Orbitofrontal cortex as a cognitive map of task space: implications for reversal learning and extinction

Robert C. Wilson, Yuji K. Takahashi, Geoffrey Schoenbaum and Yael Niv

Orbitofrontal cortex (OFC) has long been known to play an important role in learning and decision making. However, the exact nature of that role is poorly understood. Recently we have proposed that OFC plays a role in encoding the state-space for reinforcement learning [Takahashi et al. 2011]. This state-space captures the underlying structure of the task, the states, actions and the links between them. The states themselves can be both stimulus-bound - for example, Pavlov's dogs might have a state for the ringing of a bell - or complex, involving both stimulus-bound information and internal information such as the memory of one's last action. According to our theory, OFC encodes these complex states. Without it, animals are limited to stimulus bound representations, and are unable to learn complex tasks.

Here we use our theory to model the classic effects of OFC lesions on reversal learning and extinction. Our theory provides a parsimonious explanation of these deficits and makes a number of experimental predictions about post-extinction phenomena.

We model normal reversal learning using a two state model in which the states are defined by the last action and its outcome. Such a state-space allows rapid reversal learning as the model can effectively learn, and alternate between, two behavioral strategies. We model OFC lesions by assuming that these memory dependent states are no longer distinguishable, and the animal can only use a single state. In this case, initial discrimination learning proceeds normally, but reversal learning is impaired as previously acquired action preferences must be reversed.

We modeled extinction in a similar manner. Normal animals were assumed to use states that are based on the last choice and last outcome. Thus when a previously rewarded stimulus is no longer rewarded in extinction, the model transitions into a new state and rather than unlearning, the apparent "extinction" is actually new learning of action preferences in the new state. Again, with a lesioned OFC, only one state is available and extinction proceeds more slowly as it necessitates unlearning of the previous associations.

This account makes important predictions about post-extinction phenomena. Specifically that OFC lesioned animals will not show renewal, reinstatement or spontaneous recovery as extinction effectively erased their original learning [see also Monfils et al., 2010]. Similarly, we predict that reacquisition of a previously conditioned response will proceed more slowly than initial learning in contrast to the rapid reacquisition seen in normal animals.

Takahashi et al. (2011) Nature Neuroscience, 14, 1590
Monfils et al. (2009) Science, *324*, 951