined, there are usually priors at play as well. However, we did not discuss these priors in the cue-combination section because most studies in the cue-combination field ignore priors. Some studies ignore priors because they are not important—for example, when the prior is much wider than the likelihood. The majority of studies on cue combination, however, have been designed in such a way that priors have no effect. These studies use two-alternative-forced-choice (2AFC) paradigms, which remove any effect of priors.

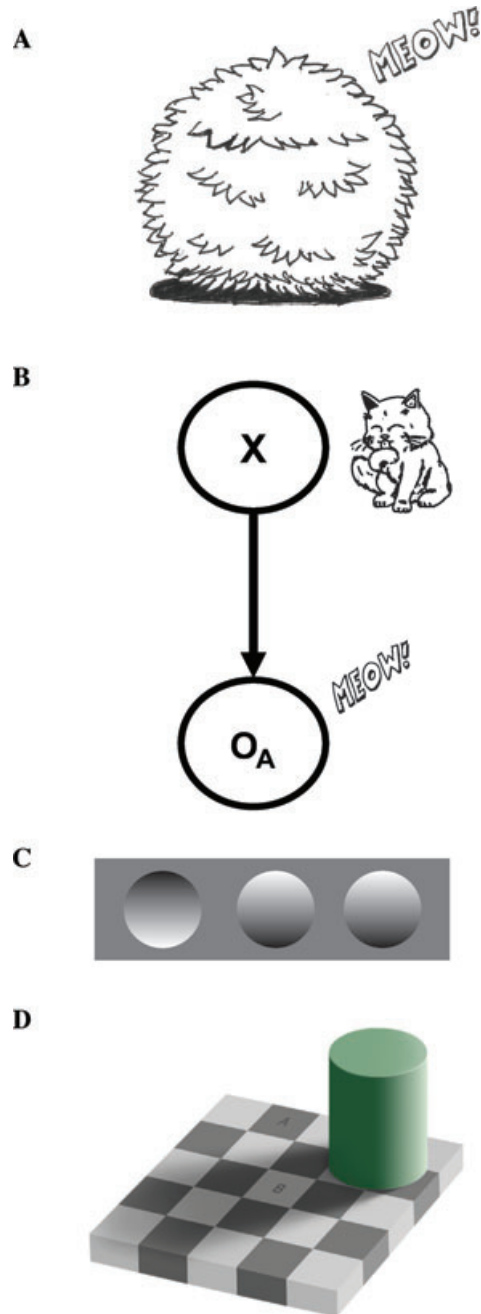


Figure 2. Combining a cue with prior knowledge. (A) Example of an indirect observation of a cat’s position, but in which only a “meow” can be heard. (B) Graphical model of what is seen in A. The variable X (position of the cat) is unobserved, but it produces the observed “meow” sound, which provides an auditory cue (O_A). (C) Example of a visual illusion. Here, people generally see one groove (left) and two bumps, but if the paper is rotated 180°, then two grooves and one bump are perceived. (D) Checker-shadow illusion.⁵⁴ In this visual illusion, the rectangle A appears to be darker than B, while in reality they have the same color.

information we get from all previous trials except the current one.⁵³ While it was generally ignored in previous models, including most of cue combination discussed above, prior knowledge clearly influences our decisions and even our perception of the world.

There are many examples of the effect of prior knowledge on perception. For example, in Figure 2C, we can see one grove and two bumps, but if we rotate the paper by 180 degrees, their perceptual depth shifts (we see two groves and one bump instead). This occurs because people have the prior assumption that light should come from above. This is also beautifully illustrated in the checker-shadow illusion, where the prior assumption of a light source that casts the observed shadow makes the rectangles A and B appear to have different brightness.⁵⁴ These sensory biases can then be explained with the incorporation of prior information on the final sensory perception.

Experimental studies have shown that priors can indeed be learned over time⁵² and that they are independent of the current sensory feedback.⁵³ A diverse set of studies has also shown that people combine previously acquired knowledge (prior) with new sensory information (likelihood) in a way that is close to the optimum prescribed by Bayesian statistics. For example, when performing arm-reaching tasks,^{55,56} pointing tasks,⁵⁷ or even timing tasks,⁵⁸ people take into account both prior and likelihood information and, moreover, they do so in a way compatible with Bayesian statistics, given more weight to the more reliable “cue” or, in other words, they rely a lot on the prior when likelihood is relatively bad and vice versa.^{55–59}

As indicated by Bayes-like behavior when combining prior and likelihood, the brain needs to represent and use uncertainty of both previously acquired information and current sensory feedback in order to optimally combine these pieces of information. How does the nervous system represent prior and likelihood and their associated uncertainties? Currently, there is relatively little known about this (although some theories have been proposed; see the last section of this review). However, when it comes to movement, it has been found that the dorsal premotor cortex and the primary motor cortex encode multiple potential reaching directions, indicating a potential representation of priors.^{60–64} However, we clearly do not understand

yet how the nervous system integrates priors and likelihoods.

Combining information across time

In the previous sections, we discussed how information from multiple senses and from priors can be combined in order to arrive at better estimates about the world. However, many hidden variables in the world change over time—for example, the cat’s position is changing as the cat moves through the garden, and we obtain information about its whereabouts at different points in time (Fig. 3A). Such cases are captured by graphical models where the hidden variable exists with potentially different values as a function of time (Fig. 3B). The graphical model highlights the so-called Markov property: conditioned on the present state of the world, the past is independent from the future.^f

The situation and graphical model depicted in Figure 3 can be seen as an extension of the cue combination and prior likelihood models, but in which the hidden variable can change over time. Thus, at any point of time, the model is identical to the prior-likelihood integration model, and the joint estimate obtained at that point of time will then be the updated prior for the next point of time.

The graphical model in Figure 3B underlies a range of related models. Models of this type that estimate discrete hidden variables are usually called hidden Markov models^g (HMM). When the variables are continuous and probability distributions are Gaussians, then the model is usually called a Kalman model.⁶⁵ When these models are used to make estimates about a given point of time, given only the past, they are generally called *filters*; when they make estimates given both past and future, they are called *smoothers*. Still, all these approaches share the same graphical model that specifies the Markov property and only differ in specific additional assumptions.

To make estimates, new sensory information obtained at each point of time needs to be combined

^f The Markov property is named after Andrey Markov, one of the pioneers of the theory of stochastic processes.

^g It should be noted here that any model with the structure of Figure 3b is a hidden Markov model, as the hidden variable has the Markov property. However, usually this name is only used for discrete models.

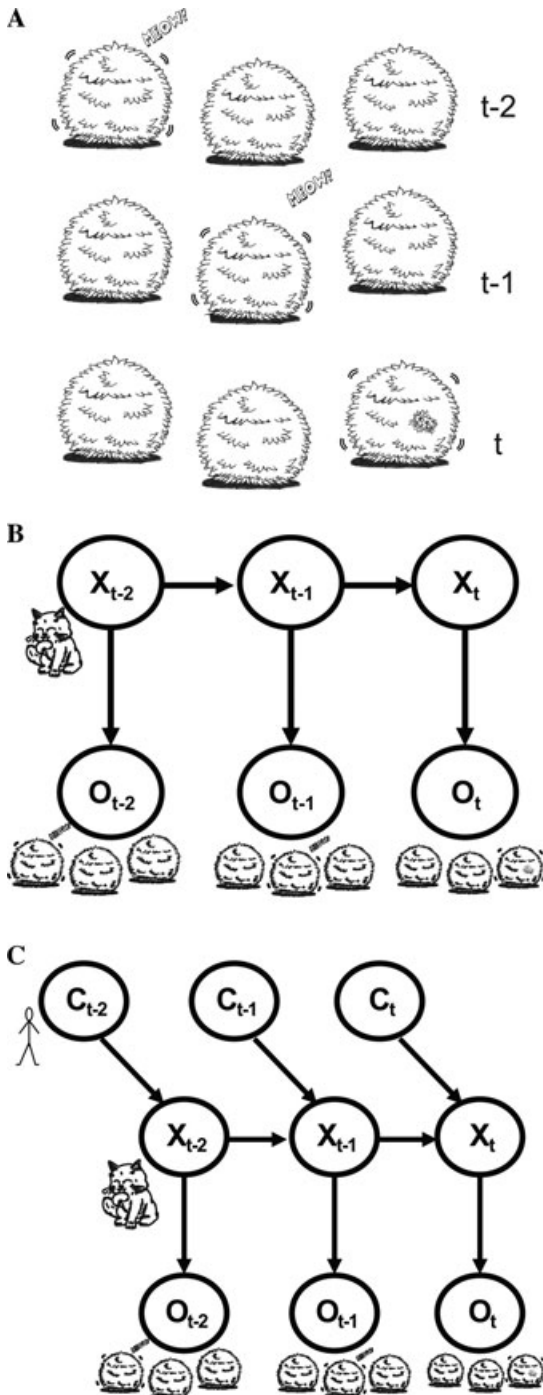


Figure 3. Combining information across time. (A) Example of an indirect observation of a cat’s position, at different points of time ($t-2$, $t-1$, and t , t being the present time). (B) Graphical model of what is seen in A. The hidden variable X (position of the cat) at each point of time produces a variable O that is observed. (C) Graphical model similar to B, but in which the external effect of a controller (in this case, a person) is incorporated in the model, which will affect X .

with the current belief (prior). A filter, e.g., the Kalman filter, can be seen as a system that alternates between two steps: (1) cues are combined with current beliefs using Bayes rule and (2) the dynamics of the world (in our example, how fast the cat changes position) affect our estimates regarding the state of the world and thus our belief. Some of the dynamics are generally unpredictable, making us less certain about the world (if the world is changing unpredictably, we become uncertain about it), while the cues we receive generally make us more certain about the world. The interplay between unpredictable changes and observation defines uncertainty in such Bayesian models.

When we interact with the world, our movements also affect the dynamics of the world. This means that the state of the world at the next point in time depends on its previous state as well as our own state. For example, our own movements will affect the cat’s movements as well, and therefore we should incorporate our own movements into the estimates. This situation is captured by the graphical model that underlies the Kalman controller (Fig. 3C). As the model retains the Markov property, and given that we know our own motor commands, solutions to this problem are very similar to those of the Kalman filter.

These kinds of situations, in which we sense and interact with the world, have previously been described by other models, such as state space models.^{66,67} In these models, a set of input, output, and state variables are related by first-order differential equations. In them, like in a Kalman filter, there is as well a kind of “internal model” or “belief” that can change with time. However, in contrast with a Kalman filter, these models do not incorporate uncertainty. Uncertainty is maybe one of the most unique features of humans and other biological controllers, as we have to sense and act in noisy and uncertain environments.⁶⁸ Therefore, interaction with an environment, like the learning of a new motor command, should depend on the uncertainty of the environment. This idea, brought by Bayesian models, that uncertainty should play a role, has led to new models and new experiments to test those models. The incorporation of uncertainty in these new models permitted them, perhaps not surprisingly, to better characterize behavior.

Models of the Kalman filter or Kalman controller type have been successfully applied to explain

movement and perceptual data. For example, they have been used to explain how human subjects estimate their hand position during and after movements.^{69,70} They have also been shown to predict salient aspects in the control of posture.^{71–74} During movement, the nervous system constantly needs to estimate the state of the body and the world, and Kalman filter–based algorithms are natural solutions to this kind of problem.

The same kind of Kalman filter model has been used to make estimates over longer periods of time, and learning has been conceptualized as such a long-term estimation.⁷⁵ The idea behind this conceptualization is that learning is the process by which we obtain knowledge about the world. If the world is changing rapidly, then we need to learn and forget rapidly, and vice versa. For example, the way human subjects adapt to force fields,⁷⁶ visuo-motor perturbations,⁷⁷ use available sensory information to minimize arm-reaching errors over time,^{78,79} and the way monkeys adapt their saccades⁸⁰ are all phenomena that have been modeled by conceptualizing learning as a form of Bayesian estimation.

Situations in which we interact with the world generally ask for Kalman controllers since our movement affects the state of the world. Such estimators are naturally part of many approaches to optimal control.^{21,81,82} Many human behaviors have been shown to be close to optimal in the optimal control sense.^{70,83–86} Overall, people tend to efficiently combine their own motor commands and sense cues into continuous estimates of the properties of the world.

Temporal models of this kind have also been extensively used to decode the state of the brain. For example, in many applications, scientists want to estimate the intent of an animal or a patient based on neural recordings. For example, say we want to estimate where a locked-in patient would like to have his hand. The position where the person wants the hand to be will probably change smoothly over time. A Kalman-filtering approach allows combining knowledge of typical movements (defining the state dynamics) with ongoing recordings of neural activities.^{87–91}

Lastly, there are a range of theories of how the nervous system can implement something like a HMM or Kalman filter.^{92,93} The brain appears to be good at integrating information over time. How exactly it achieves that is an exciting topic of ongoing research.

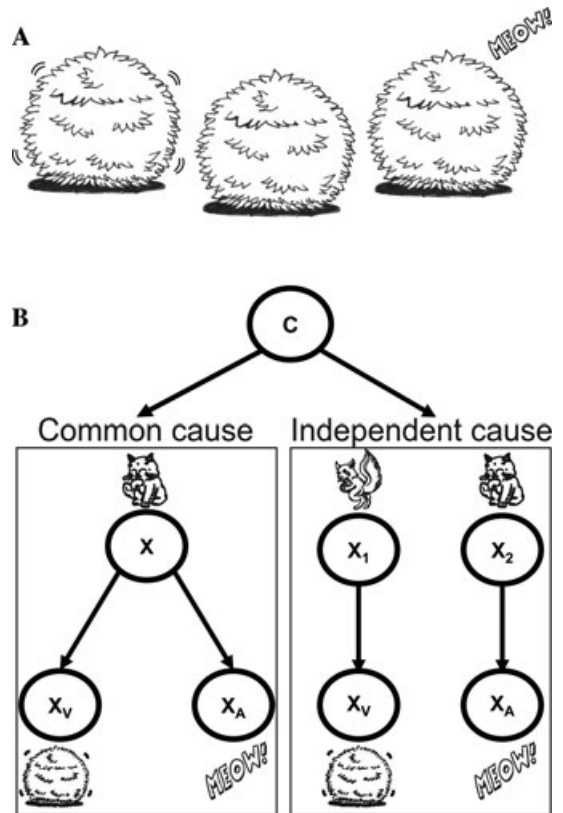


Figure 4. Inferring the causal structure of the world. (A) Example of an indirect observation of a cat’s position. In this case, the movement of the bush is seen in one direction, but the “meow” seems to come from somewhere else. (B) Graphical model of what is seen in A. The visual and the auditory cue may have a common cause (left box) or randomly co-occurring, independent causes (right box).

Inferring the causal structure of the world

Up to now, we have discussed cases where the structure of the world and thus the relations between hidden and observed variables is known. However, there are many cases where relations are not known. Let us consider again the example of estimating the cat’s position, but now imagine that the sound comes from one direction and the movement happens in a very different direction (Fig. 4A). In this case, it is unlikely that the cat caused both the movement and the meow and it is more likely that there were independent causes. In such situations we have uncertainty about the causal structure.

We can formalize such problems as a mixture problem (see Fig. 4B). If we hear and see something at the same time, there are two causal

interpretations. Either one variable (e.g., the cat) caused both cues or, alternatively, there may have been two independent causes (e.g., a cat making a meow and a squirrel making the bush move). In such cases, we can enumerate all the possible causal structures (in this case, the common cause assumption and the independent cause assumption). After enumerating the possible causal structures, assumptions are made of how likely each causal structure is, and together this specifies assumptions about the world.

In the previous examples given, we were dealing with the issue of estimating the values of hidden variables. In these mixture problems, besides estimating the hidden variables, it is also necessary to estimate the causal structure of the world (causal inference). In our example, if the cat caused both cues, then we want to combine them. However, if the cat only caused the auditory cue, then we do not want to combine it with the moving bush information. Nevertheless, the same Bayesian methods can be used—we can calculate how likely each causal structure is given the data and consider all of them or, alternatively, the most likely one.

The act of trying to find the causal structure of the world is a particularly important one. Knowing the right causal structure can allow us to do better estimates and make predictions.^{h,9} However, this is a particularly difficult task. Given that we do not observe directly the causal relation, we have to infer it from the observed sensory input,^{9,94} and sometimes this input can be quite unclear (say if the moving bush and the meow are just slightly apart). A wrong causal inference can have deleterious effects, from the simple miscalculation of the cat's position (e.g., estimating it is in the middle while it is much more on the right side) to incorrect attributions of causal relations with impact in people's life and well-being (e.g., in schizophrenia/delusion, patients often attribute a wrong causal role to, say, CIA, for common effects observed in their life).

A good amount of recent behavioral work has shown that these phenomena are in place when people combine visual cues,⁹⁵ when they combine

vision and audition for perception,^{96,97} and when they combine vision and proprioception for motor learning.⁹⁸ It has also been discussed to what extent the system chooses the best causal interpretation or considers all of them.^{99–101} A range of recent reviews discuss this general problem.^{12,38,40} Importantly, the same Bayesian framework that we explained in previous sections can be used, only now the causal structure is part of the estimation problem.

There is some emerging work on the neural implementation of causal inference.¹⁰² However, so far we know relatively little about the way the brain represents the causal structure of the environment.

We want to emphasize here that causal inference emerges quite naturally from cue combination. Given each causal structure (i.e., when we are within each box in Fig. 4B), the same rules apply as for cue combination. Only now, besides calculating the best estimates for each model, we also have to estimate how much trust we place into each causal model.

Inference in systems with switching dynamics

Above, we have seen how cue combination is extended to causal inference by assuming uncertainty about the causes behind the observed data. We have also seen how cue combination becomes filtering (e.g., in the Kalman filter) by simply assuming that relevant hidden variables change over time and thus have dynamics. These two may be combined into a switching system—a system where the dynamics of the hidden variables are not always constant but change at given points of time.

Switching systems are important in the context of movement modeling. For example, a cat may sometimes walk and sometimes run and the consequent temporal dynamics of the cat's position are different if the cat is walking or running. A good statistical model of cat locomotion should account for this—having one model for running and another for walking along with a model for transitions between the two.

These mixture approaches are of practical importance the context of neural decoding. In such applications, scientists often want to estimate the user's intent based on neural recordings—often with the objective of enabling prosthetic devices to restore function. For example, when we want to decode intended movement from neural activities, we may

^h For example, if we know that the cat caused both the movement in the bushes and the sound, we can infer that if we would hypothetically remove the cat, then the sound and the movement would disappear.

want to model the fact that there may be a number of different targets of reaches, each associated with distinct dynamics.^{103,104} Similarly, behavior may transition between times of rest and times of movement, and detecting such changes has been shown to improve decoding quality.¹⁰⁵

A recent paper has proposed how the nervous system could implement such switching system. It assumes a neurally implemented Kalman filter that can rapidly change its state estimate in a way that is triggered by specific neuromodulators.¹⁰⁶ However, we are not aware of specific neural data about the way the nervous system actually implements mixture models with temporal dynamics.

Generative models for visual scenes

So far we have focused on cases that are close to sensorimotor integration. Here, we want to discuss a set of applications of graphical models that have a very different flavor and yet use the same general framework. Specifically, we want to discuss Bayesian models that derive algorithms for object recognition and scene segmentation based on assumptions of the way visual scenes are made.

One way of conceptualizing how a visual scene is made is by assuming that we start with an empty image and keep filling it with objects until we have obtained our final image. We may fill the scene with different kinds of objects—e.g., animals, texts, and textures. In fact, this is exactly how recent objects of internet fame¹⁰⁷ are made (Fig. 5B). This conceptualization of how a visual scene is made can be readily converted into a graphical model (Fig. 5A)—although it should be noted that the number of objects, texts, and textures visible in a scene needs to be estimated as well (which is akin to causal inference) and could be drawn using plate notation.¹⁰⁸ In plate notation, instead of using a circle to denote one random variable, a rectangle with a letter is used, which implies that an unknown (potentially) infinite number of variables may exist. Each image is then caused by an unknown number of objects, textures, and texts, the number of which is also estimated according to Bayesian rules.

This Bayesian approach of dealing with a visual scene is said to be generative model-based as it assumes a scene that causes objects, texts, and textures that, by their turn, generate the observed visual image. Previous computational approaches dealing with visual scenes had a more ad-hoc approach. For

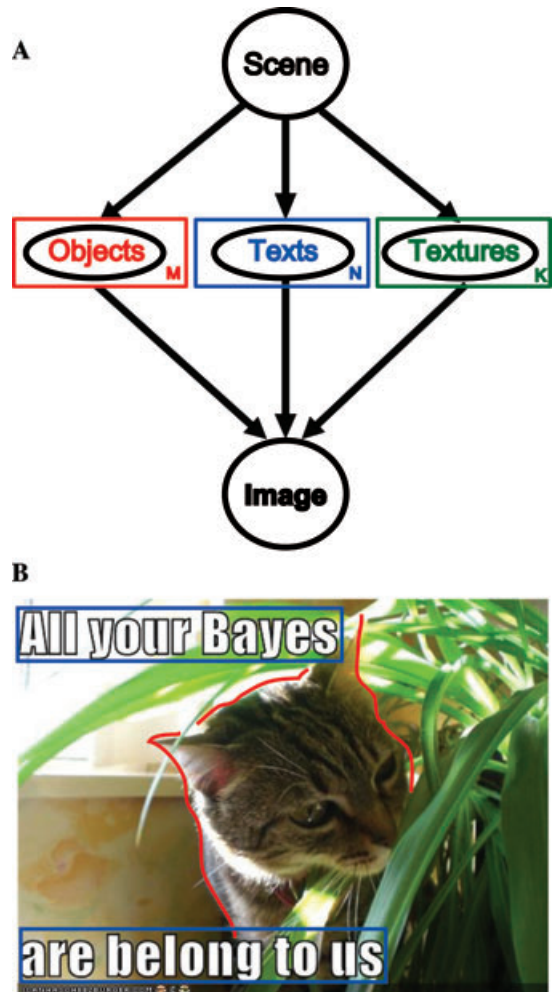


Figure 5. Generative models for visual scenes. (A) Graphical model of a generative model for visual scenes. In this model, a scene starts empty and is then filled with different concepts: objects (animals; people. . .), texts (written text), and texture. The small letters in the lower left corner of the boxes, M , N , and K represent the (unknown) number of objects, texts, and textures present in the image. (B) Example of a visual scene. Highlighted in red is the “object,” in blue is the “text,” and the rest of the image can be considered “texture.”

example, in normalized cut,¹⁰⁹ the way to segment an image follows a set of heuristics that tries to maximize similarity within groups and dissimilarity between the different groups, but these heuristics are not based on an understanding of how images are generated by the world.

The Bayesian approach has been successfully used in a number of recent studies in computer vision. For example, it has been used to model complex

scenes^{110,111} and to model objects and their parts.¹¹² The problems are very dissimilar to the ones we discussed above, and yet the mathematical approach of inferring the causal structure is very similar.

Bayesian decision making

So far, we have discussed ways of estimating hidden variables, but that is only the first step in decision-making systems. In our example, after estimating the probabilities associated for each potential cat's location around the garden, we have to choose how to capture it. Depending on the way we do that, we may incur in a cost—for example, by stepping on the cat's tail. In the field of economics, such costs are generally described as negative utility; utility measures the subjective value of any possible situation. Decision theory deals with this problem of choosing the right action given uncertainty, generally by calculating the action that maximizes (positive) expected utility.¹¹³ To make good decisions, we need to combine our uncertain knowledge of the world with the potential rewards and costs we may encounter.

Bayesian decision making can thus be seen as the important final step in all the models explained above. The models explained in previous sections give us potential optimal ways of perceiving the world, but if we then want to act upon it, we should also take into account the potential rewards and risks associated with each estimate.

Sensorimotor research has shown that human subjects, when doing a movement task, besides being able to estimate their motor uncertainties,⁸⁶ can take into account both rewards and penalties associated with the said task and aim their movements in a way that maximizes expected utility.¹¹⁴ This is in contrast to many high-level economics tasks, where human subjects exhibit a wide range of deviations from optimality.^{115,116}

To maximize expected utility, our brain has to represent not only the reward or cost value for each action, but also the associated uncertainty. The neural representation of these variables has been the focus of the emerging field of neuroeconomics. Neuroeconomics tries to understand the neural processes that occur during decision making within the framework of Bayesian decision theory.^{39,117,118} Responses to reward value and reward probability have been identified in neurons in the orbitofrontal cortex, striatum, amygdala, and dopamine neurons of the midbrain.^{119–124}

Lately, research in neuroeconomics has been shifting focus from the mere coding of expected value and magnitude of reward to neural representations of reward uncertainty, in the form of both risk and ambiguity. Uncertainty in reward has been hypothesized to be represented in some of the brain areas typically associated with reward coding, such as dopamine neurons,¹¹⁹ the amygdala,¹²⁵ the orbitofrontal cortex, and the striatum.^{126,127} However, it has been noticed that neuronal activations related to uncertainty in reward (more specifically, to risk) seem to be segregated spatially and temporally from activations due to expected reward, with activations due to risk occurring later than the immediate activations due to expected reward.¹²⁶ Besides these activations in reward-related areas, uncertainty in reward has also been associated with unique activations in the insula^{128,129} and in the cingulate cortex.^{130,131} Thus, reward value and risk have been associated with spatially and temporally distinct neural activations in specific brain areas, suggesting that the brain can use both sources of information to estimate expected utility and guide actions.

Neural representations of uncertainty

So far, we have discussed how variables, both hidden and observed, relate to one another and how these models have given rise to behavioral and neurophysiological experiments. Here, we want to more generally discuss how the nervous system may represent uncertainty. We will consider a range of theories that have been put forward to describe how neurons may represent uncertainty and also discuss actual neural data that can be related to these theories.

Imagine we are recording from a neuron somewhere in the nervous system. How could the firing of that neuron indicate the level of uncertainty? Importantly, most neurons in the nervous system appear to have tuning curves: if the world, our cognitive state, or our movement changes, then the firing rate of the neuron changes. There are thus at least two ways how uncertainty could be encoded: either there are specialized neurons that encode uncertainty and nothing else or, alternatively, neurons may have tuning curves representing some variable and encode uncertainty at the same time.

The first theory is the most simple: there may be a subset of neurons that only encode uncertainty (Fig. 6A), for example, using the neuromodulators

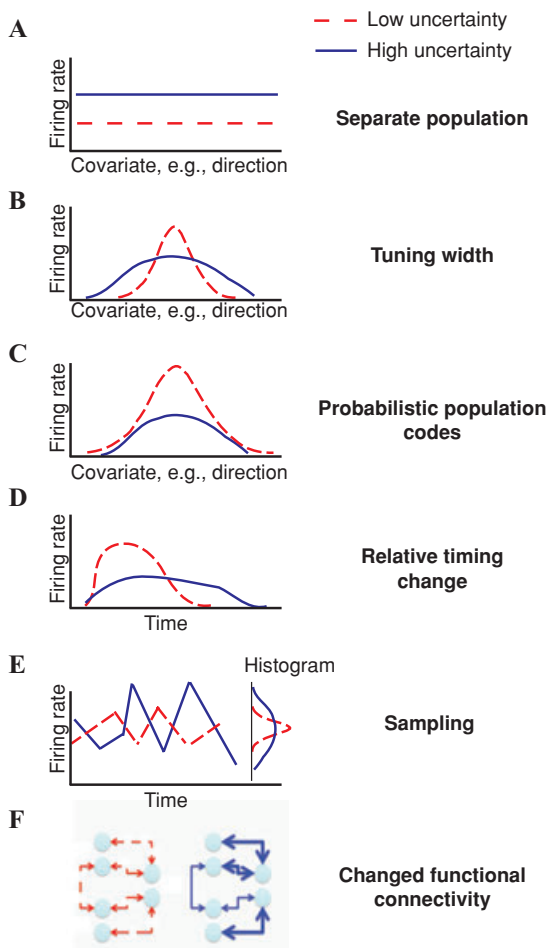


Figure 6. Possible neural representations of uncertainty. In red are the putative firing rates (or connections) in a low-uncertainty state, and in the blue are the ones occurring in a high-uncertainty state. Panels A through F represent different theories that have been proposed on how the brain could be representing uncertainty.

acetylcholine or norepinephrine.¹⁰⁶ This theory appears to have significant experimental support. For example, some experiments, as discussed above, indicate that uncertainty about reward appears to be represented by groups of dopaminergic neurons in the substantia nigra, insula, orbitofrontal cortex, cingulate cortex, and amygdala.^{119, 128, 129, 131} However, even neurons in these areas have clear tuning to other variables. Also, the experiments in support of this theory generally refer to high-level uncertainty, such as the uncertainty associated with potential rewards, and not uncertainty that is related to say sensory or motor information, which can be

considered more “low-level uncertainty” and might be represented in a different way.

A second possibility states that the width of tuning curves may change with uncertainty¹³² and that the neurons may jointly encode probability distributions (Fig. 6B). Such a joint encoding makes sense given that the visual system exhibits far more neurons than inputs¹³³ and the extra neurons could encode probability distributions instead of point estimates. When uncertainty is high, then a broad set of neurons will be active but exhibit low activity, while at low uncertainty, only few neurons are active but with high firing rates (see Fig. 6B). Support for this theory comes from early visual physiology where spatial frequency tuning curves of neurons in the retina are larger during darkness (when there is more visual uncertainty) than during the day.¹³⁴

A third influential theory of the encoding of uncertainty is the so-called probabilistic population code, or PPC (Fig. 6C). This theory starts with the observation that the Poisson-like firing observed for most neurons automatically implies uncertainty about the stimulus that is driving a neuron.¹³⁵ In this way, neurons transmit the stimulus information while at the same time jointly transmitting the uncertainty associated with that stimulus. Specifically, the standard versions of this theory predict that increased firing rates of neurons imply decreased levels of uncertainty. Some data in support of this theory come from studies on cue combination.^{50, 51} More support comes from the general finding that early visual activity is higher when contrast is higher and thus uncertainty is lower.^{136–138} For Poisson-like variability, however, not a lot of experimental support exists, and more advanced population decoding studies are needed.

Another theory that has been put forward suggests that while the tuning curves stay the same, the relative timing of signals may change (Fig. 6D).^{139, 140} If uncertainty is low, then neurons will fire a lot and quickly when a stimulus is given, but will more quickly stop firing. If, on the other hand, uncertainty is high, then neurons fire less but for a longer time. In that way, the total number of spikes may be the same, but their relative timing changes. There is some evidence for this theory coming from studies in the area MT that shows differential temporal modulation when animals are more uncertain.¹⁴¹

Another theory is the sampling hypothesis.^{142–145} According to this theory, neurons will spike a range of instantaneous firing rates that is narrow over time if the nervous system is certain about a variable and has a wider range if it is less certain (see Fig. 6E). Evidence for the sampling hypothesis comes from some recent experiments comparing the statistics of neuronal firing across different situations.^{146,147} It is also compatible with the observed contrast-invariant tuning properties of neurons in the primary visual cortex.¹⁴⁸ There are, furthermore, behavioral experiments with bistable percepts that can be interpreted in that framework.^{142,149} However, no experiments, to our knowledge, have explicitly changed probabilities and measured the resulting neuronal variability.

As a last theory, we want to mention the possibility that uncertainty could be encoded not in the firing properties of neurons but in the connections between them (see Fig. 6F)—for example, in the number and strength of synapses between them.¹⁵⁰ This type of uncertainty coding makes sense in the case of priors, as priors are acquired over long periods of time and thus there is a need to store information in a more durable way. Uncertainty, in this case, would thus change the way that neurons interact with one another.

As we discussed, there is a wide range of exciting theories of how the nervous system could represent probability distributions in the firing rates of the neurons. It is also possible that uncertainty is encoded not in the firing properties of neurons but in the connections between them. We should also not forget that the nervous system has no reason to use a code for uncertainty that is particularly easy for us humans to understand. So far, available experimental data does not in any way strongly support one theory over the others. More importantly, these theories (portrayed in Fig. 6) are not mutually exclusive. The nervous system may use any or all of these mechanisms to encode uncertainty at the same time, and use different types of coding for different types of uncertainty. The goal of future research should be not only to see which, if any, type of uncertainty coding the brain is using in each case, but also to try to figure out what could be the unifying organizing principle that could explain most of the available experimental data. As we saw in this review, there are many different types of uncertainty in the world, and Bayes theory provides a simple organizing prin-

ciple that tells us how we should act in face of any type of uncertainty. Similarly, we should strive to find an organizing principle that could underlie the different types of uncertainty representation in the brain.

Problems and future directions

Here, we have reviewed how Bayes theory can be used to formalize how we should act in face of any type of uncertainty. We have discussed how a range of different models that have been used to model aspects of processing in the nervous system derive from underlying assumptions about the structure of the world. Specifically, algorithms of cue combination assume that there is one causal event (the hidden variable that we want to estimate) that is reflected in multiple sensory cues, while prior-likelihood algorithms generally only assume one sensory cue but take into account prior information about that event. However, in both these algorithms, the variable of interest is assumed to not change with time. Models of the type of Kalman filter and Kalman controller do not have this assumption, and take into account the dynamics of the world for their estimates. In mixture models, on the other hand, besides estimating the hidden variables, the causal structure of the world itself needs to be estimated. In switching dynamic models, even the dynamics of the world are not assumed to be constant, and there is uncertainty about which type of dynamics the world has at any given point of time. Lastly, specific assumptions of how images are generated in the world give rise to Bayesian computer vision algorithms. Finally, we saw how Bayesian decision making can be regarded as an important final step in all the models explained above, incorporating the potential rewards and costs associated with each estimate. All these models derive from distinct assumptions, but they all use the idea that uncertainty should be taken into account and that different pieces of information can be combined optimally using Bayes rule.

A question that arises is how we get the information that allows us to make Bayesian inferences about sensory, motor, and causal events. There are multiple timescales over which information is acquired. Over short periods of time, the nervous system obtains information (sensory stimulation). Over longer timescales, the nervous system learns. And over evolutionary timescales,

the nervous system evolves, which implies acquiring knowledge about the kind of environment in which we live. Bayesian theories of human behavior are fundamentally about the use of information. As such, if a person behaves as should be expected from Bayesian statistics, there is always the question of the timescale over which the relevant information was acquired; i.e., if it was obtained during someone's lifetime (i.e., learned) or if it is innate (i.e., obtained by the nervous system through evolution). Further behavioral and neuroscientific work is necessary to answer this question.

It deserves to be mentioned that Bayesian behavior occasionally comes with disadvantages. Having prior beliefs introduces biases, and these biases may make us less optimal in changing environments—we become prejudiced. The existence of a strong prior can sometimes bias our perception of the world, as we do not see what is really there, making us apparently “less optimal.” Examples are the optical illusions mentioned before (Fig. 2C and D). There are also rumors that some Bayesians tend to have very strong priors that behavior is likely to be related to priors. However, cases in which priors introduce wrong biases appear to be rare, and in the majority of cases Bayesian behavior helps us to more rapidly and more efficiently sense the world.

The question naturally arises of how close people's behavior is to the Bayesian ideal. Performing optimally all the time implies that people always have the right causal model, know precisely the probability distributions associated with each event, and consistently make optimal decisions. A frequently proposed alternative to Bayesian behavior is the idea of heuristics. Instead of behavior being optimal, we use a relatively small set of strategies that just allow us to do good enough.¹⁵¹ It seems unlikely that human behavior is perfectly optimal, and it is quite possible that neurons just somehow approximate optimal behavior. However, Bayesian statistics does explain a good amount of observed behavior, and it also provides a simple coherent theoretical framework that can provide quantitative models and lead to computational insights.

Finally, we want to point out that there is currently some disconnection between Bayesian theories and experimental data about the nervous system. While there are many theoretical proposals of how the ner-

vous system might represent uncertainty, there is not much experimental support for any of them. We hope that future experiments using a wide range of technologies including behavioral, electrophysiology, and imaging studies will shine light on these issues.

Acknowledgments

We want to thank Hugo M. Martins for making all the cartoons that are shown in this paper, and Lisa Morselli for allowing us to use her cat photo. We also want to thank Hugo Fernandes and Mark Albert for helpful comments on the manuscript.

Finally, we wish to thank the International Neuroscience PhD Program, Portugal (sponsored by Fundação Calouste Gulbenkian, Fundação Champalimaud, and Fundação para a Ciência e Tecnologia; SFRH/BD/33272/2007), and the NIH grants R01NS057814, K12GM088020, and 1R01NS063399 for support.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Marr, D. 1982. *Vision: A Computational Approach*. Freeman & Co. San Francisco.
2. Kording, K. 2007. Decision theory: what “should” the nervous system do? *Science* **318**: 606–610.
3. Plato. 1991. *The Republic: The Complete and Unabridged Jowett Translation*. Vintage Books. New York, NY.
4. Smith, M.A. 2001. Alhacen's theory of visual perception. *Transactions of the American Philosophical Society*, 91.4 and 91.5. American Philosophical Society. Philadelphia, PA.
5. Helmholtz, H. 1867. *Treatise on physiological optics*, vol. III. Trans. & ed., J.P.C. Southall. Dover. New York, NY.
6. Bayes, T. 1764. An essay toward solving a problem in the doctrine of chances. *Phil. Trans R Soc London* **53**: 370–418.
7. Frey, B. 1998. *Graphical Models for Machine Learning and Digital Communication*. MIT Press. Cambridge, MA.
8. Jordan, M.I. 1998. *Learning in Graphical Models*. MIT Press. Cambridge, MA.
9. Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press. Cambridge.
10. Richardson, T. & P. Spirtes. 2002. Ancestral graph Markov models. *Ann. Stat.* **30**: 962–1030.
11. Clifford. 1990. Markov random fields in statistics. In *Disorder in Physical Systems*. G.R. Grimmett & D.J.A. Welsh, Eds.: 19–32. Oxford University Press. Oxford, UK.
12. Ernst, M.O. & H.H. Bulthoff. 2004. Merging the senses into a robust percept. *Trends Cogn. Sci.* **8**: 162–169.
13. Kersten, D., P. Mamassian & A. Yuille. 2004. Object perception as Bayesian inference. *Annu. Rev. Psychol.* **55**: 271–304.
14. Knill, D. & W. Richards. 1996. *Perception as Bayesian Inference*. Cambridge University Press. Cambridge.

15. Alais, D., F.N. Newell & P. Mamassian. 2010. Multisensory Processing in Review: from Physiology to Behaviour. *Seeing Perceiving* **23**: 3–38.
16. Jaynes, E.T. 1986. *Bayesian Methods: General Background*. Cambridge University Press. Cambridge.
17. Gelman, A. et al. 2004. *Bayesian Data Analysis*. Chapman & Hall. Boca Raton.
18. Robert, C.P. 2005. *The Bayesian Choice*. Springer. New York.
19. Yuille, A. & H.H. Bulthoff. 1996. Bayesian decision theory and psychophysics. In *Perception as Bayesian Inference*. D. Knill & W. Richards, Eds.: 123–161. Cambridge University Press. Cambridge, UK.
20. Trommershauser, J., L.T. Maloney & M.S. Landy. 2003. Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **20**: 1419–1433.
21. Todorov, E. 2006. Optimal control theory. In *Bayesian Brain*. K. Doya, Ed.: 269–298. MIT Press. Cambridge, MA.
22. Knill, D.C. & A. Pouget. 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**: 712–719.
23. Fisher, G. 1962. Resolution of spatial conflict. *Bull. Br. Psychol. Soc.* **46**: A3.
24. McGurk, H. & J. MacDonald. 1976. Hearing lips and seeing voices. *Nature* **264**: 746–748.
25. Landy, M.S. et al. 1995. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res.* **35**: 389–412.
26. Ghahramani, Z. 1995. *Computational and Psychophysics of Sensorimotor Integration*. Massachusetts Institute of Technology. Cambridge, MA.
27. van Beers, R.J., A.C. Sittig & J.J. Gon. 1999. Integration of proprioceptive and visual position-information: An experimentally supported model. *J. Neurophysiol.* **81**: 1355–1364.
28. van Beers, R.J., A.C. Sittig & J.J.D.v.d. Gon. 1996. How humans combine simultaneous proprioceptive and visual position information. *Exp. Brain Res.* **111**: 253–261.
29. Young, M.J., M.S. Landy & L.T. Maloney. 1993. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vis. Res.* **33**: 2685–2696.
30. Ernst, M.O. & M.S. Banks. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**: 429–433.
31. Landy, M.S. & H. Kojima. 2001. Ideal cue combination for localizing texture-defined edges. *J. Opt. Soc. Am. A.* **18**: 2307–2320.
32. Battaglia, P.W., R.A. Jacobs & R.N. Aslin. 2003. Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A.* **20**: 1391–1397.
33. Alais, D. & D. Burr. 2004. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* **14**: 257–262.
34. Jacobs, R.A. 1999. Optimal integration of texture and motion cues to depth. *Vis. Res.* **39**: 3621–3629.
35. Knill, D.C. & J.A. Saunders. 2003. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vis. Res.* **43**: 2539–2558.
36. Hillis, J.M. et al. 2004. Slant from texture and disparity cues: optimal cue combination. *J. Vis.* **4**: 967–992.
37. Louw, S., J. Smeets & E. Brenner. 2007. Judging surface slant for placing objects: a role for motion parallax. *Exp. Brain Res.* **183**: 149–158.
38. Berniker, M., K. Wei & K.P. Kording. 2010. Bayesian approaches to modeling action selection. In *Modeling Natural Action Selection*. A.K. Seth, Ed.: Cambridge University Press. Cambridge. In press.
39. Wolpert, D.M. 2007. Probabilistic models in human sensorimotor control. *Hum. Mov. Sci.* **26**: 511–524.
40. Ma, W.J. & A. Pouget. 2008. Linking neurons to behavior in multisensory perception: A computational review. *Brain Res.* **1242**: 4–12.
41. Shams, L. & U.R. Beierholm. 2010. Causal inference in perception. *Trends Cogn. Sci.* **14**: 425–432.
42. Wozny, D.R. & L. Shams. 2006. Integration and segregation of visual-tactile-auditory information is Bayes-optimal. *J. Vis.* **6**: 176.
43. Shams, L., Y. Kamitani & S. Shimojo. 2000. Illusions. What you see is what you hear. *Nature* **408**: 788.
44. Shams, L., W.J. Ma & U. Beierholm. 2005. Sound-induced flash illusion as an optimal percept. *Neuroreport* **16**: 1923–1927.
45. Stein, B.E. & T.R. Stanford. 2008. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**: 255–266.
46. Angelaki, D.E., Y. Gu & G.C. DeAngelis. 2009. Multisensory integration: psychophysics, neurophysiology, and computation. *Curr. Opin. Neurobiol.* **19**: 452–458.
47. Meredith, M.A. & B.E. Stein. 1983. Interactions among converging sensory inputs in the superior colliculus. *Science* **221**: 389–391.
48. Barraclough, N.E. et al. 2005. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J. Cogn. Neurosci.* **17**: 377–391.
49. Welchman, A.E. et al. 2005. 3D shape perception from combined depth cues in human visual cortex. *Nat. Neurosci.* **8**: 820–827.
50. Gu, Y., D.E. Angelaki & G.C. DeAngelis. 2008. Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* **11**: 1201–1210.
51. Morgan, M.L., G.C. Deangelis & D.E. Angelaki. 2008. Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* **59**: 662–673.
52. Berniker, M., M. Voss & K. Kording. 2010. Learning priors for Bayesian computations in the nervous system. *PLoS ONE* **5**: e12686.
53. Beierholm, U.R., S.R. Quartz & L. Shams. 2009. Bayesian priors are encoded independently from likelihoods in human multisensory perception. *J. Vis.* **9**: 21–29.
54. Adelson, E.H. 1995. *Checkershadow Illusion*. Retrieved from http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html.
55. Kording, K.P. & D.M. Wolpert. 2004. Bayesian integration in sensorimotor learning. *Nature* **427**: 244–247.
56. Brouwer, A.M. & D.C. Knill. 2009. Humans use visual and remembered information about object location to plan pointing movements. *J. Vis.* **9**: 21–19.

57. Tassinari, H., T.E. Hudson & M.S. Landy. 2006. Combining priors and noisy visual cues in a rapid pointing task. *J. Neurosci.* **26**: 10154–10163.
58. Miyazaki, M., D. Nozaki & Y. Nakajima. 2005. Testing Bayesian models of human coincidence timing. *J. Neurophysiol.* **94**: 395–399.
59. Gerardin, P., Z. Kourtzi & P. Mamassian. 2010. Prior knowledge of illumination for 3D perception in the human brain. *Proc. Natl. Acad. Sci. USA* **107**: 16309–16314.
60. Cisek, P. & J.F. Kalaska. 2005. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* **45**: 801–814.
61. Cisek, P. & J.F. Kalaska. 2002. Simultaneous encoding of multiple potential reach directions in dorsal premotor cortex. *J. Neurophysiol.* **87**: 1149–1154.
62. Riehle, A. & J. Requin. 1989. Monkey primary motor and premotor cortex: single-cell activity related to prior information about direction and extent of an intended movement. *J. Neurophysiol.* **61**: 534–549.
63. Bastian, A., G. Schoner & A. Riehle. 2003. Preshaping and continuous evolution of motor cortical representations during movement preparation. *Eur. J. Neurosci.* **18**: 2047–2058.
64. Bastian, A. *et al.* 1998. Prior information preshapes the population representation of movement direction in motor cortex. *Neuroreport* **9**: 315–319.
65. Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng. (ASME)* **82D**: 35–45.
66. Scheidt, R.A., J.B. Dingwell & F.A. Mussa-Ivaldi. 2001. Learning to move amid uncertainty. *J. Neurophysiol.* **86**: 971–985.
67. Thoroughman, K.A. & R. Shadmehr. 2000. Learning of action through adaptive combination of motor primitives. *Nature* **407**: 742–747.
68. Mussa-Ivaldi, S. & S. Solla. 2008. Models of motor control. In *The Cambridge Handbook of Computational Psychology (Cambridge Handbooks in Psychology)*. R. Sun, Ed.: 635–663. Cambridge University Press. Cambridge.
69. Wolpert, D.M., Z. Ghahramani & M.I. Jordan. 1995. An internal model for sensorimotor integration. *Science* **269**: 1880–1882.
70. Izawa, J. & R. Shadmehr. 2008. On-line processing of uncertain information in visuomotor control. *J. Neurosci.* **28**: 11360–11368.
71. Van Der Kooij, H. *et al.* 2001. An adaptive model of sensory integration in a dynamic environment applied to human stance control. *Biol. Cybern.* **84**: 103–115.
72. Peterka, R.J. & P.J. Loughlin. 2004. Dynamic regulation of sensorimotor integration in human postural control. *J. Neurophysiol.* **91**: 410–423.
73. Kuo, A.D. 2005. An optimal state estimation model of sensory integration in human postural balance. *J. Neural Eng.* **2**: S235–249.
74. Stevenson, I.H. *et al.* 2009. Bayesian integration and non-linear feedback control in a full-body motor task. *PLoS Comput. Biol.* **5**: e1000629.
75. Korenberg, A.T. & Z. Ghahramani. 2002. A Bayesian view of motor adaptation. *Curr. Psychol. Cogn.* **21**: 537–564.
76. Berniker, M. & K. Kording. 2008. Estimating the sources of motor errors for adaptation and generalization. *Nat. Neurosci.* **11**: 1454–1461.
77. van Beers, R.J. 2009. Motor learning is optimally tuned to the properties of motor noise. *Neuron* **63**: 406–417.
78. Burge, J., M.O. Ernst & M.S. Banks. 2008. The statistical determinants of adaptation rate in human reaching. *J. Vis.* **8**: 21–19.
79. Wei, K. & K. Kording. 2010. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front. Comput. Neurosci.* **4**: 1–9.
80. Kording, K.P., J.B. Tenenbaum & R. Shadmehr. 2007. The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat. Neurosci.* **10**: 779–786.
81. Shadmehr, R., M.A. Smith & J.W. Krakauer. 2010. Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* **33**: 89–108.
82. Todorov, E. 2004. Optimality principles in sensorimotor control. *Nat. Neurosci.* **7**: 907–915.
83. Diedrichsen, J., R. Shadmehr & R.B. Ivry. 2010. The coordination of movement: optimal feedback control and beyond. *Trends Cogn. Sci.* **14**: 31–39.
84. Diedrichsen, J. 2007. Optimal task-dependent changes of bimanual feedback control and adaptation. *Curr. Biol.* **17**: 1675–1679.
85. Battaglia, P.W. & P.R. Schrater. 2007. Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *J. Neurosci.* **27**: 6984–6994.
86. Christopoulos, V.N. & P.R. Schrater. 2009. Grasping objects with environmentally induced position uncertainty. *PLoS. Comput. Biol.* **5**: e1000538.
87. Wu, W. *et al.* 2004. Modeling and decoding motor cortical activity using a switching Kalman filter. *IEEE Trans. Biomed. Eng.* **51**: 933–942.
88. Kim, S.P. *et al.* 2008. Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia. *J. Neural Eng.* **5**: 455–476.
89. Mulliken, G.H., S. Musallam & R.A. Andersen. 2008. Decoding trajectories from posterior parietal cortex ensembles. *J. Neurosci.* **28**: 12913–12926.
90. Wu, W. & N.G. Hatsopoulos. 2008. Real-time decoding of nonstationary neural activity in motor cortex. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**: 213–222.
91. Wu, W. *et al.* 2009. Neural decoding of hand motion using a linear state-space model with hidden states. *IEEE Trans. Neural Syst. Rehabil. Eng.* **17**: 370–378.
92. Deneve, S., J.R. Duhamel & A. Pouget. 2007. Optimal sensorimotor integration in recurrent cortical networks: a neural implementation of Kalman filters. *J. Neurosci.* **27**: 5744–5756.
93. Gold, J.I. & M.N. Shadlen. 2002. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* **36**: 299–308.
94. Cheng, P.W. 1997. From covariation to causation: a causal power theory. *Psychol. Rev.* **104**: 367–405.
95. Knill, D.C. 2003. Mixture models and the probabilistic structure of depth cues. *Vis. Res.* **43**: 831–854.
96. Kording, K.P. *et al.* 2007. Causal inference in multisensory perception. *PLoS ONE* **2**: e943.

97. Sato, Y., T. Toyozumi & K. Aihara. 2007. Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli. *Neural Comput.* **19**: 3335–3355.
98. Wei, K. & K. Kording. 2009. Relevance of error: what drives motor adaptation? *J. Neurophysiol.* **101**: 655–664.
99. Beierholm, U. *et al.* 2008. Comparing Bayesian models for multisensory cue combination without mandatory integration. In *Neural Information Processing Systems*, Vol. 20. J. Platt, D. Koller, Y. Singer & S. Roweiss, Eds.: 81–88. MIT Press. Cambridge, MA.
100. Natarajan, R. *et al.* 2009. Characterizing response behavior in multi-sensory perception with conflicting cues. In *Advances in Neural Information Processing Systems*, Vol. 21. D. Koller *et al.*, Eds.: 1153–1180. MIT Press. Cambridge, MA.
101. Stocker, A.A. & E.P. Simoncelli. 2008. A Bayesian model of conditioned perception. *Adv. Neural Inform. Process. Syst.* **20**: 1409–1416.
102. Rowland, B., T. Stanford & B.E. Stein. 2007. A Bayesian model unifies multisensory spatial localization with the physiological properties of the superior colliculus. *Exp. Brain Res.* doi 10.1007/s00221-006-0847-2.
103. Yu, B.M. *et al.* 2007. Mixture of trajectory models for neural decoding of goal-directed movements. *J. Neurophysiol.* **97**: 3763–3780.
104. Corbett, E., E.J. Perreault & K.P. Kording. 2010. Mixture of time-warped trajectory models for movement decoding. In *Neural Information Processing Systems*. J. Lafferty, C.K.I. Williams, J. Shawe-Taylor, R.S. Zemel & A. Culotta, Eds.: 433–441. MIT Press. Cambridge, MA.
105. Kemere, C. *et al.* 2008. Detecting neural-state transitions using hidden Markov models for motor cortical prostheses. *J. Neurophysiol.* **100**: 2441–2452.
106. Yu, A.J. & P. Dayan. 2005. Uncertainty, neuromodulation, and attention. *Neuron* **46**: 681–692.
107. Brubaker, J.R. 2008. wants moar: visual media's use of text in LOLcats and Silent Film. *gnovis J.* **8**: 117–124.
108. Buntin, W.L. 1994. Operations for learning with graphical models. *J. Artif. Intell. Res.* **2**: 159–225.
109. Shi, J.B. & J. Malik. 2000. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal.* **22**: 888–905.
110. Yuille, A. & D. Kersten. 2006. Vision as Bayesian inference: analysis by synthesis? *Trends Cog. Sci.* **10**: 301–308.
111. Zhu, L., Y. Chen & A. Yuille. 2009. Unsupervised learning of probabilistic grammar-Markov models for object categories. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**: 114–128.
112. Sudderth, E. *et al.* 2008. Describing visual scenes using transformed objects and parts. *Int. J. Comput. Vis.* 291–330.
113. Bentham, J. 1780. *An Introduction to the Principles of Morals and Legislation*. Clarendon Press. Oxford.
114. Maloney, L.T., J. Trommershäuser & M.S. Landy. 2006. Questions without words: A comparison between decision making under risk and movement planning under risk. In *Integrated Models of Cognitive Systems*. W. Gray, Ed.: 297–315. Oxford University Press. New York, NY.
115. Kahneman, D. & A. Tversky. 1979. Prospect theory: an analysis of decision under risk. *Econometrica.* **XLVII**: 263–291.
116. Ariely, D. 2008. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. Harper-Collins. New York.
117. Glimcher, P. 2003. *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. MIT Press. Cambridge, MA.
118. Beck, J.M. *et al.* 2008. Probabilistic population codes for Bayesian decision making. *Neuron* **60**: 1142–1152.
119. Fiorillo, C.D., P.N. Tobler & W. Schultz. 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**: 1898–1902.
120. Tobler, P.N., C.D. Fiorillo & W. Schultz. 2005. Adaptive coding of reward value by dopamine neurons. *Science* **307**: 1642–1645.
121. Cromwell, H.C. & W. Schultz. 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* **89**: 2823–2838.
122. Padoa-Schioppa, C. & J.A. Assad. 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**: 223–226.
123. Paton, J.J. *et al.* 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**: 865–870.
124. Gottfried, J.A., J. O'Doherty & R.J. Dolan. 2003. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301**: 1104–1107.
125. Delazer, M. *et al.* 2010. Decision making under ambiguity and under risk in mesial temporal lobe epilepsy. *Neuropsychologia* **48**: 194–200.
126. Preuschoff, K., P. Bossaerts & S.R. Quartz. 2006. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**: 381–390.
127. Tobler, P.N. *et al.* 2007. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* **97**: 1621–1632.
128. Huettel, S.A., A.W. Song & G. McCarthy. 2005. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* **25**: 3304–3311.
129. Singer, T., H.D. Critchley & K. Preuschoff. 2009. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* **13**: 334–340.
130. Rushworth, M.F.S. & T.E.J. Behrens. 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* **11**: 389–397.
131. McCoy, A.N. & M.L. Platt. 2005. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* **8**: 1220–1227.
132. Anderson, C.H. 1994. Basic elements of biological computational systems. *Int. J. Modern Phys. C.* **5**: 313–315.
133. Van Essen, D.C., C.H. Anderson & D.J. Felleman. 1992. Information processing in the primate visual system: an integrated systems perspective. *Science* **255**: 419–423.
134. Barlow, H.B., R. Fitzhugh & S.W. Kuffler. 1957. Change of organization in the receptive fields of the cat's retina during dark adaptation. *J. Physiol.* **137**: 338–354.
135. Ma, W.J. *et al.* 2006. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**: 1432–1438.
136. Carandini, M. & D.J. Heeger. 1994. Summation and division by neurons in primate visual cortex. *Science* **264**: 1333–1336.
137. Shapley, R., E. Kaplan & R. Soodak. 1981. Spatial summation and contrast sensitivity of X and Y cells in the

- lateral geniculate nucleus of the macaque. *Nature* **292**: 543–545.
138. Cheng, K. *et al.* 1994. Comparison of neuronal selectivity for stimulus speed, length, and contrast in the prestriate visual cortical areas V4 and MT of the macaque monkey. *J. Neurophysiol.* **71**: 2269–2280.
 139. Deneve, S. 2008. Bayesian spiking neurons I: inference. *Neural Comput.* **20**: 91–117.
 140. Huan, Y. & R.P. Rao. 2011. Predictive coding. In *Wiley Interdisciplinary Reviews: Cognitive Science*. in press. Available at [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1939-5086](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1939-5086).
 141. Bair, W. & C. Koch. 1996. Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput.* **8**: 1185–1202.
 142. Fiser, J. *et al.* 2010. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* **14**: 119–130.
 143. Hinton, G.E. & T.J. Sejnowski. 1983. Optimal perceptual inference. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC.
 144. Hoyer, P.O. & A. Hyvärinen. 2003. Interpreting neural response variability as Monte Carlo sampling of the posterior. In *Neural Information Processing Systems*. Vol. 15, 277–284. MIT Press. Cambridge, MA.
 145. Berkes, P. *et al.* 2011. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* **331**: 83–87.
 146. Fiser, J., C. Chiu & M. Weliky. 2004. Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature* **431**: 573–578.
 147. Kenet, T. *et al.* 2003. Spontaneously emerging cortical representations of visual attributes. *Nature* **425**: 954–956.
 148. Finn, I.M., N.J. Priebe & D. Ferster. 2007. The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron* **54**: 137–152.
 149. Hoyer, P.O. & A. Hyvärinen. 2003. Interpreting neural response variability as Monte Carlo sampling of the posterior. In *Neural Information Processing Systems*, Vol. 15. S. Becker, S. Thrun & K. Obermayer, Eds.: 293–300. MIT Press. Cambridge, MA.
 150. Wu, S. & S. Amari. 2003. Neural Implementation of Bayesian Inference in Population Codes. *Adv. Neural Inform. Process. Syst.* **2003**: 1–8.
 151. Gigerenzer, G. *et al.* 1999. *Simple Heuristics That Make Us Smart*. Oxford University Press. New York.