



## Review

## Learning theory: A driving force in understanding orbitofrontal function

Michael A. McDannald<sup>a,\*</sup>, Joshua L. Jones<sup>b</sup>, Yuji K. Takahashi<sup>a</sup>, Geoffrey Schoenbaum<sup>a,\*</sup><sup>a</sup> National Institute on Drug Abuse, Intramural Research Program, Baltimore, MD, United States<sup>b</sup> University of Maryland, School of Medicine, Baltimore, MD, United States

## ARTICLE INFO

## Article history:

Available online 14 June 2013

## Keywords:

Learning theory  
 Orbitofrontal  
 Reversal  
 Devaluation  
 Overexpectation  
 Sensory preconditioning  
 Blocking  
 Transfer

## ABSTRACT

Since it was demonstrated the orbitofrontal cortex (OFC) is critical to reversal learning, there has been considerable interest in specifying its role in flexible, outcome-guided behavior. Behavioral paradigms from the learning theory tradition, such as outcome devaluation, blocking, Pavlovian to instrumental transfer, and overexpectation have been a driving force in this research. The use of these procedures has revealed OFC's unique role in forming and integrating information about specific features of events and outcomes to drive behavior and learning. These studies highlight the power and importance of learning theory principles in guiding neuroscience research.

Published by Elsevier Inc.

Perhaps the first clear demonstrations of the OFC's role in flexible, outcome-guided behavior and learning came from studies of reversal learning. In reversal studies, subjects must first discriminate a cue that predicts a rewarding outcome from another cue that predicts an aversive outcome or nothing. Discrimination is achieved when subjects respond to the rewarding cue, but withhold responding to the aversive cue. Next, the cue-outcome relationships are reversed. The previously rewarded cue now leads to the aversive outcome and the previously aversive cue now leads to the rewarding outcome. Reversal learning is achieved when behavioral responding reflects the new cue-outcome relationships. It has been demonstrated in a wide variety of species and settings that the OFC is critical for rapid reversal learning (Chudasama & Robbins, 2003; Dias, Robbins, & Roberts, 1996; Izquierdo, Suda, & Murray, 2004; Jones & Mishkin, 1972; Schoenbaum, Nugent, Saddoris, & Setlow, 2002; Teitelbaum, 1964). This contribution is specific to the reversal, as impairing OFC function does not typically impair acquisition of the initial discrimination. The OFC is then not necessary for learning, per se, but is necessary for learning in the face of changing cue-outcome relationships.

At first glance reversal learning appears simple: subjects learn two cue-outcome relationships, these relationships are reversed, and subject's behavior reflects the reversal. A closer look reveals the simultaneous and concurrent changes taking place during reversal, which make pinpointing the contribution of the OFC difficult. In the initial discrimination subjects learn many associations

between cues, responses and outcomes. These associations range from simple, habit-like associations to more cognitive associations by which cues contain information about predicted value and/or sensory features of outcomes, all of which may influence responding. These associations may be rewarding or aversive, adding the dimension of value. Reversal does not consist of changing only one of these associations, but all of them... simultaneously. This requires learning, which invokes the need to signal accurate predictions, in order to engage error and attentional mechanisms to drive learning. Reversal also introduces the concurrent need for behavioral inhibition and engagement. Subjects must withhold a previously learned response while in the same setting engage a previously withheld response. Because reversal performance is assessed in a single format, wherein several processes are concurrently happening, it is difficult to distinguish deficits in these learning mechanisms from deficits in performance.

In some ways the difficulty in interpreting reversal data may account for its popularity. Because reversal learning consists of simultaneous changes in a variety of factors it can provide support for many potential theories of orbitofrontal and prefrontal function. For example, response theories have long claimed the role of the prefrontal cortex – and the OFC in particular – is to inhibit previously learned or prepotent responses (Eagle et al., 2008; Ferrier, 1876; Fineberg et al., 2010; Izquierdo & Jentsch, 2012; Jentsch & Taylor, 1999; Jones & Mishkin, 1972; Man, Clarke, & Roberts, 2009). Value theories generally claim the OFC is necessary for the use or assignment of value (Levy & Glimcher, 2011, 2012; Padoa-Schioppa, 2011; Padoa-Schioppa & Assad, 2006; Plassmann, O'Doherty, & Rangel, 2007). This is a common currency in which incommensurable goods (apples and oranges) may be compared.

\* Corresponding authors.

E-mail addresses: [michael.mcdannald@nih.gov](mailto:michael.mcdannald@nih.gov) (M.A. McDannald), [g\\_scho002@gmail.com](mailto:g_scho002@gmail.com) (G. Schoenbaum).

Specific-outcome theories state the OFC is necessary for predicting and integrating specific features of upcoming events, potentially independent of value (Clark, Hollon, & Phillips, 2012; Delamater, 2007; McDannald et al., 2012; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009). All of the above hypotheses claim reversal deficits as evidence. Yet while reversal learning was an excellent task to begin investigation of OFC function, this task would not be able to pinpoint a precise role for the OFC in flexible behavior.

## 1. Outcome devaluation

A giant leap in our understanding of OFC function has come from studies using outcome devaluation (Colwill & Rescorla, 1985a,b; Holland & Rescorla, 1975; Holland & Straub, 1979). Now known by behavioral neuroscientists, economists and clinicians alike, outcome devaluation has become the gold standard in assessing whether learned behavior is guided by the current value of the outcome. In a typical procedure, subjects first learn that a cue or action leads to a particular reward, for example, a piece of chocolate. Learning is demonstrated when subjects approach the location of chocolate during the predictive cue or perform actions leading to chocolate. Next, the value of the reward is reduced. In the case of chocolate this is achieved either by unlimited consumption, producing selective satiety, or through pairing it's consumption with nausea, forming a conditioned taste aversion. The critical test comes when the cue is presented or action made available in the absence of reward. During such tests, subjects spontaneously decrease cue-evoked responding or actions, reflecting the current value of the reward. Successful outcome devaluation requires two key capacities. First, in initial learning, cues or actions must come to represent detailed information about the rewards they predict. Second, this information must be immediately integrated with current information about that reward's value, so that responding can be adjusted before the reward is ever encountered.

To say this has been a transformative finding for the field of behavioral neuroscience would be a gross understatement. Outcome devaluation procedures have provided a concrete and reliable means of assessing outcome-guided behavior and its underlying neural circuitry. Such studies have repeatedly identified a critical role for the OFC. Lesions or transient inactivation of OFC in rodents (Gallagher, McMahan, & Schoenbaum, 1999; Pickens, Saddoris, Gallagher, & Holland, 2005; Pickens et al., 2003) and macaques (Izquierdo & Murray, 2004, 2010; Izquierdo, Suda, & Murray, 2004; Machado & Bachevalier, 2007; West, Desjardin, Gale, & Malkova, 2011) impair performance in outcome devaluation. Importantly, the deficits observed in these studies are not due to general impairments in learning. Interfering with OFC function typically does not impair acquisition of either Pavlovian or instrumental responding in acquisition. Further, the OFC is not necessary for choosing between valued and devalued rewards when they are present. The OFC is specifically necessary for integrating information about reward-predictive cues and the new value of those rewards. Consistent with work from rodents and macaques, human neuroimaging has found that BOLD signals in the OFC reflect both a cues relationship with an outcome and this BOLD response is decreased in outcome devaluation (Gottfried, O'Doherty, & Dolan, 2003). Thus, across a range of species the OFC is both necessary for and encodes information relevant to outcome devaluation procedures (Murray, O'Doherty, & Schoenbaum, 2007).

Despite the requirement that specific, cue-outcome associations be integrated to successfully decrease responding, the OFC's role in outcome devaluation has been cited as evidence for roles in both value and response inhibition. These descriptions have some merit. Successful demonstration of outcome devaluation involves knowl-

edge of a reward's current value as well as the ability to decrease responding in the face of this knowledge. Fortunately, studies of learning theory have led to a wealth of behavioral procedures that isolate contributions of specific-outcome, value and response inhibition processes to learning and behavior. In the following sections we describe a handful of these procedures that have been especially useful.

## 2. Unblocking and transfer

Studies of unblocking and Pavlovian-to-instrumental transfer (PIT) have been highly valuable in isolating learning and behavior driven by general value versus specific outcomes. Blocking and unblocking are at the heart of modern learning theory. In a typical blocking study subjects first learn that a cue predicts reward ( $A \rightarrow \text{Food}$ ). Next a novel cue is added and followed by the same reward ( $AX \rightarrow \text{Food}$ ). The critical test comes when reward responding to the added cue X is tested in isolation. When this is done subjects show little or no responding to X, compared to another group that did not receive  $A \rightarrow \text{Food}$  learning (Kamin, 1969). Prior training of A 'blocked' learning to X.

In a popular explanation, blocking of learning to the added cue occurs because the previously trained cue fully and accurately predicts all aspects of the reward. If blocking is about information then changing aspects of the reward should result in 'unblocking' or learning to the added cue. This is because the added cue is now in a position to provide new information about the reward. Studies by Dr. Peter Holland and colleagues have repeatedly demonstrated such unblocking effects (Holland, 1984, 1988; Holland & Gallagher, 1993; Holland & Kenmuir, 2005). In one form of unblocking, rats receive extensive training of a cue and a specific amount of reward. Once well-learned, a novel cue is added and the amount of reward is increased (Holland, 1984). Because the novel cue now predicts a larger reward, unblocking occurs. Importantly, this learning need not rely on specific-outcome properties of reward. This form of unblocking might be said to require information *about value*. Rescorla hypothesized that unblocking ought to occur even when there is no explicit change in value but only a change in the quality, or sensory-feature of rewards. To demonstrate this Rescorla first trained rats that two separate cues predicted two discriminable, yet equally valued rewards ( $A \rightarrow O1$ ,  $B \rightarrow O2$ ). Next novel cues were compounded with the initially trained cues and the reward changed in one condition ( $AX \rightarrow O1$ ,  $BY \rightarrow O1$ ). Using a transfer test he reported blocking of learning to the added cue X, that provided no information about reward but substantial unblocking to cue Y, which selectively provided information about specific features of reward. Thus, learning may occur in situations in which value is held constant, but specific features of rewards are altered (Rescorla, 1999).

The unblocking data reported by Rescorla are consistent with results of selective Pavlovian-to-instrumental transfer (PIT) routinely reported by Dr. Bernard Balleine and colleagues (Corbit & Balleine, 2005, 2011; Corbit, Janak, & Balleine, 2007). In these studies subjects are taught that two separate instrumental responses lead to two discriminable yet equally preferred rewards ( $R1 \rightarrow O1$ ,  $R2 \rightarrow O2$ ). At the same time subjects also learn that two environmental cues lead to these different rewards ( $A \rightarrow O1$ ,  $B \rightarrow O2$ ). Once both instrumental and Pavlovian learning is established, a transfer test is given in which the cues are presented over the top of instrumental responding but no rewards are present. During presentation of a specific cue, subjects will typically perform actions leading to that same reward much more than action leading to the alternate reward. For example, cue A selectively augments responding for R1, while B selectively augments responding for R2. These results demonstrate that in initial learning both the

Download English Version:

<https://daneshyari.com/en/article/936559>

Download Persian Version:

<https://daneshyari.com/article/936559>

[Daneshyari.com](https://daneshyari.com)