Learning from silver linings

R. McKell Carter[1,2]* and Scott A. Huettel[1,2,3]*

[1] Center for Cognitive Neuroscience, Duke University, Durham, NC, USA
[2] Brain Imaging and Analysis Center, Duke University, Durham, NC, USA
[3] Department of Psychology and Neuroscience, Duke University, Durham, NC, USA
*Correspondence: mckell.carter@duke.edu; scott.huettel@duke.edu

Edited by:
Hauke R. Heekeren, Freie Universität Berlin, Germany

A commentary on

From bad to worse: striatal coding of the relative value of painful decisions
by Brooks, A. M., Pammi, V. S. C., Noussair, C., Capra, C. M., Engelmann, J. B., and Berns, G. S. (2010). Front. Neurosci. 4:176. doi: 10.3389/fnins.2010.00176

The cardinal goal of decision neuroscience is to identify the neural mechanisms that translate external rewards into an internal sense of value (Rangel et al., 2008; Huettel, 2010; Glimcher, 2011; Rangel and Clithero, 2012). The brain region most commonly associated with "reward" has been the ventral striatum (vStr, neurosynth.org forward inference, 1/2013). Studies of fMRI activation and neuronal activity in the vStr have converged on a standard model in which the vStr computes a reward prediction error [RPE (Schultz, 1997; Pagnoni et al., 2002; O'Doherty et al., 2003; Bayer and Glimcher, 2005)] that then facilitates the learning of new associations (Sutton and Barto, 1998). The RPE mechanism provides flexibility for encoding the relationships between cues and positive rewards across a wide dynamic range, but was historically thought to ignore aversive stimuli (Delgado et al., 2000; Knutson et al., 2001).

A strong split between representations of positively and negatively valued stimuli would be consistent with evidence from classical conditioning (Martin-Soelch et al., 2007) and emotion (Russell, 1980) representation. But, it is inconsistent with a growing body of research that indicates that the vStr can represent the magnitude of value for both positive and negative stimuli (Seymour et al., 2005, 2007; Carter et al., 2009). Research by Brooks and colleagues (Brooks et al., 2010) has pointed toward a way to reconcile this conflicting evidence: negative expectations create a baseline that allows two negative stimuli to be distinguished in the vStr.

Brooks and colleagues presented participants with a series of choices between a constant number of electric shocks and a number based on the outcome of a gamble. The use of primary sensory stimuli in choice tasks is rare but potentially valuable (O'Doherty et al., 2006). Electric shocks in particular provide primary sensory input that is very aversive and can take place in an entirely negative context (i.e., without the use of an endowment). Prior to choice, on each trial, participants were given a standard number of shocks to establish a negative expectation for each trial. The participants' choices were then used to characterize their utility curves. Consistent with prospect theory (Kahneman and Tversky, 1979), the convex shape of these curves indicates that participants, in spite of their preference for fewer shocks, still viewed fewer shocks as a negative outcome. The authors next show that the vStr not only represents aversive choice options, but also does so in a manner that is counter to traditional salience arguments (Blackburn et al., 1992; Salamone, 1994). Less aversive—and therefore less salient—options produce greater activations. This surprising finding has important implications for how associations are learned and subsequent choices are made.

By embedding local context in an RPE framework, Brooks and colleagues can explain two puzzles in the decision neuroscience literature. First, the lack of vStr activation for aversive but unexpectedly positive experiences can now be understood as a problem with firing-rate sensitivity. Negative stimuli that evoke low firing rates can be extraordinarily difficult to distinguish from one another. But, in the local context provided by the negative reference shocks, negative stimuli generate higher overall firing rates and could be easily distinguished.

Second, because these negative stimuli can generate different representations in vStr, the neural machinery responsible for learning positive rewards can be co-opted to choose the more positive of two negative options; widening the applicability of temporal difference models of learning (Sutton and Barto, 1998). The use of local context in temporal difference learning has been described in work on relief from pain (Seymour et al., 2005), and is a potential explanation for striatal representation of monetary losses in one of our own studies where gain and loss contexts were held constant within runs (Carter et al., 2009).

While being able to distinguish negative stimuli in using the vStr expands the applicability of reinforcement models of learning, a number of questions regarding the representation of negative stimuli during choice remain. Work from Hikosaka and colleagues [reviewed in Bromberg-Martin et al. (2010)] has indicated that the vStr may also incorporate signal from dopamine neurons, anatomically distinct in origin, that fire more strongly to both rewards and punishments. Such findings raise the intriguing possibility that a single experiment could reveal anatomically distinct regions within the vStr that evince distinct reward and salience coding—following a clever manipulation of local context. We also note that modern models of decision value have difficulty predicting choices for gambles containing mixed outcomes that include potential gains and losses (Payne, 2005). In order to address these potential shortcomings, choice sets consisting of true mixed gambles may provide important methodological advantages (Venkatraman et al., 2009).

Although much attention has been paid to the representation of aversive stimuli in the vStr, this study by Brooks

and colleagues provides an important and novel reminder: subtle differences in experimental protocol can drastically change the neural response to a simple stimulus.

REFERENCES

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Blackburn, J., Pfaus, J., and Phillips, A. (1992). Dopamine functions in appetitive and defensive behaviours. *Prog. Neurobiol.* 39, 247–279.

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834.

Brooks, A. M., Pammi, V. S. C., Noussair, C., Capra, C. M., Engelmann, J. B., and Berns, G. S. (2010). From bad to worse: striatal coding of the relative value of painful decisions. *Front. Neurosci.* 4:176. doi: 10.3389/fnins.2010.00176

Carter, R. M., Macinnes, J. J., Huettel, S. A., and Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Front. Behav. Neurosci.* 3:21. doi: 10.3389/neuro.08.021.2009

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., and Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 3), 15647–15654.

Huettel, S. A. (2010). Ten challenges for decision neuroscience. *Front. Neurosci.* 4:171. doi: 10.3389/fnins.2010.00171

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263.

Knutson, B., Adams, C. M., Fong, G. W., and Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.

Martin-Soelch, C., Linthicum, J., and Ernst, M. (2007). Appetitive conditioning: neural bases and implications for psychopathology. *Neurosci. Biobehav. Rev.* 31, 426–440.

O'Doherty, J. P., Buchanan, T. W., Seymour, B., and Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49, 157–166.

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.

Pagnoni, G., Zink, C. F., Montague, P. R., and Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nat. Neurosci.* 5, 97–98.

Payne, J. W. (2005). It is whether you win or lose: the importance of the overall probabilities of winning or losing in risky choice. *J. Risk Uncertain.* 30, 5–19.

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.

Rangel, A., and Clithero, J. A. (2012). Value normalization in decision making: theory and evidence. *Curr. Opin. Neurobiol.* 22, 970–981.

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178.

Salamone, J. D. (1994). The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. *Behav. Brain Res.* 61, 117–133.

Schultz, W. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831.

Seymour, B., O'Doherty, J. P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K., et al. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8, 1234–1240.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F., and Huettel, S. A. (2009). Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron* 62, 593–602.