Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Neurobiology

Two tales of how expectation of reward modulates behavior

Long Ding[1] and David J Perkel[2,3]

Expectation of reward modulates many types of behaviors. Here we highlight two lines of research on reward-modulated perceptual decision making in primates and social context-modulated singing in songbirds, respectively. These two seemingly distinct behaviors are both known to involve cortico-basal ganglia-thalamic circuits. The underlying computations may be conceptualized using a simple, common framework. We summarize and compare our current knowledge of the two fields to motivate new experiments for each field, with the goal of finding general principles for how the brain implements reward-modulated behavior.

Addresses
[1] Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104, USA
[2] Department of Biology, University of Washington, Seattle, WA 98195-1800, USA
[3] Department of Otolaryngology, University of Washington, Seattle, WA 98195-6515, USA

Corresponding author: Perkel, David J (perkel@uw.edu)

Current Opinion in Neurobiology 2014, **29**:142–147

This review comes from a themed issue on **Neuromodulation**

Edited by **David McCormick** and **Michael P Nusbaum**

For a complete overview see the Issue and the Editorial

Available online 23rd July 2014

http://dx.doi.org/10.1016/j.conb.2014.07.011

0959-4388/© 2014 Elsevier Ltd. All right reserved.

Introduction

In goal-directed behaviors, expectations of reinforcement are an integral part of the definition of goals. For example, for animals foraging for food, it is often beneficial to have some knowledge of how much food reward may be expected of currently available food sources and then act accordingly. For animals trying to attract a mate, it is likely beneficial to appreciate what the potential mates prefer, which may be thought of as an expectation of successful mate attraction given certain attributes of the mate-attracting behavior. In this review, we highlight two lines of research that deal with seemingly distinct types of goal-directed behaviors: reward-modulated perceptual decision making in primates and social context-dependent singing behavior in songbirds. We suggest that, despite the apparent differences in species, behavioral modality and the general premise of goals, these two lines of research share many similarities and could inform each other to advance our understanding of the neural basis of how expectation of reinforcement modulates behavior in general.

In the following sections, we first describe the behaviors of interest. We then introduce a common framework to describe the general computations that are needed for these two types of behaviors. Next we describe the cortico-basal ganglia-thalamo-cortical circuits that are (likely) involved in mediating these behaviors, drawing evidence from fMRI studies in humans and neurophysiological studies in monkeys and songbirds. Finally, we put forth some open questions for future studies, guided by a comparison of existing knowledge between the two lines of research.

The behaviors and general framework

Reward-modulated perceptual decision making

To illustrate some key properties of reward-modulated perceptual decision behavior, we focus on examples of adding explicit reward expectation manipulations to a well-known perceptual decision task, the visual motion direction discrimination task [1,2]. On a basic motion discrimination task, a subject sees a field of moving dots on a screen and needs to report the global motion of these dots. On each trial, there are two possible directions (i.e. two possible choices) and a proportion of the dots move in one direction while the other dots move in random directions. This proportion specifies the motion stimulus strength and influences the difficulty of the perceptual judgment: if the motion strength is high, that is, most dots are moving coherently in one direction with only a few dots moving randomly, it is easy to tell that the global motion is in that direction; conversely, if the motion strength is low, that is, only a few dots are moving coherently in one direction in a field of randomly moving dots, it is hard to determine the true global motion. Another factor influencing the task difficulty is the viewing time: when the viewing time is controlled by the experimenter, the longer the viewing time, the more likely the subject is to make a correct decision; when the viewing time is controlled by the subject, trials with lower motion strength tend to have longer viewing times (i.e. longer response times).

With the basic task design, reward expectation is not explicitly manipulated, except that all correct choices earn a fixed reward and error choices are not rewarded and/or incur a fixed penalty (e.g. time out). One way to explicitly manipulate reward expectation is thus to make the reward size associated with the two choices unequal. In humans and monkeys, this kind of manipulation has robust effects on behavioral performance [3,4,5••,6–8]. These effects include: firstly, subjects are more likely to pick the choice associated with the larger reward; such a choice bias is more prominent on trials with low motion strength, while minimal on trials with very high motion strength; secondly, when subjects control the viewing time, for the same motion strength, response time is shorter for the choice associated with the larger reward; such response time bias also depends on motion strength, with the largest bias occurring on trials with low motion strength.

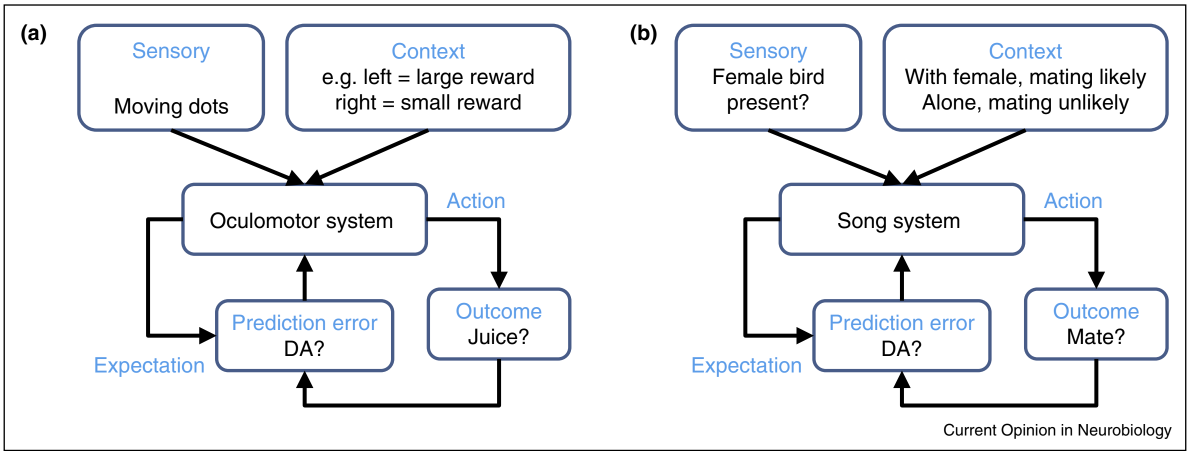
To understand performance on this task, consider a simple framework illustrated in Figure 1a. At the onset of the decision process, the brain combines sensory information related to the moving dots and contextual information related to the reward contingency, and produces a categorical 'left' or 'right' choice. Using concepts from the perception literature, a perceptual categorization process can generally be thought of as having two components: sensory representation, reflecting the conversion of sensory inputs into task-relevant evidence, and readout, reflecting the conversion of the task-relevant evidence into a categorical decision [9]. *A priori*, the contextual information may influence both components. As an example of reward modulation of sensory representation, when the 'left' choice is paired with the larger reward, there may appear to be more leftward-moving dots than are actually present, biasing the evidence toward the 'left' choice. As an example of reward modulation of readout, in the same left-choice-with-larger-reward scenario, the subject may require observing a total of 100 rightward-moving dots before making a 'right' choice, but only 20 leftward-moving dots to make a 'left' choice, thereby reading out the same evidence with a leftward bias. Based on model fitting using 'accumulate-to-bound' models, reward-biased readout better recapitulates observed behavioral effects on reward-biased motion discrimination tasks [3,5••,6] and this seems to be a common finding for many unequal-payoff perceptual decision tasks [7,10–13]. During decision formation, the brain may also generate predictions of how much reward is expected. A comparison between these predictions and actual reward received can then instruct changes of the categorical decision-formation process to improve future decisions.

Social context-dependent singing

Another example in which behavior is altered depending on the context, perhaps including the expectation of reward, is in avian song. Songbirds use song for courtship, territory defense and individual recognition. In zebra finches, a young male learns a tutor song through a trial-and-error process [14], in which the young bird's own songs are compared against a memorized template [15]. This learning occurs in an unsupervised manner; no external rewards are needed to guide progress or indicate success. Over a period of ~3 months, songs show gradual improvement until a good match is achieved [14]. During this period, song variability also decreases such that the adult song becomes acoustically stereotyped, save for a small amount of residual variability.

This residual variability is modulated by social context, potentially in anticipation of social reward. When a male sings as part of female-directed courtship (directed song), trial-to-trial song variability drops to about half of the value measured when the bird sings alone (undirected song) [16–18]. Females care about this distinction, spending more time near a speaker playing the mate's directed song than one playing undirected song [19]. One practical

Figure 1



A common conceptual framework for reward-modulated perceptual decision making **(a)** and social context-modulated vocal behavior **(b)**. Blue text indicates the conceptual functions (identical for the two panels) and black text provides behavior-specific information.

Download English Version:

https://daneshyari.com/en/article/4334206

Download Persian Version:

https://daneshyari.com/article/4334206

Daneshyari.com