itself show that outputs from the frontal cortex are responsible for the task-dependent modulation of A1 neurons, but it does suggest that top-down attentional control can be directed in a frequency-specific fashion that reflects the tonotopic organization of A1.

Particular mention should be made of the animal model used in this study. Although previous work on the influence of prefrontal cortex on processing in sensory cortical areas has been dominated by research on primates[5,6], ferrets provide another option for studying the effects of attention and learning on hearing. They can be trained on a variety of tasks and are now widely used in auditory neuroscience research. Nevertheless, it is not a straightforward matter to show that activity in frontal cortex reflects target recognition as opposed to motor planning in the behavioral procedure, which was chosen by Fritz and colleagues[3] for its ease of use with ferrets. By correlating activity with task-related licking, the authors attempted to isolate any contributions resulting from motor behavior. However, there is no doubt that the use of an approach that dissociates sensory and motor effects, such as a delayed matching-to-sample procedure[7], would be desirable, although it remains to be seen whether ferrets can be trained to perform such a task.

Another issue concerns our primitive understanding of frontal cortex in nonprimate species. Fritz and colleagues[3] propose that the region they recorded from is likely to be homologous to primate prefrontal cortex and this is supported by recent anatomical data[8]. Several questions naturally follow from this. Foremost among these is whether the domain specificity hypothesis put forward previously[9] to describe the functional organization of prefrontal cortex in both human and nonhuman primates also applies to ferrets. Do equivalent regions exist that are devoted to spatial working memory and to the processing of object features or object-related working memory[10]? Fritz and colleagues[3] found that neurons responding during auditory and/or visual tasks were intermixed in ferret frontal cortex, suggesting a lack of modality specificity. Given how widespread multisensory convergence is, even in ferret auditory cortex[11], this is not surprising, but further work is obviously needed to characterize the inputs to different parts of the frontal cortex in this species.

The frontal cortex is not the only forebrain region thought to be involved in adjusting the cortical processing of sensory signals. Cholinergic inputs from the basal forebrain also seem to mediate attention and learning by modulating the responsiveness and tuning properties of cortical neurons in ways that enhance the representation of behaviorally relevant stimuli[12]. Recent evidence points to the existence of a sensory cortex–prefrontal cortex–basal forebrain loop[13] comprising neural components capable of integrating bottom-up sensory input with top-down attentional processing and inputs from the limbic system. Such a loop has the potential to respond to real-time changes in task demands by way of the prefrontal cortex and to induce changes in neuronal firing in auditory cortex by the action of the cholinergic inputs from the basal forebrain.

These circuits are also likely to provide the mechanism by which top-down inputs dictate how the response properties of A1 change in a task-specific fashion during perceptual learning[14]. In turn, their influences may be transmitted by way of the auditory cortex to lower parts of the auditory pathway, as suggested by the loss of learning-induced behavioral plasticity after the selective elimination of layer V pyramidal neurons that project to the inferior colliculus[15]. The study by Fritz and colleagues[3] implies that prefrontal cortical neurons might create a flexible framework for focusing attention on behaviorally relevant sounds and provides an intriguing insight into the brain circuitry that allows the dynamic processing of sensory information.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Dahmen, J.C., Keating, P., Nodal, F.R., Schulz, A.L. & King, A.J. *Neuron* **66**, 937–948 (2010).
2. Fritz, J.B., Elhilali, M., David, S.V. & Shamma, S.A. *Hear. Res.* **229**, 186–203 (2007).
3. Fritz, J.B., David, S.V., Radtke-Schuller, S., Yin, P. & Shamma, S.A. *Nat. Neurosci.* **13**, 1011–1019 (2010).
4. Atiani, S., Elhilali, M., David, S.V., Fritz, J.B. & Shamma, S.A. *Neuron* **61**, 467–480 (2009).
5. Miller, E.K. & Cohen, J.D. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
6. Gregoriou, G.G., Gotts, S.J., Zhou, H. & Desimone, R. *Prog. Brain Res.* **176**, 35–45 (2009).
7. Rainer, G., Asaad, W.F. & Miller, E.K. *Nature* **393**, 577–579 (1998).
8. Duque, A. & McCormick, D.A. *Cereb. Cortex* **20**, 1020–1036 (2010).
9. Goldman-Rakic, P.S. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **351**, 1445–1453 (1996).
10. Romanski, L.M. *Cogn. Affect. Behav. Neurosci.* **4**, 421–429 (2004).
11. Bizley, J.K., Nodal, F.R., Bajo, V.M., Nelken, I. & King, A.J. *Cereb. Cortex* **17**, 2172–2189 (2007).
12. Sarter, M., Hasselmo, M.E., Bruno, J.P. & Givens, B. *Brain Res. Brain Res. Rev.* **48**, 98–111 (2005).
13. Rasmusson, D.D., Smith, S.A. & Semba, K. *Neuroscience* **149**, 232–241 (2007).
14. Polley, D.B., Steinberg, E.E. & Merzenich, M.M. *J. Neurosci.* **26**, 4970–4982 (2006).
15. Bajo, V.M., Nodal, F.R., Moore, D.R. & King, A.J. *Nat. Neurosci.* **13**, 253–260 (2010).

It's that time again

Pascal Mamassian & Michael S Landy

How do we estimate the duration of a temporal interval in a familiar context? A new study finds that it is appropriate, perhaps even advantageous, to tolerate a small bias in our estimate to reduce the overall temporal uncertainty.

Your feet are in the starting blocks and the race official calls, "Ready... Set... Go!" To win the race, you must push off on time. If you start too late, you may not be able to make up the lost time and catch up with your competitors. If you start too early, you may be disqualified for a false start. You cannot wait for the starting pistol to plan your start, as that will certainly lead to a slow start. Rather, you must listen for "Ready" and "Set" and predict the correct starting time on the basis of that time interval and prior knowledge of the rhythms typical of race calls. Using a laboratory model, Jazayeri and Shadlen[1] studied this task to understand how humans combine uncertain sensory information and prior knowledge in the estimation of duration.

We speak of the perception of time and temporal duration even though, unlike for primary sensations (for example, loudness, pitch, luminance or pressure), there is no specific sensory organ that encodes duration *per se* and one can in fact estimate durations both in and across sensory modalities. However, the perception of duration shares several traits with other sensory capabilities. Estimation of temporal duration follows Weber's law (also called the scalar variability property[2]): uncertainty scales with duration. Perceived duration is often biased[3]. For example, when humans are asked to reproduce

Pascal Mamassian is at the Laboratoire Psychologie de la Perception (CNRS UMR 8158), Université Paris Descartes, Paris, France, and Michael S. Landy is in the Department of Psychology & Center for Neural Science, New York University, New York, New York, USA.
e-mail: pascal.mamassian@parisdescartes.fr

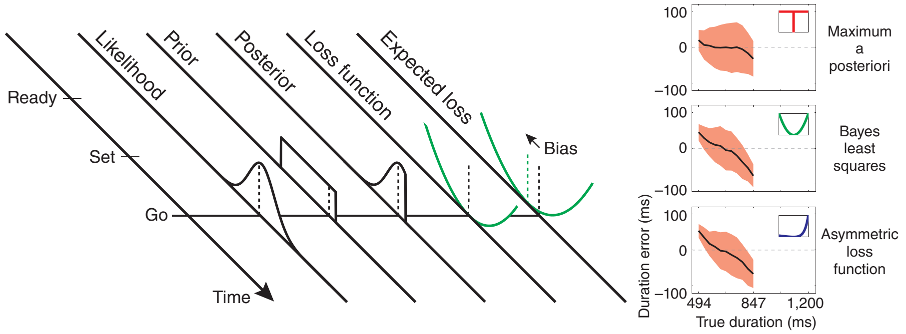


Figure 1 Bayesian models of duration estimation. Left, the participant must anticipate the Go signal after seeing the Ready and Set preparation signals. Knowing that the variable interval between Ready and Set is identical to the interval between Set and Go, what is the best strategy to predict the occurrence of the Go signal? A Bayesian decision maker answers this question by combining several pieces of information. The first piece of information is the likelihood function, which represents the probability of making the measurement the participant has just made of the present duration from Ready to Set given various possible true durations. The second is the prior probability distribution, which represents the accumulated knowledge of interval durations over past races. The product of the likelihood function and the prior distribution determines the posterior distribution: the probability of various possible estimates of the interval duration given the current measurement. The final piece of information is the loss function, which represents the costs associated with correct and incorrect estimates. Combining the posterior distribution with the loss function gives the expected loss: the anticipated cost associated with different duration estimations. The minimum of the expected loss (dashed green line) corresponds to the optimal Bayesian decision. It is negatively biased relative to the correct duration as a result of the prior (most previous race start cadences had a shorter duration). Right, three Bayesian models of duration estimation for true durations chosen from the lowest range of durations used in the Jazayeri and Shadlen study[1]. Each plot shows mean (±1 s.d.) estimates for 1,000 simulated races. For maximum a posteriori estimation, the loss function (inset) penalizes all errors equally. This model is accurate (weakly biased) but not very precise (large variability of the estimates). For Bayes least-squares estimation, the loss function is quadratic. This model has smaller variability, but larger bias, especially for long durations as a result of the increased uncertainty of the likelihood function for longer durations (the scalar variability property). For the asymmetric loss function model, late starts are penalized more than early starts. The accuracy and precision of this model are similar to those of the Bayes least-squares model. The larger biases for longer durations are now the result of the high cost for overestimates, even if the participant's internal model of likelihood is incorrect (no scalar variability).

various temporal intervals, longer durations are perceived as being shorter than the reference and the opposite is true for short durations. This phenomenon of regression to the mean, known as Vierordt's law[4], is one of the most robust effects in time perception, but has not been convincingly explained. In this issue, Jazayeri and Shadlen[1] find that the bias to underestimate long intervals and to overestimate short ones is a consequence of becoming familiar with the range of event durations in a particular context. Moreover, this context-based bias improves performance in the sense of optimizing the tradeoff between bias and variability.

In Jazayeri and Shadlen's experiment[1], human participants saw two flashes of light in succession (Ready and Set) and were required to press a key (Go) so that the temporal interval between the first and second flash was equal to the interval from the second flash to the keypress[5]. In each block of experimental trials, the displayed durations were chosen randomly and uniformly across a fixed range; for example, this 'temporal context' might range from approximately 500 to 850 ms. The authors found that participants produced durations that were biased toward the mean of the temporal context and that this bias increased with increases in the mean of the context.

Bayesian models, now prevalent in models of perceptual-motor behavior (for example, ref. 6), offer a natural setting to explain this temporal bias. Here, we only discuss the perceptual side of the problem. Bayesian models are based on three elements: the likelihood function, the prior probability distribution (usually called the prior) and the loss function (**Fig. 1**). The likelihood function represents the observer's knowledge of measurement uncertainty—that is, the probability of obtaining the current duration measurement given possible durations that might have actually occurred. The prior is the observer's knowledge of the statistics of the world; for the Jazayeri and Shadlen experiment[1], the prior represents knowledge of the temporal context. Finally, the loss function represents the cost to the observer for making a particular estimate; in this case, the cost of reproducing an estimated duration that differs slightly from the true duration.

The loss function for perceptual estimation can be thought of as the compromise between accuracy (the amount of bias) and precision (the variability of the estimate). Consider two extreme strategies. First, participants could ignore the current stimulus and instead estimate the duration to be the mean of recently experienced durations. This first strategy would result in excellent precision (no variability), but clearly this is achieved at the expense of increased bias. Alternatively, participants could ignore the prior and simply reproduce the current duration measurement; this corresponds approximately to maximum-likelihood estimation. This second strategy would result in maximal accuracy, but lower precision, as sensory measurements are uncertain and the participant is ignoring a potential source of information (the prior). The best trade-off between accuracy and precision will depend on the magnitude of the uncertainty. In a Bayesian model, the greater the uncertainty, the more the estimate regresses to the mean of the prior. The amount of bias will depend on the choice of the loss function. Small biases are obtained when the loss function rewards only correct estimates: maximum a posteriori estimation (**Fig. 1**). In contrast, larger biases, but smaller variability, are obtained when a quadratic loss function is used: Bayes least-squares estimation. This latter loss function may penalize large errors more than human participants do[7]. One should also consider asymmetric loss functions[8]. Asymmetric loss functions are particularly critical in the time domain[5,9]; consider the different costs of under- and over-estimation of the time to arrive at an appointment or the time until you should have your car brakes checked.

To what extent did the participants in Jazayeri and Shadlen's experiment[1] have access to the appropriate prior, their temporal uncertainty and a rational loss function? The authors showed that participants had knowledge of the prior because timing bias differed across the three temporal contexts. In particular, the same cued duration was reproduced differently in different contexts, always biased toward the mean of the current context. The authors also confirmed that participants had some knowledge of their temporal uncertainty[5,9] (at least implicitly). In particular, they argue that participants knew about the scalar variability that rules their temporal uncertainty; variability was larger for larger mean durations, so a greater bias was obtained. Unfortunately, temporal uncertainty was not measured independently so it is not clear whether participants have perfect knowledge of their own uncertainty. Finally, Jazayeri and Shadlen[1] found that participants' behavior was consistent with a quadratic loss function so that the optimal estimate is the mean of the posterior distribution. However, other combinations of assumed priors, temporal uncertainty and loss functions might have been consistent with their results (**Fig. 1**).

For the Bayesian modeler, there are three unknown functions: the likelihood, prior and loss function. This multiplicity of unknowns is particularly vexing, as the data only tell us the participant's average response to any given stimulus. This stimulus-response function might result from more than one Bayesian model. For example, larger bias for long durations results from a likelihood implementing scalar variability combined with a quadratic loss function, but similar results are obtained if the participant assumes variability is constant and uses an asymmetrical loss function with high cost for overestimation (**Fig. 1**). Previous studies have used a variety of approaches to reduce the number of unknowns. For sensory experiments, the likelihood function can be measured by determining the observer's ability to discriminate similar stimuli (for example, ref. 10). The prior distribution can be imposed, as it is in Jazayeri and Shadlen's study[1], with training sessions so that participants have an opportunity to learn the prior. It can be measured from the environment and one can ask whether perceptual biases are consistent with the participant computing a Bayesian estimate using this natural prior (for example, ref. 11). Alternatively, experiments can be designed to estimate the shape of the prior used by the observer[12]. The loss function can be imposed by the experimenter[13]. Finally, sensory information can be removed so that performance can only be based on the prior and loss function.

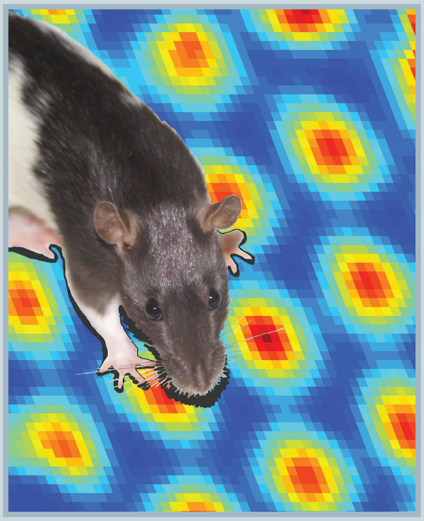
How can the experimenter be assured the participant is truly carrying out the Bayesian computation, rather than a simpler heuristic that has the same net effect? In Jazayeri and Shadlen's experiment[1], participants received feedback for responses that were sufficiently close to the correct duration. Thus, it is reasonable to ask whether the resulting biases were learned by effectively computing a linear regression of produced intervals that led to positive feedback as a function of the corresponding measured intervals. A true Bayesian computation would imply that the participant has knowledge of the likelihood, prior and loss function, and can use those elements when one of them, such as the context (prior), is changed[14]. One strength of the Jazayeri and Shadlen study[1] is that observers were naturally exposed, in different sessions, to three different temporal contexts (ranges of displayed durations) and these priors were learned effortlessly. Other priors can be updated given sufficiently convincing feedback, such as the default prior knowledge that light comes from above our head[15]. It is also surprisingly easy for participants to form nearly optimal strategies in pointing tasks to a visual target with arbitrary, experiment-imposed payoffs and penalties[13]. The ability of human participants to readily adapt to changes in context or loss function constitutes strong evidence for Bayesian decision theories of perceptual-motor performance[14].

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Jazayeri, M. & Shadlen, M.N. *Nat. Neurosci.* **13**, 1020–1026 (2010).
2. Gibbon, J., Malapani, C., Dale, C.L. & Gallistel, C.R. *Curr. Opin. Neurobiol.* **7**, 170–184 (1997).
3. Eagleman, D.M. *Curr. Opin. Neurobiol.* **18**, 131–136 (2008).
4. Vierordt, K. *Der Zeitsinn nach Versuchen [The Time-Sense According to Experiments]* (Laupp, Tübingen, Germany, 1868).
5. Mamassian, P. *Psychol. Sci.* **19**, 601–606 (2008).
6. Geisler, W.S. & Kersten, D. *Nat. Neurosci.* **5**, 508–510 (2002).
7. Körding, K.P. & Wolpert, D.M. *Proc. Natl. Acad. Sci. USA* **101**, 9839–9842 (2004).
8. Zellner, A. *J. Am. Stat. Assoc.* **81**, 446–451 (1986).
9. Hudson, T.E., Maloney, L.T. & Landy, M.S. *PLOS Comput. Biol.* **4**, e1000130 (2008).
10. Knill, D.C. & Saunders, J.A. *Vision Res.* **43**, 2539–2558 (2003).
11. Burge, J., Fowlkes, C.C. & Banks, M.S. *J. Neurosci.* **30**, 7269–7280 (2010).
12. Stocker, A.A. & Simoncelli, E.P. *Nat. Neurosci.* **9**, 578–585 (2006).
13. Trommershäuser, J., Maloney, L.T. & Landy, M.S. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **20**, 1419–1433 (2003).
14. Maloney, L.T. & Mamassian, P. *Vis. Neurosci.* **26**, 147–155 (2009).
15. Adams, W.J., Graf, E.W. & Ernst, M. *Nat. Neurosci.* **7**, 1057–1058 (2004).

An abundance of grid cells



Several different classes of neurons are involved in the mapping of allocentric space. Two such classes, place cells and head-direction cells, have been found throughout the medial temporal lobe memory areas, including the presubiculum, parasubiculum and entorhinal cortex. Place cells represent spatial locations and features of the environment and head-direction cells are sensitive to the orientation of a rat's head with respect to the environment. A third major class of cells, known as grid cells, was recently identified in the medial entorhinal cortex (MEC). Grid cells, which encode abstract spatial structure, are of great interest because, unlike place cells, their response properties are independent of any particular environment, suggesting that they are involved in path integration mechanisms. On page 987, Boccara and colleagues report that grid cells are not unique to the MEC, where they were first found, but are also abundant in the pre- and parasubiculum.

Boccara and colleagues recorded from neurons throughout the presubiculum, parasubiculum and MEC of rats during food-motivated running in an open environment. They found grid cells in all three of these areas, interspersed with head-direction cells and border cells (another recently reported cell class that encodes the boundaries of a local environment). The relative proportions of each of these cell classes were comparable across presubiculum, parasubiculum and the deep layers of MEC. However, the rotational symmetry of the grid pattern and the theta modulation of presubiculum neurons were significantly weaker than those in MEC.

The pre- and parasubiculum project strongly to MEC, raising the possibility that they may be the sources of grid-cell properties in MEC neurons. However, it is also possible that grid properties are generated locally in each of these regions. Although the existence of grid cells in multiple areas does not definitively identify the mechanism for generating these unique response properties, it further specifies the neural network that supports the mapping of allocentric space.

Hannah Bayer