State-dependent value representation: evidence from the striatum

Christopher J. Burke[1], Jean-Claude Dreher[2], Ben Seymour[3,4] and Philippe N. Tobler[1]*

[1] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland
[2] Neuroeconomics Laboratory: Reward and Decision-Making, CNRS, UMR 5229, Université de Lyon, Université Claude Bernard Lyon 1, Lyon, France
[3] Center for Information and Neural Networks, National Institute for Information and Communications Technology, Osaka, Japan
[4] Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge, UK
*Correspondence: phil.tobler@econ.uzh.ch

A commentary on

Encoding of aversion by dopamine and the nucleus accumbens
by McCutcheon, J. E., Ebner, S. R., Loriaux, A. L., and Roitman, M. F. (2012). Front. Neurosci. 6:137. doi: 10.3389/fnins.2012.00137

Acute stress influences neural circuits of reward processing
by Porcelli, A. J., Lewis, A. H., and Delgado, M. R. (2012). Front. Neurosci. 6:157. doi: 10.3389/fnins.2012.00157

The ability to distinguish good from bad options, to approach the former and avoid the latter, forms the basis of successful behavior. This ability is expressed in value-based decisions, which in turn is thought to depend largely on the process of reinforcement learning. In order to adaptively determine the value of different actions, organisms need to take external as well as internal states into account (e.g., Rangel et al., 2008). For example, finding shelter may be more valuable in a cold environment than in a warm environment. Internal states can also affect valuation, as illustrated for instance by salt appetite (Berridge et al., 1984; Tindell et al., 2009; Robinson and Berridge, 2013): In the normal (non-salt-deficient) state, rats do not usually ingest extremely salty solutions or approach cues that predict them. However, in a salt-deficient state, they do. This pattern of behavior is compatible with the notion that state information can have such a profound impact

on value computation that a previously bad option becomes good (Dayan and Berridge, 2014).

Since its inception in the seventeenth century, economic choice theory has gradually come to recognize the importance of internal and external states on valuation. While it was initially thought that a given monetary unit was worth the same no matter how wealthy one is (Pascal), it was later proposed that the value of a given monetary unit is greater when one is in a state of poverty as compared to one of affluence (Bernoulli). In the last century, researchers found that our expectations also affect valuation and accounted for this finding by incorporating reference points into value functions (e.g., prospect theory: Kahneman and Tversky, 1979; see also Koszegi and Rabin, 2006). For example, we value a salary raise of USD 200 less when we originally expected to receive a raise of USD 400 than when we did not expect to receive any raise at all. Thus, the value of an option can also depend on cognitive states.

The states and other variables that influence valuation are manifold and include not only financial status and expectations, but also mood, emotion, motivation, previous learning, and social aspects. Still, little is known about how state information influences valuation at the neural level. Two recent publications (McCutcheon et al., 2012; Porcelli et al., 2012) addressed this question. The two studies used different techniques (dopamine voltammetry vs. functional magnetic resonance imaging), different model organisms (rats vs. humans), and

different value-impacting state parameters (previous learning vs. stress). Despite these differences, both studies found that striatal value signals are state-dependent.

In the first study, McCutcheon et al. (2012) measured dopamine release in the nucleus accumbens shell following intra-oral infusion of sucrose. In half of the rats, sucrose was rendered aversive by pairing it with induced nausea (via injection of lithium chloride just after sucrose consumption). In the other half of the rats, nausea induction and sucrose consumption occurred on different days, so sucrose remained appetitive. As expected, aversive sucrose elicited fewer appetitive and more negative orofacial responses. Neurobiologically, unlike appetitive sucrose, aversive sucrose reduced accumbens dopamine concentration compared to baseline, even though the sensory properties of the sucrose were held constant in the two conditions. This finding converges with several previous reports of reduced dopamine firing and concentrations induced by aversive stimuli (for review, see McCutcheon et al., 2012). Conversely, appetitive sucrose elicited a (weak) increase in dopamine, in line with previous voltammetry data (Roitman et al., 2008) and a wealth of previous findings implicating dopamine in reward processing (for a review, see Daw and Tobler, 2013).

The reduction in accumbens dopamine concentration in response to aversive sucrose shows that learning can change the value of a primary appetitive stimulus and provides a pharmacological foundation

to the reduction in striatal activation observed in imaging studies in which participants' reward expectations were disappointed (e.g., McClure et al., 2003; O'Doherty et al., 2003; Pessiglione et al., 2006; Burke et al., 2010; Kahnt et al., 2012).

In the second study, Porcelli and colleagues investigated the flexibility of reward processing in response to induced stress (an internal state factor). They adapted the cold pressor task, an established stress induction procedure (Schwabe et al., 2008), for use in the scanner. They used MRI-compatible gel-pacs to subject half of the participants to stress-inducing cold temperatures and half to room temperature (for details on the procedure, see Porcelli, 2014). Both groups then performed a card-guessing task in which they could win $5 or $0.50, or lose $0.25 or $2.50. The asymmetry between the gain and loss domains aimed to ensure that the impact of reward and punishment on behavior and brain activity was similar and thus to compensate for the fact that people generally exhibit loss aversion (another feature of prospect theory; Kahneman and Tversky, 1979). Throughout the experiment, the level of salivary cortisol (a stress-inducing hormone) was measured at 15 minute intervals using an oral swab.

The adapted cold pressor task resulted in greater feelings of discomfort and a higher cortisol level in the stress induction group than in the control group. However, the most striking results from this study pertain to the striatal BOLD response: Responses to rewards as compared to punishments were much larger in the caudate and putamen in the control group than in the stress induction group. Participants in the latter group only showed reward-related activation in this region when outcomes were of high magnitude, suggesting that stress causes a desensitization of the reward network. In addition, the reduction in response to rewards under stress was observed in the dorsal, but not the ventral striatum. These regions have been shown to correlate with reward learning in a manner consistent with computational models of reinforcement learning (specifically, the actor-critic model proposed by Barto, 1995), with dorsal actor regions processing action contingencies that guide future choices, and ventral critic regions

making predictions about future rewards (O'Doherty et al., 2004). If stress causes a desensitization to reward in the "actor" areas of the striatum, action-outcome contingencies may not be processed accurately enough to guide future choices, forcing the organism to rely on more habitual responses. This may manifest itself in a return to otherwise suboptimal reward-seeking behavior, such as the relapses commonly seen in recovering addicts (Everitt and Robbins, 2005).

Moreover, one could hypothesize that stress, which by itself increases tonic levels of dopamine (e.g., Inoue et al., 1994), may prevent the detection of phasic reductions in dopamine that could be elicited by stimuli that predict negative drug-related effects (see also Weiss et al., 2001; Schultz, 2011). This notion would predict that stress reduces sensitivity to losses through a dopamine-dependent mechanism. Incidentally, one possible target region for implementing this mechanism is suggested by the Porcelli study (Porcelli et al., 2012), which reports reduced magnitude discrimination under stress in the inferior frontal gyrus. Given that the inferior frontal gyrus appears to play a role in response inhibition and cognitive control (e.g., Bari and Robbins, 2013), it might be worth investigating the role of dopamine on punishment sensitivity under withdrawal-induced stress in that region.

Taken together, the two papers discussed here support the notion that reward processing in the brain is flexible and highly dependent on internal and external states. The interaction between the environment and internal states allows organisms to prioritize their goals and adapt their behavior accordingly. We suggest that this inherent flexibility can be detrimental when reward and learning systems are artificially challenged, for example, through the use of addictive drugs. While the use of drugs themselves may affect internal states (such as by tonically enhancing dopamine levels), subsequent withdrawal may induce stress, causing a change in behavior, e.g., in the form of a shift toward habitual responding and enhanced drug-seeking. Moreover, state processes could interact with drug effects and further research may wish to investigate how this interaction contributes to addiction and relapse. While the two papers (McCutcheon et al.,

2012; Porcelli et al., 2012) give some leads for that endeavor they more generally highlight the importance of external and internal states for brain and behavior.

REFERENCES

Bari, A., and Robbins, T. W. (2013). Inhibition and impulsivity: behavioral and neural basis of response control. *Prog. Neurobiol.* 108, 44–79. doi: 10.1016/j.pneurobio.2013.06.005

Barto, A. G. (1995). "Adaptive critics and the basal ganglia," in *Models of Information Processing in the Basal Ganglia*, eds J. C. Houk, J. L. Davis, and D. G. Beiser (Cambridge, MA: MIT Press), 215–232.

Berridge, K. C., Flynn, F. W., Schulkin, J., and Grill, H. J. (1984). Sodium depletion enhances salt palatability in rats. *Behav. Neurosci.* 98, 652–660. doi: 10.1037/0735-7044.98.4.652

Burke, C. J., Tobler, P. N., Baddeley, M., and Schultz, W. (2010). Neural mechanisms of observational learning. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14431–14436. doi: 10.1073/pnas.1003111107

Daw, N. D., and Tobler, P. N. (2013). "Value learning through reinforcement: the basics of dopamine and reinforcement learning," in *Neuroeconomics, 2nd Edn.*, eds P. W. Glimcher and E. Fehr (Oxford: Academic Press), 283–298.

Dayan, P., and Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* 14, 473–492. doi: 10.3758/s13415-014-0277-8

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489. doi: 10.1038/nn1579

Inoue, T., Tsuchiya, K., and Koyama, T. (1994). Regional changes in dopamine and serotonin activation with various intensity of physical and psychological stress in the rat brain. *Pharmacol. Biochem. Behav.* 49, 911–920. doi: 10.1016/0091-3057(94)90243-7

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291. doi: 10.2307/1914185

Kahnt, T., Park, S. Q., Burke, C. J., and Tobler, P. N. (2012). How glitter relates to gold: similarity-dependent reward prediction errors in the human striatum. *J. Neurosci.* 32, 16521–16529. doi: 10.1523/JNEUROSCI.2383-12.2012

Koszegi, B., and Rabin, M. (2006). A model of reference-dependent preferences. *Q. J. Econ.* 121, 1133–1165. doi: 10.1093/qje/121.4.1133

McClure, S. M., Berns, G. S., and Montague, P. R. (2003). Temporal prediction errors in a passive

learning task activate human striatum. *Neuron* 38, 339–346. doi: 10.1016/S0896-6273(03)00154-5

McCutcheon, J. E., Ebner, S. R., Loriaux, A. L., and Roitman, M. F. (2012). Encoding of aversion by dopamine and the nucleus accumbens. *Front. Neurosci.* 6:137. doi: 10.3389/fnins.2012.00137

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454. doi: 10.1126/science.1094285

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337. doi: 10.1016/S0896-6273(03)00169-7

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045. doi: 10.1038/nature05051

Porcelli, A. J. (2014). An alternative to the traditional cold pressor test: the cold pressor arm wrap. *J. Vis. Exp.* 83, 1–5. doi: 10.3791/50849

Porcelli, A. J., Lewis, A. H., and Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Front. Neurosci.* 6:157. doi: 10.3389/fnins.2012.00157

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. doi: 10.1038/nrn2357

Robinson, M. J. F., and Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational "wanting". *Curr. Biol.* 23, 282–289. doi: 10.1016/j.cub.2013.01.016

Roitman, M. F., Wheeler, R. A., Wightman, R. M., and Carelli, R. M. (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nat. Neurosci.* 11, 1376–1377. doi: 10.1038/nn.2219

Schultz, W. (2011). Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. *Neuron* 69, 603–617. doi: 10.1016/j.neuron.2011.02.014

Schwabe, L., Haddad, L., and Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology* 33, 890–895. doi: 10.1016/j.psyneuen.2008.03.001

Tindell, A. J., Smith, K. S., Berridge, K. C., and Aldridge, J. W. (2009). Dynamic computation of incentive salience: "wanting" what was never "liked". *J. Neurosci.* 29, 12220–12228. doi: 10.1523/JNEUROSCI.2499-09.2009

Weiss, F., Ciccocioppo, R., Parsons, L. H., Katner, S., Liu, X., Zorrilla, E. P., et al. (2001). Compulsive drug-seeking behavior and relapse. Neuroadaptation, stress, and conditioning factors.*Ann. N.Y. Acad. Sci.* 937, 1–26. doi: 10.1111/j.1749-6632.2001.tb03556.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.