

Dispatch

Vision: How to Train Visual Cortex to Predict Reward Time

Balázs Hangya^{1,2} and Adam Kepecs¹

Little is known about how the brain learns to anticipate the timing of reward. A new study demonstrates that optogenetic activation of basal forebrain input is sufficient to train reward timing activity in the primary visual cortex.

“Your transaction is being processed.” When you see this message on the ATM screen, you are expecting a timed reward while the machine is shuffling out the appropriate notes into the cash dispenser: your money. But which part of your brain is actually involved in generating this expectation? In a series of recent papers [1–3], one published in this issue of *Current Biology* [4], Shuler and colleagues present a surprising answer: it is, at least in part, your primary visual cortex.

Primary visual cortex (V1) is the first station for cortical processing of visual information. It is textbook knowledge that V1 extracts specific aspects of the visual world and represents elementary features such as edges [5]. According to this classical feedforward view of processing, sensory information propagates from V1 to later stages of cortical processing where more and more complex features of the sensory world are extracted, and eventually to higher-order centers that assign behavioral significance to visual features. This framework successfully explains not

only early visual representations but also rapid object recognition, a key function of the primate visual system [6].

In recent years, the feedforward view of visual processing has undergone significant revision, with increasing appreciation for the role of feedback from higher cortical centers, as well as highly precise recurrent and lateral connectivity [7,8]. For instance, lateral connections are thought to mediate response modulation specific to the geometry of object boundaries, an important process for visual scene segmentation [8,9]. Top-down feedback allows V1 to act as an adaptive processor influenced by brain states; for instance, it can lead to attentional modulation that may even contribute to visual awareness [7,10]. A simple, yet dramatic example for how behavioral state impacts V1 is the observation that when mice run, the stimulus-evoked firing of V1 neurons can double while retaining stimulus selectivity [11,12,13]. In fact, primary sensory cortices have dedicated neurons that can represent not only low-level stimulus features but even behavioral contingencies such as reinforcers [14,15].

A particularly intriguing line of investigation into non-sensory representations in visual cortex was initiated by Shuler and Bear [1]. They fitted rats with head-mounted goggles that delivered full-field retinal stimulation to either the left or the right eye. These stimuli were cues that predicted the delayed delivery of a drop of water a few seconds later (Figure 1A). Their surprising discovery was that many rat V1 neurons modify their cue-evoked firing to predict the expected time of future rewards — coined ‘reward timing activity’ [1]. These reward-timing responses come in three different varieties: some neurons show sustained activation from stimulus presentation to the expected time of reward, others show inhibition during the same period, and a third group exhibits a firing rate peak at the expected reward time (Figure 1B). Liu *et al.* [4] in this

issue build on this work to better understand the mechanisms by which V1 neurons can be trained to exhibit such responses.

What might be the circuit origin of reward timing activity in V1? Is it a reflection of a ‘cognitive’ brain function that is relayed from other, higher cortical areas, such as the prefrontal cortex, via top-down feedback connections? Alternatively, cue-reward intervals may be generated within V1 circuitry, so that their timing needs to be learned with the help of an external reinforcement signal [4]. Neuromodulatory systems are a good candidate for providing such reinforcement signals as they are able to broadcast behaviorally relevant events broadly across cortical areas. Dopaminergic neurons in particular have been shown to represent reward prediction error signals that could be used for learning timed intervals [16]. However, V1 does not receive strong dopaminergic input and therefore Liu *et al.* [4] focused on another classic neuromodulator: acetylcholine.

Cholinergic neurons of the basal forebrain (BF) densely innervate V1, and are known to control cortical plasticity and enable sensory map reorganization [17,18]. To test the idea that they may provide the reinforcement signal required for reward timing, Liu *et al.* [4] rendered basal forebrain neurons light-responsive by expressing channelrhodopsin-2, and directed light onto axon terminals projecting to V1. They trained mice with an optogenetic conditioning protocol, in which they first provided a light cue to one of the eyes, and then after a delay activated the BF→V1 projection. Using this training protocol outside of a behavioral task led to the emergence of the same three types of reward timing activity as if the reward had been delivered. The optogenetically-conditioned responses retained their plasticity and were modifiable by further conditioning to either shorter or longer intervals. Thus replacing actual rewards with optogenetic activation of basal forebrain inputs to V1 is sufficient to recapitulate reward timing activity in V1 neurons.

The basal forebrain contains a neurochemically heterogeneous population of projection neurons, not only cholinergic ones. Therefore, Liu *et al.* [4] repeated this experiment, this time using cholinergic-specific optogenetic stimulation within visual cortex, and found that it was indeed sufficient to induce timing activity. One potential caveat to this experiment, however, is that in addition to the cholinergic BF→V1 projection they likely also activated a class of local interneurons that are both cholinergic and express the neuropeptide VIP. These VIP interneurons were shown to be activated by both reward and punishment in another primary sensory area, auditory cortex — and could contribute to this effect [14]. On the other hand, Chubykin *et al.* [2] also demonstrated that cholinergic-specific lesions of basal forebrain prevent the acquisition of reward timing activity in V1. In addition, even an isolated V1 slice *in vitro* could express reward timing activity via cholinergic mechanisms, further supporting the notion that reward timing activity emerges in V1 ‘de novo’ and not simply transmitted from higher cortical regions [2]. Taken together these studies provide a compelling case that basal forebrain cholinergic neurons are both necessary and sufficient for inducing reward timing activity.

Liu *et al.* [4] went on to establish a remarkable property of reward timing activity. Theoretically, a population of V1 neurons could encode not only the mean but also the animal’s temporal uncertainty about reward timing. Reward timing uncertainty should decrease with experience and increase with the duration of the time interval represented. Indeed, the authors found that further optogenetic conditioning decreased the variability of neural report of time, suggesting an experience-dependent decrease in the reported temporal uncertainty (Figure 1C). A general property of elapsed time estimation is that it follows Weber’s law, that is, the variance of time estimates increases linearly with elapsed time, referred to as scalar timing. The authors tested this hypothesis and determined that the variance of V1 timing activity increases with the learned

time duration, thus exhibiting the scalar timing property (Figure 1C). This result meshes nicely with a recent report by some of the same authors [19] demonstrating that Weber's law in time estimation leads to Weber's law in neural representations of subjective value and reward magnitude. They showed mathematically that multiple sources of noise result in scalar properties under the assumptions of an ecological model of decision making. This result is likely rooted in the generality of the Poisson limit theorem postulating that discrete distributions combining multiple independent sources of influence (such as the number of spikes fired by a neuron) can be approximated by Poisson distributions — and may explain the ubiquity of scalar representations in the brain. These results suggest that timing is represented in a population code in V1 that provides an estimate of both the mean and temporal uncertainty about anticipated reward time (Figure 1C-D).

Regardless of the precise mechanisms, these observations provide the clearest indication that visual cortex activity represents an internal model of the world beyond sensory signals. But is there behavioral relevance to time coding in V1? Another recent paper by Namboodiri and colleagues [3] take a big step forward in addressing this question. They trained rats on a visually cued interval timing task in which rats received reward proportional to their waiting time — but only up to a threshold, beyond which no reward was delivered. Rats learned to wait for the optimal interval, neither too short nor too long, albeit with natural variability from trial-to-trial. They found that in a large population of V1 neurons trial-to-trial variation in reward timing activity was correlated with behavioral waiting time. Moreover, optogenetic perturbation of V1 during the timed intervals led to an increase in waiting times, suggesting that V1 is causally involved in visually cued timed actions.

[h1] megjegyzést írt: no this is the 'thus' I meant. increase is not scalar timing, either omit 'thus' or say 'linearly increase'

These exciting findings raise a host of new questions about reward timing activity. One issue is whether reward has a special role in entraining V1 neurons, or might punishment also recruit similar basal forebrain mechanisms? Further studies will be required to understand the circuit and molecular mechanisms by which basal forebrain achieves training. Previous *in vitro* experiments revealed muscarinic effects [2] — however, precise timing may necessitate faster, nicotinic mechanisms as well. Potentially the cholinergic system might also rapidly engage specific cortical cell types and circuits [20] and exert some of its impact via cortical disinhibition [14,15].

Finally, perhaps the most burning open question raised by these studies is how reward timing activity maps onto other, better understood functions of V1? Are visual feature detection and temporal anticipation segregated at the circuit level, involving partially non-overlapping cell types or cortical layers? Or instead are these processes interwoven reflecting a broader unified function of visual cortex in making behaviorally relevant predictions based on visual information? These new questions will undoubtedly move research beyond the textbook paradigm of feedforward visual processing and lead to the exploration of novel principles for cortical computation as the construction of internal models of the world.

References

1. Shuler, M.G., and Bear, M.F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606–1609.
2. Chubykin, A.A., Roach, E.B., Bear, M.F., and Shuler, M.G.H. (2013). A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* 77, 723–35.

3. Namboodiri, V.M.K., Huertas, M.A., Monk, K.J., Shouval, H.Z., and Shuler, M.G.H. (2015). Visually Cued Action Timing in the Primary Visual Cortex. *Neuron* 86, 319–330.
4. Liu, C.-H., Coleman, J.E., Davoudi, H., Zhang, K., and Shuler, M.G.H. (2015). Selective activation of a putative reinforcement signal conditions cued interval timing in primary visual cortex. *Curr. Biol.* this issue.
5. Hubel, D.H., and Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148, 574–591.
6. Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6424–6429.
7. Gilbert, C.D., and Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron* 54, 677–696.
8. Stettler, D.D., Das, A., Bennett, J., and Gilbert, C.D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* 36, 739–750.
9. Self, M.W., van Kerkoerle, T., Supèr, H., and Roelfsema, P.R. (2013). Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation. *Curr. Biol.*, 2121–2129.
10. Lamme, V.A., and Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579.
11. Keller, G.B., Bonhoeffer, T., and Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron* 74, 809–815.

12. Niell, C.M., and Stryker, M.P. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron* 65, 472–479.
13. Saleem, A.B., Ayaz, A., Jeffery, K.J., Harris, K.D., and Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nat. Neurosci.* 16, 1864–1869.
14. Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J. I., Huang, Z. J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. *Nature* 503, 521–524.
15. Letzkus, J.J., Wolff, S.B.E., Meyer, E.M.M., Tovote, P., Courtin, J., Herry, C., and Lüthi, A. (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* 480, 331–5.
16. Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
17. Froemke, R.C., Merzenich, M.M., and Schreiner, C.E. (2007). A synaptic memory trace for cortical receptive field plasticity. *Nature* 450, 425–9.
18. Kilgard, M.P., and Merzenich, M.M. (1998). Plasticity of temporal information processing in the primary auditory cortex. *Nat. Neurosci.* 1, 727–31.
19. Namboodiri, V.M.K., Mihalas, S., and Shuler, M.G.H. (2014). A temporal basis for Weber’s law in value perception. 8, 1–11.
20. Alitto, H.J., and Dan, Y. (2012). Cell-type-specific modulation of neocortical activity by basal forebrain input. *Front. Syst. Neurosci.* 6, 79.

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 11724, USA. ²Laboratory of Cerebral Cortex Research, Institute of Experimental Medicine, Hungarian Academy of Sciences, Budapest, H-1083, Hungary. E-mail: kepecs@cshl.edu

Figure 1. Reward timing activity in neural populations of visual cortex.

(A) Experimental paradigm: mice received visual stimuli that predicted water reward or laser stimulation of basal forebrain projections to V1 after a fixed delay. (B) Responses of V1 neurons from the visual cue (green dashed line) to the reward (blue dashed line) anticipate reward time. Three types of interval coding neurons were found in V1: sustained activated, sustained inhibited and peak firing neurons corresponding to neural expectation of waiting time. (C) V1 activity can be decoded to generate a neural report of waiting time. The variance of the neural report of time increased with waiting time and decreased with experience. (D) The neural report of time encodes uncertainty and follows scalar timing for waiting time. With more experience there is less increase of uncertainty with longer waiting times.

In Brief:

Little is known about how the brain learns to anticipate the timing of reward. A new study demonstrates that optogenetic activation of basal forebrain input is sufficient to train reward timing activity in the primary visual cortex.