

On the value of information and other rewards

Yael Niv & Stephanie Chan

Knowledge is not just power. Even if advance information can not influence an upcoming event, people (and animals) prefer to know ahead of time what the outcome will be. According to the firing patterns of neurons in the lateral habenula, from the brain's perspective, knowledge is also water—or at least its equivalent in terms of reward.

That advance knowledge is valuable is intuitively clear: knowing that it will rain this afternoon allows us to leave home with an umbrella. But how important is it to know in advance whether something will or will not happen in 270 days? Apparently, quite a bit, especially if the event in question is the birth of your baby. In fact, some people will happily pay up to 40% more for a test that can detect pregnancy 1 day earlier than conventional tests, attesting to the high value of even slightly earlier advance information.

It turns out that this paradoxical behavior is not unique to humans. In this issue of *Nature Neuroscience*, Bromberg-Martin and Hikosaka¹ show that monkeys also prefer to have advance knowledge of impending rewards. Moreover, by recording neurons in the lateral habenula (a nucleus in the epithalamus that encodes prediction error signals that are related to the value of impending rewards²), they find that monkeys value cues that indicate the availability of advance information more than they do cues that indicate that such information will not be forthcoming. That is, if they could read the labels at the pharmacy, monkeys too would value an early pregnancy test more than a regular pregnancy test.

Notably, Bromberg-Martin and Hikosaka¹ show that the value of information (or, more precisely, the change in predicted information) is encoded by the same habenula and dopamine neurons that encode (changes in) predicted primary reward, and in a common currency. Thus, from the brain's perspective, information and reward are exchangeable.

To measure both separate and joint coding of information and reward values, Bromberg-Martin and Hikosaka¹ devised a simple, but

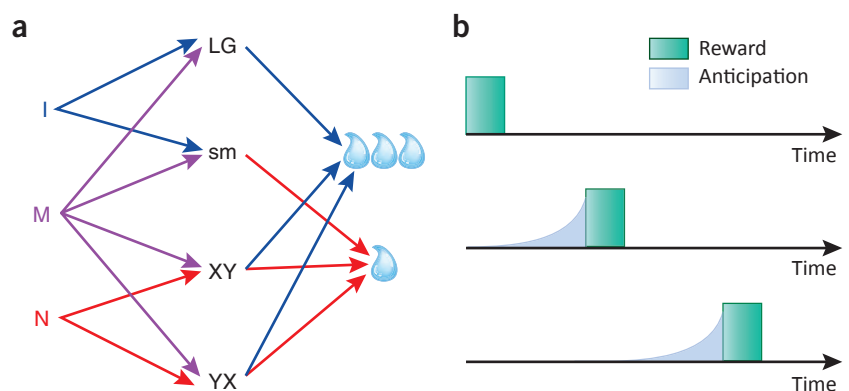


Figure 1 Anticipation and the value of information. (a) The structure of a trial in Bromberg-Martin and Hikosaka's experiment¹. All arrows emanating from one node are equally likely to occur. (b) If we know of an upcoming reward in advance, its anticipation value is added to its reward value. The hypothesized temporal profile of anticipation means that, compared to an immediate reward (top), a reward in the intermediate future will benefit from the added bonus of anticipation (middle). However, a reward very far in the future (bottom) will not recruit more anticipation value than the intermediate-term reward.

ingeniously balanced, task. Monkeys chose between two of three possible visual cues (Fig. 1a). One cue, 'I' (for information), was followed by one of two stimuli, 'LG' or 'sm', which told the monkey whether the upcoming reward would be a large (0.88 ml) or small (0.04 ml) amount of water. Another cue, 'N' (no information), was followed by one of two cues (XY or YX) that were each equally likely to be followed by a large or small amount of water, and thus gave the monkey no advance information about the upcoming reward. A third cue, 'M' (for maybe), was equally likely to be followed by the informative LG or sm cues from the I trials or the random XY or YX cues from the N trials. In any case, each trial was equally likely to result in a small or large water reward, and all the monkey was deciding was whether to get advance information about the size of the upcoming reward or not.

That cues signaling large rewards have more value than cues signaling small rewards

is well established and makes sense; if we did not attach more value to an invitation to a free dinner at a fancy restaurant than at the corner fast-food joint, we would not know to seek the former invitation rather than the latter. However, the value of advance information is much less straightforward. Imagine you were to join a 'daily gourmet dinner club' that delivers a chef-cooked fresh meal every evening for a week. Would you pay more for a membership plan that added a daily email at 5 p.m. announcing the evening's menu?

In the monkey experiment, any extra value attached to the I cue is equivalent to the higher value of the email-added plan. Specifically, the I cue is equally likely to be followed by a LG or sm cue. Its value in terms of future expected reward should therefore be the mean of the two water amounts. Similarly, the N and M cues are also equally likely to lead to either reward, and their value should therefore be identical to that of the I cue. Still, the monkeys

Yael Niv and Stephanie Chan are at the Princeton Neuroscience Institute and the Psychology Department, Princeton University, Princeton, New Jersey, USA.
e-mail: yael@princeton.edu

unequivocally preferred advance knowledge, choosing the I cue over the M cue and the M cue over the N cue almost exclusively. In fact, the monkeys were more consistent in choosing advance information than they typically are in choosing a more probable reward over a less probable one (compare ref. 3), even though their choices had no bearing on the actual impending reward.

Economists might not be surprised by this finding: such preference for 'temporal resolution of uncertainty' is documented even in cases in which the advance information has no bearing on subsequent behavior⁴. But people don't seek advance information in all cases: they consistently prefer early information about good outcomes such as monetary reward, but they often choose to delay (or totally avoid) information about potential bad outcomes such as pain⁵. The surprising corollary is that people prefer to delay good outcomes such as a meal in a fancy restaurant, but would like aversive events to happen sooner rather than later^{6–8}, which stands in direct contradiction to temporal discounting, that is, the discounted value of outcomes far in the future.

The traditional explanation for this pattern of preferences, dating back to the philosopher Jeremy Bentham in the late eighteenth century, is that knowing that a reward is forthcoming allows one to enjoy not only the reward, but also its anticipation (termed 'savoring' by the economist George Loewenstein⁶). Thus, the value of a future reward that we know about in advance (and can therefore savor) is higher than the value of the same reward without the advance knowledge, or the same reward obtained immediately. Conversely, advance knowledge about aversive events induces dread, a negatively valued emotion, and thus it is better to not know about these events in advance, or to simply get them over with.

Of course, we don't want to delay rewards indefinitely; while you might prefer that fancy dinner in a week rather than right now, you would not want to delay it by a year. This suggests that anticipation only has a limited ability to enhance (or decrease) the value of future outcomes and therefore cannot overcome temporal discounting indefinitely. The reason for this, according to economists, is that anticipation has a characteristic temporal profile: it starts a certain amount of time before the outcome and increases until the time of the outcome itself (Fig. 1b). The value that advance knowledge adds to the value of the outcome is the total anticipation over time, which is thus bounded^{6,7}.

If all Bromberg-Martin and Hikosaka¹ found was this behavioral preference of information in monkeys, theirs would be an interesting,

if not all that surprising result. Their report, however, goes far beyond the behavior: it investigates the neural underpinnings of these behavioral preferences. Bromberg-Martin and Hikosaka previously found that midbrain dopaminergic neurons attach value to advance information signals just as they do to cues predicting reward⁹. That is, these neurons transiently increase their firing to the appearance of both types of cues. Here they found that this is also true for a subpopulation of neurons in the lateral habenula, a structure that inhibits dopamine neurons and tends to be excited by negative events (such as reward omission), and inhibited by positive events. Previous work has shown that neurons in the habenula transiently decrease firing to a cue that predicts more reward². Bromberg-Martin and Hikosaka's results¹ show that these same neurons decrease firing to cues that predict more information. The converse is also true: habenula neurons increase firing to cues that predict less reward and to those that predict less information¹.

For the record, consistent with previous reports, the recordings from Bromberg-Martin and Hikosaka^{1,9} show that lateral habenula neurons, as well as dopamine neurons, do not signal the (information or reward) value of cues *per se*. Rather, they signal errors in the prediction of value. That is, these neurons fire when there is a discrepancy between the currently anticipated value and the value anticipated at the previous time point, in surprisingly tight correspondence with theoretical predictions from reinforcement learning algorithms (the expected value itself, a signal with a different temporal profile, has been shown to reside elsewhere in the brain^{10,11}). This is why both habenula and dopamine neurons fired transiently to the onset of the high-value LG or low-value sm cues; before the arrival of the cue, the monkey could not predict which reward would appear on this trial, and presumably expected their average (0.46 ml reward). On seeing the cue, the predicted reward value increased to 0.88 ml for the LG cue (or decreased to 0.04 ml for the sm cue), inducing a positive (or negative) prediction error of 0.42 ml. For this same reason, the neurons did not fire to the reward itself when it was fully predicted by a preceding LG or sm cue, but did respond to the receipt of water in trials in which the preceding cue, XY or YX, was not informative.

Thus, with a little math, one can derive the value of expected rewards from these prediction error signals at any point during the task. Crucially, from the fact that both habenula and dopamine neurons showed different responses to the information-predicting I, N and M cues that had the same value in terms of predicted reward and only differed in terms of predicted

information, one can deduce that information itself has value. In fact, in the context of the current task, the value of impending information was equivalent, in terms of neural firing, to as much as 0.17 ml of water—striking given that this information only preceded the reward by about 2 s!

The report goes on to show that, as would be expected of a true information prediction error signal, uninformative cues (for instance, XY) cause habenula neurons to fire only if this lack of information was unexpected. Furthermore, consistent with the idea that information and reward have commensurable value that is simply added up, the high-value LG cue caused an even larger dip in firing rate when this cue appeared after an M choice than after an I choice, as the monkey did not know that the information would be forthcoming. Similarly, the low-value sm cue was not as disappointing when it appeared in an M trial as in an I trial, attesting the additive value of surprising information, even if that information indicates that reward is in fact not forthcoming.

At first blush these results also make perfect sense, joining other recent work that suggests that neural prediction errors are not concerned only with primary reward^{12,13}. However, that the value of reward and the value of information are signaled by the same brain areas and in a common (neural) currency is far from trivial. Behaviorally, this predicts that monkeys should be willing to forfeit actual reward in return for advance information about an outcome, suboptimal behavior that has recently been demonstrated in pigeons¹⁴. Computationally and neurally, these results pose a challenge to existing models of reinforcement learning and their associated prediction error signals: because information prediction errors are encoded in exactly the same way as reward prediction errors, theories that have explained the provenance of the latter must now also explain the former.

But this is easier said than done. Bromberg-Martin and Hikosaka¹ model their data using an 'information bonus' model in which informative cues carry reward value. The big question, left unanswered in the economics literature as well, is how does the brain know how much reward an information cue is worth? Human choice data show that the value of information is not merely a reflection of the reduction in uncertainty in information-theoretic terms. Otherwise, the value of information would not depend on the timing and magnitude of the forthcoming reward. The alternative Bentham-inspired scheme in which the reward value of information derives from the value of (time-varying) anticipation is, unfortunately, not easy (and perhaps

impossible) to learn using prediction errors such as are encoded by dopamine and habenula neurons. This contradicts reinforcement learning theory, which contends that the values feeding in to the calculation of prediction errors are learned using these same prediction errors¹⁵. The resolution, either a modification of reinforcement learning algorithms, or reconciliation with other mechanisms by which these values might be computed, is yet to come.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Bromberg-Martin, E.S. & Hikosaka, O. *Nat. Neurosci.* **14**, 1209–1216 (2011).
2. Matsumoto, M. & Hikosaka, O. *Nature* **447**, 1111–1115 (2007).
3. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. *Nat. Neurosci.* **9**, 1057–1063 (2006).
4. Eliaz, K. & Schotter, A. *Games Econ. Behav.* **70**, 304–324 (2010).
5. Averill, J.R. & Rosenn, M. J. *Pers. Soc. Psychol.* **23**, 128–141 (1972).
6. Loewenstein, G. *Econ. J.* **97**, 666–684 (1987).
7. Harris, C.R. *J. Behav. Decis. Mak.* published online, doi:10.1002/bdm.709 (2 July 2010).
8. Berns, G.S. *et al. Science* **312**, 754–758 (2006).
9. Bromberg-Martin, E.S. & Hikosaka, O. *Neuron* **63**, 119–126 (2009).
10. Schultz, W., Tremblay, L. & Hollerman, J.R. *Neuropharmacology* **37**, 421–429 (1998).
11. Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W. & Rangel, A. *J. Neurosci.* **28**, 5623–5630 (2008).
12. Wittmann, B.C., Daw, N.D., Seymour, B. & Dolan, R.J. *Neuron* **58**, 967–973 (2008).
13. Ribas-Fernandes, J.J.F. *et al. Neuron* **71**, 370–379 (2011).
14. Zentall, T.R. & Stagner, J. *Proc. R. Soc. Lond. B* **278**, 1203–1208 (2010).
15. Niv, Y. & Schoenbaum, G. *Trends Cogn. Sci.* **12**, 265–272 (2008).

Grilled RIBEYE stakes a claim for synaptic ribbons

Jeffrey S Diamond

Photo-inactivation of RIBEYE, an important presynaptic protein at ribbon synapses in the retina, reveals a new role for ribbons in the vesicle priming process.

Synaptic ribbons are enigmatic specializations at presynaptic active zones in neurons that encode sensory information with small, graded voltage signals instead of large, all-or-none action potentials. In retinal photoreceptors and bipolar cells, ribbons are distinctive, planar structures that are oriented orthogonally to the presynaptic membrane and tethered to dozens of synaptic vesicles (Fig. 1). Ribbon synapses transmit analog sensory information, suggesting that they have evolved to sustain continuous vesicle release for long periods, but their specific role in the release process remains unclear. In this issue of *Nature Neuroscience*, Snellman *et al.*¹ photolytically inactivate ribbons and find that they help prime vesicles before release¹.

The synaptic ribbon has been proposed to function as either a conveyor belt, actively delivering synaptic vesicles to release sites at the presynaptic membrane, or as a safety belt that stabilizes vesicles next to each other, possibly so that they can fuse together before fusing with the plasma membrane². Evidence that replenishment of the readily releasable pool (RRP), those vesicles at the base of the ribbon that are primed for release (Fig. 1a), is relatively slow^{3,4} suggests that the ribbon may act as a timing belt³ to regulate occupancy of the RRP. Whether or how ribbons actually accomplish any of these tasks remains unclear. Genetically disconnecting ribbons from the plasma membrane disrupts release and active zone organization⁵, but until now it has not

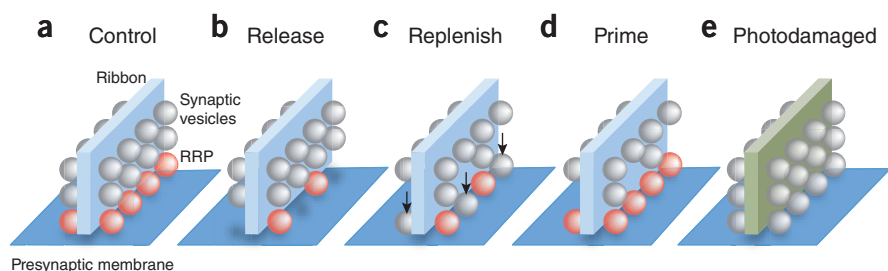


Figure 1 The life cycle of vesicles on a synaptic ribbon. (a) Ribbons in the retina tether synaptic vesicles to both sides of their plate-like structure. The vesicles at the base of the ribbon, closest to the presynaptic membrane, are thought to make up the primed RRP. (b) Membrane depolarization elicits vesicle release and depletion of the RRP. (c) Vacated sites at the base of the ribbon become reoccupied, probably by other vesicles on the ribbon. (d) Vesicles at the base of the ribbon must undergo an ATP-dependent priming step before release. (e) When RIBEYE has been photodamaged, vesicles can occupy vacated sites at the base of the ribbon, but are not primed to be released.

been possible to study the same synapses before and after eliminating ribbon function.

To inactivate ribbons acutely, Snellman *et al.*¹ used a fluorescein-labeled peptide⁶ that binds RIBEYE, the principal protein component of synaptic ribbons⁷, and delivered it through a patch electrode into the cytosol of mouse rod bipolar cells (RBCs) in retinal slices. They also performed similar experiments in salamander cone photoreceptors. Fluorescence microscopy confirmed that the peptide localized specifically to presynaptic active zones, and paired recordings between synaptically coupled RBCs and postsynaptic amacrine cells revealed that the peptide itself did not disrupt ribbon function. As usual, synaptic responses comprised transient and sustained components, likely reflecting rapid depletion and subsequent replenishment, respectively, of the RRP^{3,8,9} (Fig. 1b,c).

After recording control responses, the authors exposed the slice to light to photobleach the fluorescein, generating singlet oxygen molecules that damage proteins in the immediate vicinity

(within ~40 Å)¹⁰, a process that is known as fluorophore-assisted light inactivation (FALI)¹¹. Ribbon fluorescence did not recover after FALI, indicating that RIBEYE was either permanently damaged or at least no longer able to bind the peptide. In the first synaptic response evoked after FALI, the transient component was relatively normal, but the sustained component was reduced, suggesting that, immediately after RIBEYE damage, vesicles in the pre-existing RRP could be released, but not replaced. Accordingly, subsequent responses revealed greatly diminished transient and sustained components. Repeated FALI steps eventually eliminated the responses altogether, arguing that all evoked release from RBCs requires functional ribbons. In salamander cones, in which FALI caused similar reductions in transmitter release, membrane capacitance measurements also confirmed that RIBEYE disruption did not affect endocytosis, effectively ruling out one possible mechanism by which RRP replenishment could have been impaired.

Jeffrey S. Diamond is in the Synaptic Physiology Section, National Institute of Neurological Disorders and Stroke, US National Institutes of Health, Bethesda, Maryland, USA.
e-mail: diamondj@ninds.nih.gov