Volume 18, issue 2   April 2008
ISSN 0959-4388

ELSEVIER

# Current Opinion in
# Neurobiology

Tobias Bonhoeffer & Marc Tessier-Lavigne, Editors

## April 2008

### Cognitive neuroscience
Edited by Read Montague and John Assad

June 2008 Signalling mechanisms
August 2008 Sensory systems
October 2008 Neuronal and glial cell biology
December 2008 Motor systems • New technologies
February 2009 Development

ScienceDirect   Access *CONB* articles online up to one month before
makes sense.   they appear in your print journal www.sciencedirect.com

CURRENT
OPINION

# Spiking networks for Bayesian inference and choice

Wei Ji Ma[1,3], Jeffrey M Beck[1] and Alexandre Pouget[1,2]

Systems neuroscience traditionally conceptualizes a population of spiking neurons as merely encoding the value of a stimulus. Yet, psychophysics has revealed that people take into account stimulus uncertainty when performing sensory or motor computations and do so in a nearly Bayes-optimal way. This suggests that neural populations do not encode just a single value but an entire probability distribution over the stimulus. Several such probabilistic codes have been proposed, including one that utilizes the structure of neural variability to enable simple neural implementations of probabilistic computations such as optimal cue integration. This approach provides a quantitative link between Bayes-optimal behaviors and specific neural operations. It allows for novel ways to evaluate probabilistic codes and for predictions for physiological population recordings.

**Address**
[1] Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA
[2] Theoretical Neuroscience Group, Collège de France, 75005 Paris, France
[3] Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA

Corresponding author: Pouget, Alexandre (alex@bcs.rochester.edu)

## Introduction

When performing even mundane actions such as changing lanes on the highway, sensory information is manipulated in many ways: Features are assigned to objects, objects are segregated from the background, auditory and visual cues are integrated, current and remembered information about locations of other cars are combined, reward and cost information are taken into account, and task-relevant variables are extracted (e.g. whether changing lanes is safe). Understanding the neural basis of all of these computations constitutes a key problem in neuroscience.

In the past decade, models of neural representation and computation have started to explore the possibility that neurons encode probability distributions and that neural computation is equivalent to probabilistic inference [1,2]. This work was inspired by psychophysical findings showing that human perception [3–11] and motor control [12,13] are nearly optimal in a Bayesian sense. Bayesian optimality means that pieces of uncertain information about task-relevant stimuli are combined with each other or with prior information in a way that takes into account their uncertainty. This requires a neural code that simultaneously represents both the value and the uncertainty of a stimulus and uses this representation in its computations. Since the most general way to represent uncertainty is with a probability distribution, guiding questions in the field have been how spiking neurons might represent probability distributions, and how probabilistic inference can be implemented using biologically plausible neural operations.

## Neural coding of probability distributions

Consider a stimulus $s$ evoking activity in a population of $N$ neurons, which we denote by a vector $\mathbf{r} = (r_1, \ldots, r_N)$. In the standard view of population coding, $\mathbf{r}$ encodes a particular value of the stimulus $s$, and much work has focused on methods for decoding (also called estimating or reading out) this value (e.g. [14–18]). Downstream neural computations are often described in terms of computations on this estimate [19].

However, there is an alternative way to think about population coding and neural computation. If the form of the neural variability is known (that is, if the probability distribution $p(\mathbf{r}|s)$ is known), an entire probability distribution over the stimulus (known as the posterior distribution, $p(s|\mathbf{r})$) can be decoded on each trial, through the use of Bayes' rule (see Box 1) [20,21]:

$$p(s|\mathbf{r}) \propto p(\mathbf{r}|s)\,p(s). \qquad (1)$$

We will call a code that utilizes this fact a probabilistic population code (PPC) [22••]. It is by no means the only neural code that has been suggested for probability distributions ([23–30]; see Box 2 and Figure 1), but it has the advantage that it is consistent with the variability in the nervous system. This is because it is the very presence of variability in $\mathbf{r}$, modeled by $p(\mathbf{r}|s)$, that induces uncertainty about $s$, given by $p(s|\mathbf{r})$, and Bayes' rule describes the exact relationship between these two quantities.

Since a PPC uses the neural variability, that is $p(\mathbf{r}|s)$, to infer the uncertainty about the stimulus, it is essential with such a code to characterize precisely the form of this neural variability. A common assumption is that neural activity can be described as a set of independent Poisson processes. Neurons, however, are not exactly Poisson and

---

**Box 1** Bayes' rule

Bayes' rule is a mathematical identity relating two conditional probabilities to each other. In neural coding theory, it describes how a population of neurons, of which the activity is denoted by a vector **r**, encodes a probability distribution (the posterior) $p(s|\mathbf{r})$ over the stimulus $s$: $p(s|\mathbf{r}) = \frac{p(\mathbf{r}|s)p(s)}{p(\mathbf{r})}$. Bayes' rule requires that the form of neural variability, $p(\mathbf{r}|s)$, is known, either to the experimenter who tries to decode $p(s|\mathbf{r})$ from **r**, or to downstream neurons that use **r** for further computation. Because $p(\mathbf{r})$ is merely a normalization factor that ensures that $p(s|\mathbf{r})$ is a proper probability distribution over $s$, Bayes' rule is often written in the form of Eq. (1). Contrary to widespread belief, the key aspect of Bayes' rule for systems neuroscience is not the prior distribution, $p(s)$, but the fact that a population encodes an entire probability distribution over $s$, and therefore also uncertainty. This is true even if the prior distribution is flat (uniform). It is necessary to consider a population rather than a single neuron (or a scalar 'internal representation', as in signal detection theory [50]), because the set of distributions that can be encoded by a single neuron is very limited. For example, a normal distribution already has two independent parameters, mean and variance, and thus requires at least two neurons to represent it.

---

are correlated. Recently, a more general family of distributions was proposed: the exponential family with linear sufficient statistics, which has been called Poisson-like variability [22••]. This not only includes the Poisson distribution but also can model neural activity for which the Fano factor (the ratio of the variance over the mean) differs from 1 (which is important because the Fano factors of neurons vary from 0.3 to 1.8 [31,32]), as well as stimulus-dependent noise correlations between neurons. This family is parameterized by

$$p(\mathbf{r}|s) = \Phi(\mathbf{r})e^{\mathbf{h}(s)\cdot\mathbf{r}}, \qquad (2)$$

where $\Phi(\mathbf{r})$ is an arbitrary function of **r**, · is the dot product, and the vector-valued function $\mathbf{h}(s)$ is related to the neurons' tuning curves $\mathbf{f}(s)$ and their covariance matrix $\Sigma(s)$ through $\mathbf{h}'(s) = \Sigma^{-1}(s)\mathbf{f}'(s)$. The PPC obtained using Poisson-like variability is related to other forms of prob-abilistic coding (see Box 2). For instance, if elements of the vector $\mathbf{h}(s)$ are Dirac functions, the PPC becomes what is known as a log probability code, a code that has been explored by several investigators [25,27,33,34].

## Bayes-optimal computation with probabilistic codes

The notion that populations encode probability distributions suggests a new perspective on neural computation. Traditionally, neural computation is thought to be a form of nonlinear function approximation [19]. In a probabilistic setting, computations are better formalized as probabilistic inferences. This means that when a function is computed, the output population should encode a probability distribution over function values. Probabilistic inference also allows one to deal with important computations that cannot be cast as function approximations, such as cue combination. Psychophysical experiments have revealed that the nervous system indeed performs probabilistic inference and does so in a nearly optimal way in a variety of tasks [3–13]. The challenge now is to determine how these computations are implemented at the neural level.

The answer to that question depends on the nature of the neural code. For instance, let us consider a task in which human subjects must combine tactile and visual information to determine the width of an object, as in [4]. The optimal solution to this task involves multiplying the probability distribution over the width given by vision, with the one specified by touch. If the brain uses the Poisson-like PPC described above, one can show that this multiplication of distributions can be implemented by linearly combining the patterns of neural activity corresponding to vision and touch [22••], see Figure 2. This operation can be approximated well by a network of spiking integrate-and-fire neurons. This framework

---

**Box 2** A glossary of probabilistic codes

Many schemes for coding probability distributions in populations have been proposed. In explicit probability codes, the mean activity of a neuron is linearly related to the probability that the stimulus is the preferred stimulus of that neuron. In log probability codes, single-trial activity is linearly related to the logarithm of the same probability. Convolution codes consider the activity of a neuron as a vote for a particular function associated with that neuron. Explicit probability codes and PPCs are compared in Figure 1. The following table compares several types of codes.

| Code | Encoding | Decoding | Refs. |
|------|----------|----------|-------|
| Explicit probability code | $\langle r_i \rangle \propto p(s = s_i) + \text{constant}$ | Unclear; requires prior over distributions $p(s)$ | [23,24] |
| Log probability code | $r_i = [a\log p(s = s_i) + b]_+$ | Winner-take-all | [25,27] |
| Log likelihood ratio code | $r_1 = \left[a\log\frac{p(s=s_1)}{p(s=s_2)} + b\right]_+$ | Winner-take-all; limited to binary variables $s$ | [28,30] |
| Convolution code | $r_i \propto \int \varphi_i(s)\,p(s)\,\mathrm{d}s$ | $\hat{p}(s) = \frac{\sum_i r_i \psi_i(s)}{\sum_i r_i}$ | [26,51] |
| Probabilistic population code | Observed variability, $p(\mathbf{r}|s)$ | $p(s|\mathbf{r}) \propto p(\mathbf{r}|s)\,p(s)$ | [20,21] |

Here, $[\cdot]_+$ denotes rectification, that is, nonzero values of the argument are set to zero. A PPC with Poisson-like variability, Eq. (2), is related to other probabilistic codes. Applying Bayes' rule and assuming a flat prior, we find $\log p(s|\mathbf{r}) = \mathbf{h}(s)\cdot\mathbf{r} + \tilde{\Phi}(\mathbf{r})$, where $\tilde{\Phi}(\mathbf{r})$ is such that $p(s|\mathbf{r})$ integrates to 1. This PPC is therefore similar to a convolution code, except for the logarithm. The kernel $\mathbf{h}(s)$ is specified by the statistics of the population and can be estimated through logistic regression. The log probability code is a special case, for which the components of the kernel are delta functions, $h_i(s) = \delta(s - s_i)$. This kernel would, however, not be compatible with tuning curves and covariance structures found in real neurons.

---

Figure 1
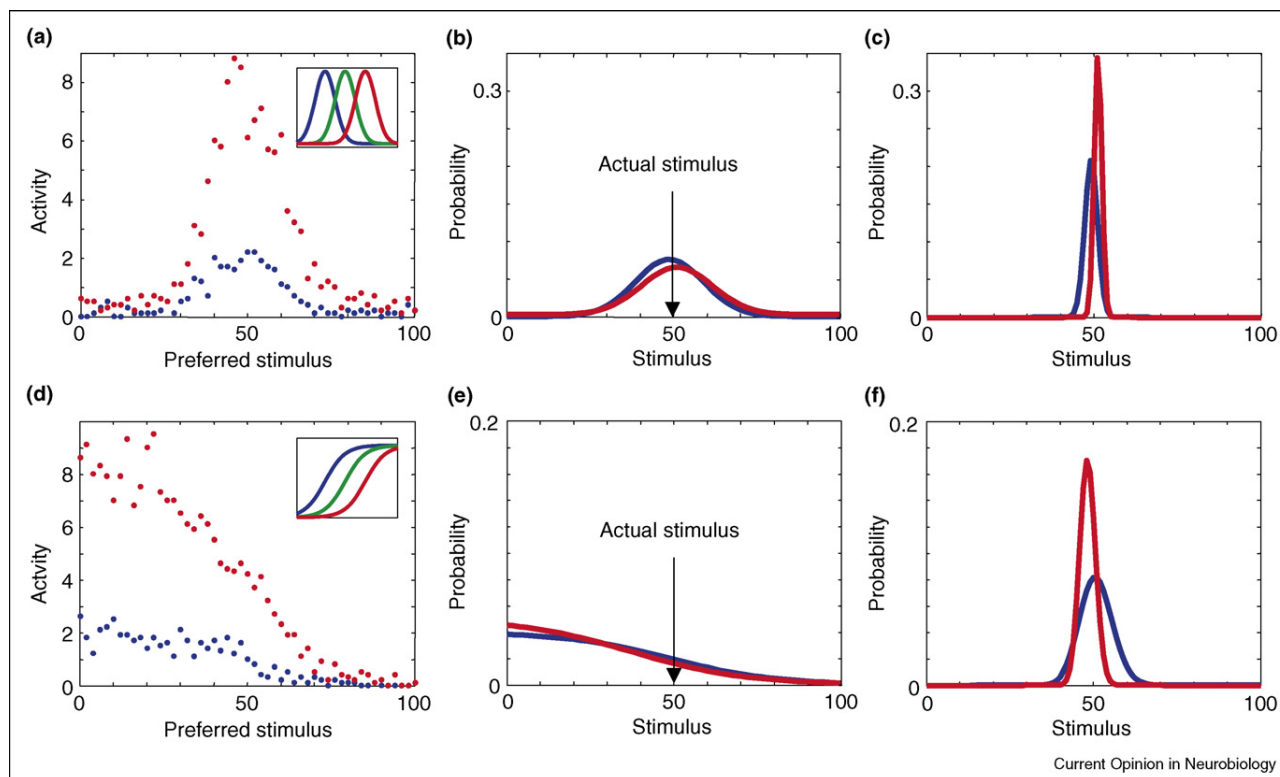


Comparison of explicit probability coding (EPC) and probabilistic population coding (PPC) for populations with bell-shaped (**a–c**) or monotonic (**d–f**) tuning curves and independent Poisson noise. (**a, d**) Population patterns of activity (averaged over 10 trials for clarity) when the actual stimulus had value 50 (arbitrary units). Example tuning curves are shown in the insets, and neurons are ordered by their tuning curves (in **d**, the 'preferred stimulus' is the stimulus for which the tuning curve has the highest slope). The same population is shown for low (blue) and high (red) gain. (**b, c, e, f**) Distributions over the stimulus encoded in the populations *on a single trial*, using EPC (**b, e**) or PPC (**c, f**), when gain is low (blue) or high (red). All distributions will vary somewhat from trial to trial. In PPC, higher neural gain yields a sharper distribution, thus higher certainty (see **c, f**). In EPC, gain has no effect on certainty (see **b, e**) and the distributions do not peak close to the actual value of the stimulus when tuning curves are monotonic (compare **e, f**).

predicts that the response of a multisensory population to a multisensory stimulus is equal to the sum of its responses to each of the corresponding unisensory stimuli (plus or minus a constant, because, in a PPC, constant terms have no effect on the posterior encoded in a population [35]). Recent experiments in multisensory integration have revealed that the majority of multisensory neurons do indeed exhibit this form of additivity [36]. If this finding holds in awake animals performing optimal cue integration, this would provide strong support for a Poisson-like PPC.

Cue integration is only one of many probabilistic computations that the brain has to perform. Others include integrating information over time in perceptual decision making [37], deciding whether multisensory stimuli come from the same physical object [38], running a Kalman filter in sensorimotor control [39*], and combining information from many locations in visual search. Recent studies have even considered situations in which the stimulus varies on a time scale comparable to the time

scale of interspike intervals. In this case, the timing of individual spikes becomes crucial [28,40*]. (Even for constant stimuli, ignoring individual spike timing can incur information loss [41].) More complex neural network schemes have also been designed for (hierarchical) Bayesian inference [27,29,42,52,53], but these were not meant to describe actual neurons, or there is little experimental data to test their validity.

## Binary choice
As we have just pointed out, Bayesian inference can be applied to a wide variety of problems, and this is particularly true for decision making. When deciding between two (or more) alternatives, the relevant quantity is the probability that each of the alternatives is the correct decision given the available evidence. Studies with awake monkeys have revealed that the responses of LIP neurons might indeed encode this probability [30]. In the most recent of these, monkeys were sequentially presented with four shapes, each of which provided probabilistic evidence that choosing one of two targets would lead to

**Figure 2**
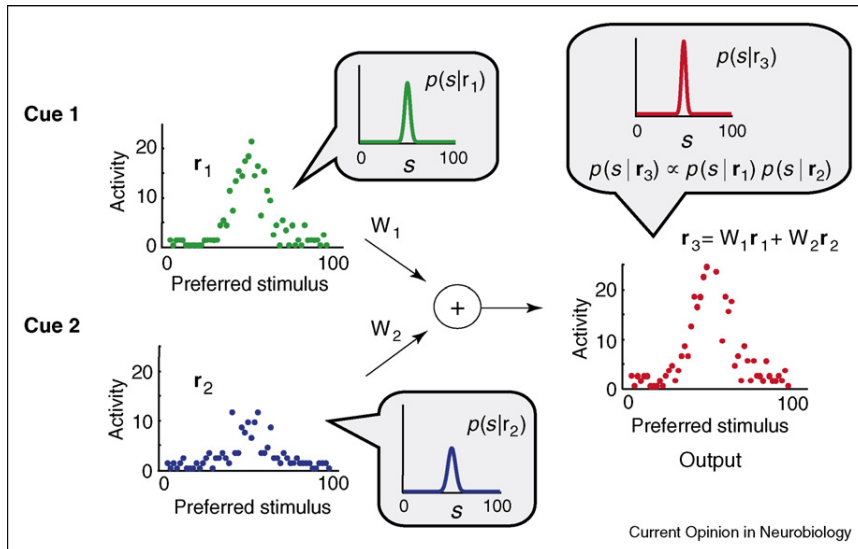


Optimal cue integration with probabilistic population codes [22••]. The cues elicit activity in input populations $r_1$ and $r_2$, indicated by green and blue dots (neurons are ordered by their preferred stimulus). A simple linear combination of the population patterns of activity, $r_3 = W_1r_1 + W_2r_2$, guarantees optimal cue integration, if neural variability is Poisson-like. The dialog boxes show the probability distributions over the stimulus encoded in each population on a single trial. Optimal cue integration is characterized by a multiplication of probability distributions over the stimulus, $p(s|r_3) \propto p(s|r_1)p(s|r_2)$. The synaptic weight matrices $W_1$ and $W_2$ depend on the statistics of the input populations, but do not have to be adjusted over trials.

reward [43••]. The weight of the evidence and the favored target were different for different shapes. It was found that monkeys assigned subjective weights to the shapes that were close to the shapes' true weights. Moreover, firing rates of LIP neurons varied linearly with the log likelihood ratio in each epoch of a trial.

At first sight, these experiments suggest that neurons respond in proportion to log likelihood ratios, or log probabilities (see Box 2). There are, however, a variety of arguments suggesting that this is unlikely to be the case [22••,44]; for instance, it is not always possible to recover the log odds of a decision solely from LIP activity under

**Figure 3**
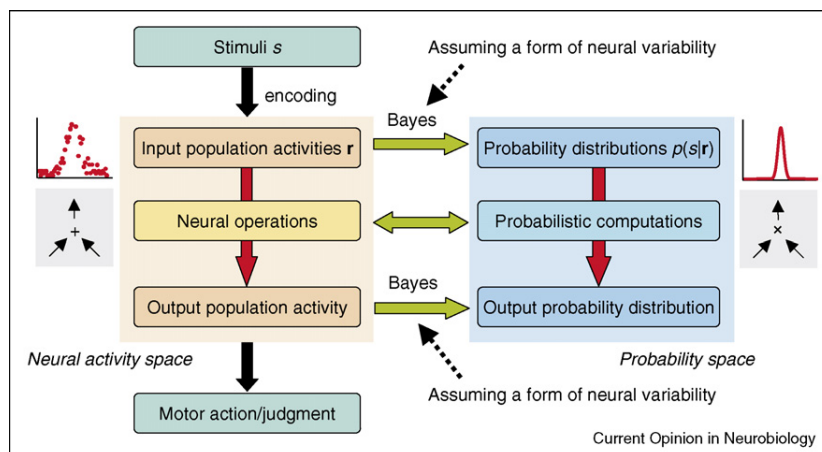


Schematic of probabilistic population coding for perceptual computation. One or multiple stimuli elicit population patterns of activity. Each pattern encodes a probability distribution over the stimulus through Bayes' rule. In perceptual tasks like cue combination, decision making, or visual search, these probability distributions have to be manipulated in specific ways to achieve optimality. Now the key problem is to establish a 'dictionary' between such probabilistic computations and neural operations on population patterns of activity, assuming a form of neural variability. Using those neural operations, the brain will retain full probabilistic information about the variable(s) of interest at all intermediate stages of computation. Eventually, a motor action is generated or a high-level judgment is made (for example, about target presence in a visual search task).

the assumption of a log probability code [44]. Moreover, the data available so far are compatible with the PPC approach we have just described [37]. Therefore, it is indeed quite likely that LIP, and other cortical areas, encode probability distributions for decisions, but the coding scheme may not be as simple as a log probability code.

## Conclusions

We have outlined recent work that shows that populations of neurons can represent probability distributions (instead of only single values) and perform Bayes-optimal computations on them. Such computations, including cue combination and decision making, provide important quantitative tests for probabilistic codes (see Figure 3). A good code should allow each computation to be implemented using biologically realistic neural operations and make physiological predictions for Bayes-optimal behaviors.

Although studies on the neural basis of Bayesian inference have focused on purely sensory processing, further work should investigate generalizations to other domains, such as motor control [12,45] and decision making with a utility function [46]. Sensory and reward-based decisions are often studied separately [47], but a complete theory of Bayes-optimal neural computation should account for tasks with both sensory uncertainty and nontrivial reward schedules. Bayes-optimal behavior can also occur in high-level cognitive tasks, such as generalizing from sparse data [48••,49]. While little is known about the physiological substrates of such behaviors, it is likely that they are already amenable to theoretical approaches.

## References and recommended reading

Papers of particular interest, published over the period of the review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Pouget A, Dayan P, Zemel RS: **Inference and computation with population codes**. *Annu Rev Neurosci* 2003, **26**:381-410.

2. Knill DC, Pouget A: **The Bayesian brain: the role of uncertainty in neural coding and computation**. *Trends Neurosci* 2004, **27**:712-719.

3. Jacobs RA: **Optimal integration of texture and motion cues to depth**. *Vision Res* 1999, **39**:3621-3629.

4. Ernst MO, Banks MS: **Humans integrate visual and haptic information in a statistically optimal fashion**. *Nature* 2002, **415**:429-433.

5. Alais D, Burr D: **The ventriloquist effect results from near-optimal bimodal integration**. *Curr Biol* 2004, **14**:257-262.

6. van Beers RJ, Sittig AC, Denier van der Gon JJ: **How humans combine simultaneous proprioceptive and visual position information**. *Exp Brain Res* 1996, **111**:253-261.

7. Battaglia PW, Jacobs RA, Aslin RN: **Bayesian integration of visual and auditory signals for spatial localization**. *J Opt Soc Am A Opt Image Sci Vis* 2003, **20**:1391-1397.

8. Knill DC, Saunders JA: **Do humans optimally integrate stereo and texture information for judgments of surface slant?** *Vision Res* 2003, **43**:2539-2558.

9. Kersten D, Mamassian P, Yuille A: **Object perception as Bayesian inference**. *Annu Rev Psychol* 2004, **55**:271-304.

10. Weiss Y, Simoncelli EP, Adelson EH: **Motion illusions as optimal percepts**. *Nat Neurosci* 2002, **5**:598-604.

11. Mamassian P: **Bayesian inference of form and shape**. *Progr Brain Res* 2006, **154**:265-270.

12. Wolpert D, Ghahramani Z, Jordan M: **An internal model for sensorimotor integration**. *Science* 1995, **269**:1880-1882.

13. Kording KP, Wolpert DM: **Bayesian integration in sensorimotor learning**. *Nature* 2004, **427**:244-247.

14. Seung H, Sompolinsky H: **Simple model for reading neuronal population codes**. *Proc Natl Acad Sci U S A* 1993, **90**:10749-10753.

15. Wu S, Amari S, Nakahara H: **Population coding and decoding in a neural field: a computational study**. *Neural Comput* 2002, **14**:999-1026.

16. Georgopoulos A, Kalaska J, Caminiti R: **On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex**. *J Neurosci* 1982, **2**:1527-1537.

17. Deneve S, Latham P, Pouget A: **Reading population codes: a neural implementation of ideal observers**. *Nat Neurosci* 1999, **2**:740-745.

18. Paradiso M: **A theory of the use of visual orientation information which exploits the columnar structure of striate cortex**. *Biol Cybern* 1988, **58**:35-49.

19. Poggio T: **A theory of how the brain might work**. *Cold Spring Harbor Sympos Quant Biol* 1990, **55**:899-910.

20. Foldiak P: **The 'ideal homunculus': statistical inference from neural population responses**. In *Computation and Neural Systems.* Edited by Eeckman F, Bower J. Kluwer Academic Publishers; 1993:55-60.

21. Sanger T: **Probability density estimation for the interpretation of neural population codes**. *J Neurophysiol* 1996, **76**:2790-2793.

22. Ma WJ, Beck JM, Latham PE, Pouget A: **Bayesian inference**
•• **with probabilistic population codes**. *Nat Neurosci* 2006, **9**:1432-1438.
A probabilistic interpretation of neural activity and neural operations. It was found that Bayes-optimal cue integration could be implemented using simple linear operations on population activities, given that neural variability is Poisson-like. This lends new significance to the format of neural variability (sometimes confusingly called noise).

23. Hoyer PO, Hyvarinen A: **Interpreting neural response variability as Monte Carlo sampling of the posterior**. *Neural Information Processing Systems*. MIT Press; 2003.

24. Anastasio TJ, Patton PE, Belkacem-Boussaid K: **Using Bayes' rule to model multisensory enhancement in the superior colliculus**. *Neural Comput* 2000, **12**:1165-1187.

25. Barlow HB: **Pattern recognition and the responses of sensory neurons**. *Ann NY Acad Sci* 1969, **156**:872-881.

26. Zemel R, Dayan P, Pouget A: **Probabilistic interpretation of population code**. *Neural Computat* 1998, **10**:403-430.

27. Rao RP: **Bayesian computation in recurrent neural circuits**. *Neural Comput* 2004, **16**:1-38.

28. Deneve S: **Bayesian spiking neurons I: inference**. *Neural Comput* 2008, **20**:91-117.

29. Koechlin E, Anton JL, Burnod Y: **Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area MT**. *Biol Cybern* 1999, **80**:25-44.

30. Gold JI, Shadlen MN: **Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward**. *Neuron* 2002, **36**:299-308.

31. Tolhurst D, Movshon J, Dean A: **The statistical reliability of signals in single neurons in cat and monkey visual cortex**. *Vision Res* 1982, **23**:775-785.

32. Gur M, Snodderly DM: **High response reliability of neurons in primary visual cortex (V1) of alert, trained monkeys**. *Cereb Cortex* 2006, **16(6)**:888-895.

33. Beck JM, Pouget A: **Exact inferences in a neural implementation of a hidden Markov model**. *Neural Comput* 2007, **19**:1344-1361.

34. Jazayeri M, Movshon JA: **Optimal representation of sensory information by neural populations**. *Nat Neurosci* 2006, **9**:690-696.

35. Beck J, Ma WJ, Latham P, Pouget A: **Probabilistic population codes and the exponential family of distributions**. *Prog Brain Res* 2007, **165**:509-519.

36. Stanford TR, Quessy S, Stein BE: **Evaluating the operations underlying multisensory integration in the cat superior colliculus**. *J Neurosci* 2005, **25**:6499-6508.

37. Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman JD, Shadlen MN, Latham PE, Pouget A, Probabilistic population codes for Bayesian decision making, Neuron 2008, in press.

38. Kording KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L: **Causal inference in multisensory perception**. *PLoS ONE* 2007, **2(9)**:e943.

39. Deneve S, Duhamel JR, Pouget A: **Optimal sensorimotor**
•   **integration in recurrent cortical networks: a neural implementation of Kalman filters**. *J Neurosci* 2007, **27**:5744-5756.
A neural model for inference on time-varying stimuli, such as the position of a moving arm. The problem with this model is that it infers only the mean of the posterior distribution, not the full distribution. This problem can be corrected by using the scheme proposed in [22••].

40. Huys Q, Zemel RS, Natarajan R, Dayan P: **Fast population**
•   **coding**. *Neural Comput* 2007, **19**:404-441.
A probabilistic interpretation of neural variability similar to the one proposed in [22••], but applied to stimuli varying on very fast time scale.

41. Toyoizumi T, Aihara K, Amari S: **Fisher information for spike-based population decoding**. *Phys Rev Lett* 2006, **97**:e098102.

42. Sahani M, Dayan P: **Doubly distributional population codes: simultaneous representation of uncertainty and multiplicity**. *Neural Comput* 2003, **15**:2255-2279.

43. Yang T, Shadlen MN: **Probabilistic reasoning by neurons**.
••  *Nature* 2007, **447**:1075-1080.
One of the experiments showing that monkeys can perform Bayesian inference and that the responses of LIP neurons encode probability distributions. Although the data seem to suggest a very specific neural code, namely that LIP neurons respond in proportion to log probability, this is by no means the only interpretation. Other schemes, such as probabilistic population codes, are possible. We emphasize this point because the notion that neurons respond in proportion to log probability is hard to reconcile with other experimental data.

44. Gold JI, Shadlen MN: **Neural computations that underlie decisions about sensory stimuli**. *Trends Cogn Sci* 2001, **5**:10-16.

45. Kording K: **Decision theory: what 'should' the nervous system do?** *Science* 2007, **318**:606-610.

46. Lee D: **Neural basis of quasi-rational decision-making**. *Curr Opin Neurobiol* 2006, **16**:191-198.

47. Sugrue LP, Corrado GS, Newsome WT: **Choosing the greater of two goods: neural currencies for valuation and decision making**. *Nat Rev Neurosci* 2005, **6**:363-375.

48. Tenenbaum JB, Griffiths TL, Kemp C: **Theory-based Bayesian**
••  **models of inductive learning and reasoning**. *Trends Cogn Sci* 2006, **10**:309-318.
A generalization of the Bayesian approach to high-level cognitive learning and reasoning. There is currently no theory of how this type of inference might be implemented at the neural level.

49. Chater N, Tenenbaum JB, Yuille A: **Probabilistic models of cognition: conceptual foundations**. *Trends Cogn Sci* 2006, **10**:287-291.

50. Green DM, Swets JA: **Signal Detection Theory and Psychophysics**. Los Altos, CA: John Wiley & Sons; 1966.

51. Anderson C: **Neurobiological computational systems**. *Computational Intelligence Imitating Life*. IEEE Press; 1994:. pp. 213–222.

52. Kononenko: **Bayesian Neural Networks**. *Biol. Cybernetics* 1989, **61**:361-370.

53. Lansner, Holst: **A higher order Bayesian neural network with spiking units**. *International J. Neural Systems* 1996, **7(2)**:115-128.