

# **Dopaminergic and Non-Dopaminergic Value Systems in Conditioning and Outcome-Specific Revaluation**

Mark R. Dranias, Stephen Grossberg, and Daniel Bullock<sup>1</sup>

Department of Cognitive and Neural Systems

Center for Adaptive Systems

and

Center of Excellence for Learning in Education, Science, and Technology

Boston University

677 Beacon Street, Boston, MA 02215

Technical Report CAS/CNS TR-2007-020

Submitted: December 6, 2007

*All correspondence should be addressed to*

Professor Stephen Grossberg

Department of Cognitive and Neural Systems

Boston University

677 Beacon Street

Boston, MA 02215

Phone: 617-353-7858

Fax: 617-353-7755

Email: [steve@bu.edu](mailto:steve@bu.edu)

<sup>1</sup>Authorship is in rotated alphabetical order.

## ABSTRACT

Animals are motivated to choose environmental options that can best satisfy current needs. To explain such choices, this paper introduces the MOTIVATOR (Matching Objects To Internal VAlues Triggers Option Revaluations) neural model. MOTIVATOR describes cognitive-emotional interactions between higher-order sensory cortices and an evaluative neuraxis composed of the hypothalamus, amygdala, and orbitofrontal cortex. Given a conditioned stimulus (CS), the model amygdala and lateral hypothalamus interact to calculate the expected current value of the subjective outcome that the CS predicts, constrained by the current state of deprivation or satiation. The amygdala relays the expected value information to orbitofrontal cells that receive inputs from anterior inferotemporal cells, and medial orbitofrontal cells that receive inputs from rhinal cortex. The activations of these orbitofrontal cells code the subjective values of objects. These values guide behavioral choices. The model basal ganglia detect errors in CS-specific predictions of the value and timing of rewards. Excitatory inputs from the pedunculopontine nucleus interact with timed inhibitory inputs from model striosomes in the ventral striatum to regulate dopamine burst and dip responses from cells in the substantia nigra pars compacta and ventral tegmental area. Learning in cortical and striatal regions is strongly modulated by dopamine. The model is used to address tasks that examine food-specific satiety, Pavlovian conditioning, reinforcer devaluation, and simultaneous visual discrimination. Model simulations successfully reproduce discharge dynamics of known cell types, including signals that predict saccadic reaction times and CS-dependent changes in systolic blood pressure.

**Section:** Computational and Theoretical Neuroscience.

**Keywords:** amygdala, orbitofrontal cortex, rhinal cortex, lateral hypothalamus, inferotemporal cortex, basal ganglia, conditioning, motivation, devaluation, food-specific satiety, dopamine, cognitive-emotional interactions, decision-making, discrimination learning

## 1. INTRODUCTION

Animal behavior is fundamentally opportunistic. Animals choose actions whose consummatory responses serve their basic biological needs, such as avoidance of damage, regulation of body temperature, and replenishment of energy stores. Many of these needs vary over life cycles, seasons, and days, as do the environmental opportunities for making appropriate consummatory responses. Choosing options that can best satisfy currently pressing needs often requires temporarily ignoring options that, under different subjective conditions, would be evaluated as highly attractive. This may require temporarily ignoring some current needs that would be strong enough to dominate behavioral choices if the animal were relocated to an environment that supported consummatory responses matched to those needs.

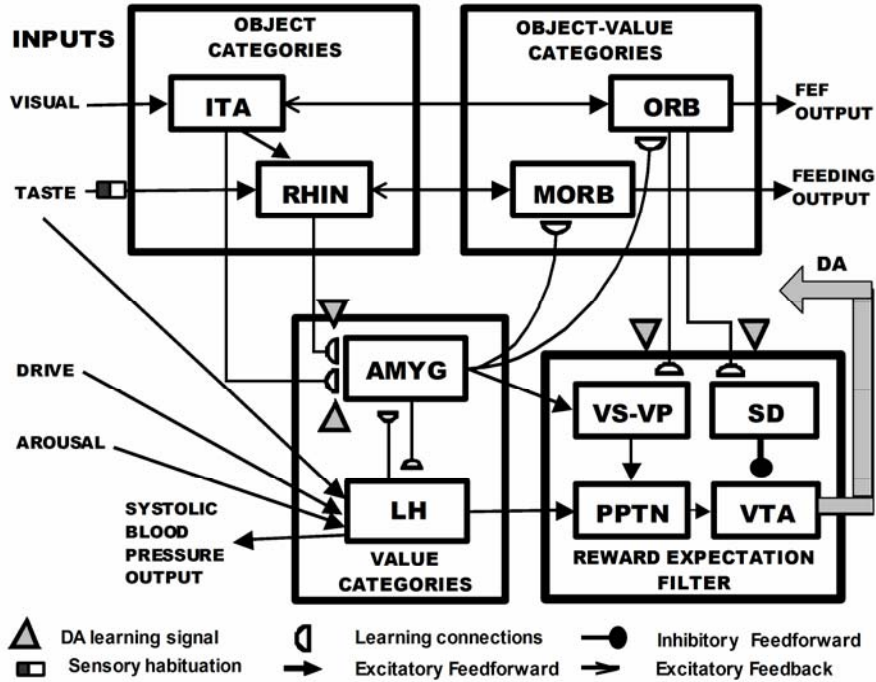
What brain processes allow an animal to use cues to quickly assess the options in its environment and estimate their values relative to the animal's current needs? How are strong needs ignored when the environment affords no opportunity for their satisfaction? How are normally attractive and highly available options ignored for a time after the needs that they consummate have been satisfied? To address such questions, a neural model is proposed and simulated to explain laboratory phenomena such as: the conditioning of cues that predict specific outcomes in a task setting, the automatic revaluation of conditioned stimuli (conditioned reinforcers) following food-specific satiety, and motivational and emotive influences on decision processes, reaction time, response vigor, and blood pressure. The phenomenon of automatic revaluation has only recently been thoroughly investigated and requires additional explanation (Dickinson and Balleine, 2001; Corbit and Balleine, 2005). Revaluation refers to the observation that motivational shifts can alter the vigor of conditioned responses.

Outcome-specific revaluation occurs when shifts in motivation alter conditioned responding in a manner that respects the different reward associations of these responses and how this motivational shift differentially impacts the consumption value of these outcomes (Corbit and Balleine, 2005). Normally, changes in conditioned responding follow the law of effect, and the value of a CS only reflects the experienced value of its associated food reward. However, for first-order and second-order conditioned stimuli, revaluation automatically occurs in an outcome-specific fashion (Corbit and Balleine, 2003, 2005; Hall 2001). The effect is automatic in that changes in the value of rewards impact the vigor of conditioned responding without new CS-US pairings. In contrast, motivational shifts alter the vigor of higher-order conditioned responses in an outcome-specific fashion only after additional training trials have taken place during which the reward is experienced in the new motivational state (Balleine et al., 1995).

---

Abbreviations: AMYG, Amygdala; BG, Basal ganglia; BP, Systolic blood pressure; CS, Conditioned stimulus; DA, Dopamine; FEF, Frontal eye fields; FSS, Food-specific satiety; GUS, Gustatory inputs; ITA, Anterior inferotemporal cortex; LDT, Laterodorsal tegmental nuclei; LH, Lateral hypothalamus; LH\_gus, Gustatory-receptive lateral hypothalamic cells; LH\_in, Drive input receptive lateral hypothalamic cells; LH\_out, Lateral hypothalamic output cells; LTM, Long-term memory; MORB, Medial orbitofrontal cortex; MTM, Medium-term memory; ORB, Orbitofrontal cortex; PIT, Posterior inferotemporal cortex; PPTN, Pedunculopontine nucleus; RHIN, Rhinal cortex; RT, Reaction time; SD, Striosomal delay cells in the ventral striatum; SNc, Substantia nigra pars compacta; STM, Short-term memory; SVD, Simultaneous visual discrimination; TD, Temporal difference; US, Unconditioned stimulus; VIS, Visual inputs; VP, Ventral pallidum; VS, Ventral striatum; VTA, Ventral tegmental area.

Key aspects of these phenomena are explained within a neural circuit that integrates homeostatic, hedonic and emotional information to calculate the current value of conditioned and unconditioned cues. The model serves to detail, contrast, and elaborate the roles of dopaminergic and non-dopaminergic value systems and mechanisms that are engaged by most evaluative tasks, including Pavlovian and operant conditioning (Berridge, 2001; Berridge and Robinson, 1998). These results were reported in preliminary form in Dranias, Bullock and Grossberg (2006, 2007a, 2007b).



**Figure 1: Overview of MOTIVATOR model:** Brain areas in the MOTIVATOR circuit can be divided into four regions that process information about conditioned stimuli (CSs) and unconditioned stimuli (USs). (a) Object Categories represent visual or gustatory inputs, in anterior inferotemporal (ITA) and rhinal (RHIN) cortices; (b) Value Categories represent the value of anticipated outcomes on the basis of hunger and satiety inputs, in amygdala (AMYG) and lateral hypothalamus (LH); (c) Object-Value Categories resolve the value of competing perceptual stimuli in medial (MORB) and lateral (ORB) orbitofrontal cortex; and (d) the Reward Expectation Filter detects the omission or delivery of rewards using a circuit that spans ventral striatum (VS), ventral pallidum (VP), striosomes of the striatum, the pedunculo pontine nucleus (PPTN) and midbrain dopaminergic neurons of the SNc/VTA (substantia nigra pars compacta/ventral tegmental area). The circuit that processes CS-related visual information (ITA, AMYG, ORBL) operates in parallel with a circuit that processes US-related visual and gustatory information (RHIN, AMYG, ORBM). The model captures systematic changes in processing of the same stimuli at different times, due to processes of learned category formation, sensory habituation, satiation or deprivation of particular rewarding outcomes,

CS-US associative learning, and violations of expectations based on learned regularities. Model outputs modulate saccadic choice and reaction time and blood pressure changes.

The **MOTIVATOR (Matching Objects To Internal Values Triggers Option Revaluations)** model focuses on cognitive-emotional processing wherein sensory and cognitive neocortex interacts with an evaluative neuraxis composed of the hypothalamus, amygdala, orbitofrontal cortex, and basal ganglia. An overview of the model, which has been specified as a real-time dynamical system and simulated in Matlab, is shown in Fig. 1. This model unifies and further develops the Cognitive-Emotional-Motor, or CogEM, model of cognitive-emotional learning and performance (Grossberg, 1971, 1972a, 1972b, 1975, 1984; 2000a, 2000b; Grossberg and Gutowski, 1987; Grossberg and Levine, 1987; Grossberg, Levine and Schmajuk, 1987; Grossberg and Merrill, 1992; Grossberg and Schmajuk, 1987) and the TELOS model of how an animal learns to balance reactive vs. planned behaviors through learning based on reward expectation and its disconfirmation (Brown, Bullock, and Grossberg, 1999, 2004). The CogEM model focused on how affective brain regions, such as the lateral hypothalamus and amygdala, interact with sensory and cognitive areas, such as inferotemporal cortex and orbitofrontal cortex. The TELOS model focused on how the basal ganglia regulate attention and reinforcement-based learning in thalamocortical systems. The current model proposes how both amygdala and basal-ganglia processes interact to control reward-based processes.

In the MOTIVATOR model, visual inputs activate view-invariant representations of visual objects in the anterior inferotemporal cortex (ITA). Gustatory cortex relays the taste properties salty, sweet, umami, and fatty to rhinal cortex (RHIN) and to gustatory-responsive lateral hypothalamic cells (LH\_gus). RHIN cells also receive ITA inputs, and can thereby code gustatory-visual properties of food rewards. Endogenous drive and arousal inputs project to lateral hypothalamic input cells (LH\_in). LH\_in cells represent the homeostatic state of the animal by reporting fat, salt, amino acid, and sugar levels. LH\_gus cells correlate gustatory tastes with corresponding homeostatic features and excite lateral hypothalamic output cells (LH\_out), which project to amygdala (AMYG) cells that categorize LH\_out states. The LH-AMYG network computes the net subjective outcome associated with a consummatory act. It thereby defines a neural representation of US (unconditioned stimulus) reward value. Because the AMYG also receives conditionable CS-activated signals from ITA and RHIN, it can mediate CS-US learning. Given a CS, the AMYG and LH interact to calculate the expected current value of the subjective outcome that the CS predicts, given the current state of deprivation or satiation for that outcome. The AMYG relays the expected value information to ITA-recipient orbitofrontal (ORB) and RHIN-recipient medial orbitofrontal (MORB) cells, whose activations code the relative subjective values of objects. These values guide behavioral choices.

The model basal ganglia (BG) detect errors in CS-specific predictions of the value and timing of rewards. Striosomes (SD) of the ventral striatum (VS) prevent predicted rewards from generating SNc/VTA responses by inhibiting dopamine cells in the SNc/VTA with adaptively timed signals (Fig. 1). Inputs from the LH\_gus and the ventral striatum (VS) excite the pedunculopontine nucleus (PPTN/LDT) whenever a conditioned (CS) or unconditioned (US) rewarding cue occurs. Cells in the PPTN/LDT, in turn, excite dopamine cells in the SNc/VTA. When inhibitory signals from the SD and excitatory signals from the PPTN/LDT mismatch, a dopamine dip or dopamine burst may occur. A dopamine burst occurs in the SNc/VTA when an unexpected rewarding CS or US is presented. When an unexpected rewarding cue is presented, SD cells are unable to relay anticipatory inhibitory signals to the SNc/VTA and reward-related

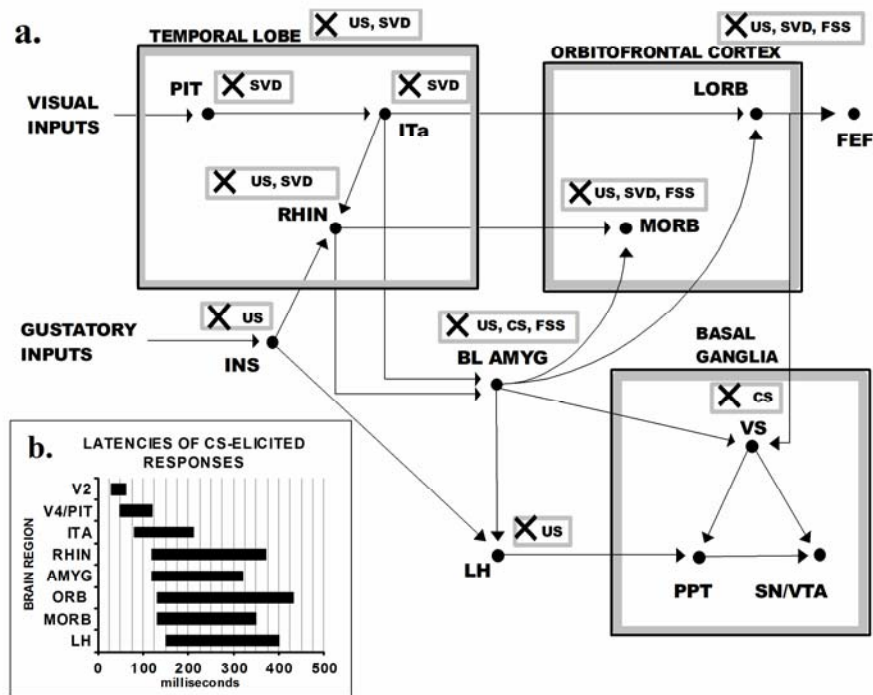
excitation is relayed from the PPTN/LDT to dopaminergic cells in the SNc/VTA, eliciting a dopamine burst. When an expected reward is omitted, a dopamine dip occurs. In this case, a rewarding CS is presented and SD cells send an adaptively timed inhibitory input to the SNc/VTA at the expected time of reward. When US presentation is omitted, dopaminergic SNc/VTA cells never receive a reward-related excitatory signal from the PPTN/LDT and are instead transiently suppressed by inhibitory signals from the SD. Model simulations reproduce discharge dynamics of known cell types, including signals that predict saccadic reaction times and CS-dependent changes in systolic blood pressure. Learning in cortical and striatal regions is strongly modulated by dopamine, whereas learning between the AMYG and LH\_out cells is not.

**1.1 Task Selection.** Three basic tasks are neurally explained and simulated by the model: a ‘free reward’ or unconditioned stimulus (US) learning task, a conditioned stimulus (CS) to US (hereafter CS-US) associative learning task, and a simultaneous visual discrimination (SVD) task that involves dual associative learning (CS1-US1 and CS2-US2). Variants of these three tasks were also simulated to explain data about food-specific satiety (FSS), devaluation, extinction, and reversal learning.

In the US learning task, the model learns to associate stimulus features of a food reward with internal representations of the value and outcomes elicited by that food. The model simulates the US learning task of Nakamura and Ono (1986) because the literature related to this task contains electrophysiological data against which simulated neural dynamics can be compared. Three simulations of the US task were performed. The first simulation shows learning in the model amygdala of US-specific internal representations that encode specific drives and the identity of rewarding food stimuli. The second simulation shows US devaluation and satiety curves. The third simulation demonstrates the specificity of the satiation response. In particular, when a single food is eaten to satiety, it is known that satiation acts in a food-specific fashion by devaluing the consumed food more than other foods. In summary, simulations of the US task demonstrate how cells in the AMYG learn to encode a US-specific motivational representation, devaluation effects, satiation curves, and food-specific satiety. While achieving these functions, the model replicates the dynamics of experimentally observed neurophysiological cell types in the AMYG, LH, ORB, PPTN/LDT, SNc/VTA, and VS.

The CS-US learning task reported by Ono et al. (1986a) was also selected for simulation, both because it reported electrophysiological data and because Pavlovian conditioning has been the task of choice for laboratories investigating which neural circuits underlie the automatic revaluation of conditioned stimuli (Hatfield et al., 1996; Gallagher et al., 1999; Balleine, 2005). The primary computational issues address how conditioned cue representations associate with US-specific outcome representations and track the current value of prospective outcomes.

The simultaneous visual discrimination (SVD) task allows an animal to choose the more preferred of two simultaneously presented visual conditioned stimuli (CSs). It incorporates properties of reinforcement learning and decision-making, and has a lengthy history of study in behavioral neuroscience (Easton and Gaffan, 2000; Murray and Mishkin, 1998; Voytko, 1985). Simulations of brain processes sufficient to perform the SVD task demonstrate how value is assigned and choices made between competing stimuli. This task provides an opportunity to study how different valuation mechanisms elicit changes in cue preference. Dopamine-dependent conditioned reversals of CS preference were simulated and compared with reversals of CS preference that arise from changes in neural representations of organismic needs.



**Figure 2: Brain lesions and neural response timing:** Part 2a shows feedforward connections among the major brain regions of the model (see Table 1) and 2b shows representative neural response latencies to visual stimuli. In 2a, brain regions for which there is evidence of a disruption in tasks substantially similar to the **US**, **CS**, **SVD**, or **FSS**-related tasks are marked with a bold 'X' together with a task designator. **US-related** tasks include: Lidocaine inactivation during **US** task, sham-feeding, general feeding, or food preference tests. **CS-related** tasks include: Lidocaine inactivation during **CS** task, Pavlovian conditioning (either autoshaping or place), and disruptions of extinction or reversal in these tasks. **SVD-related** tasks include: The **SVD** task and **SVD** reversal. **FSS-related** tasks include **CS-FSS**, **SVD-FSS** tasks, and **CS** revaluation tasks where the food is poisoned with LiCl. Part 2b shows latencies in monkeys for the initial response to a visual conditioned stimulus by neurons in various brain regions. V2 latencies (Luck et al., 1997); V4/PIT latencies (Ibid.); ITA latencies (Wachsmuth et al., 1994; Liu and Richmond, 2000); RHIN latencies (Suzuki et al., 1997); AMYG latencies (Wilson and Rolls, 1993); ORB latencies (Rolls, 2000; Tremblay and Schultz, 2000a); LH latencies (Rolls et al., 1979).

**1.2 Neurobiological Basis of the Model.** A neural network involving higher-order sensory cortices, the hypothalamus, the amygdala, the basal ganglia, and the orbitofrontal cortex enables evaluative and emotional processing of cues that predict appetitive and aversive outcomes. Fig. 2a summarizes evidence from lesion studies that these brain regions play a critical role in evaluation during the **US** learning, **CS-US** learning, and **SVD** tasks. Lesions of the posterior inferotemporal (PIT) cortex impair **SVD** tasks (Voytko, 1986). RHIN lesions impair **US** tasks (Parker and Gaffan, 1998) and **SVD** tasks (Buffalo et al., 1999). Gustatory insula (INS) lesions impair **US** tasks (Dunn and Everitt, 1988). ITA lesions impair **SVD** tasks (Voytko, 1986).

Temporal lobe lesions impair US tasks (Barefoot et al., 2000) and SVD tasks (Voytko, 1986). AMYG lesions impair US tasks (Gaffan, 1994), CS tasks (Kantak et al., 2001; Cardinal et al., 2002), and FSS tasks (Murray et al., 1996; Malkova et al., 1997). LH lesions impair US tasks (Nakamura et al., 1987; Bernardis and Bellinger, 1996; Touzani and Sclafani, 2002). Orbitofrontal cortex (MORB and ORB) lesions impair US tasks (Baylis and Gaffan, 1991), SVD tasks (Easton and Gaffan, 2000) and FSS tasks (Baxter et al., 2000; Cardinal et al., 2002). Ventral Striatum lesions impair CS tasks (Cardinal et al., 2002; Schoenbaum et al., 2003). Table 1 summarizes pathway-tracing studies that verify the existence of significant links between the brain regions identified in Figs. 1 and 2a.

#### *Anatomical Connections of Model Brain Regions*

| Anatomical Connection  | References   |   |
|------------------------|--|---|
|                        | Monkey   | Rat   |
| <b>PIT to ITa</b>      | Suzuki, et al. (2000)                                | Shi & Cassell (1997)                            |
| <b>ITa to ORB</b>      | Barbas (1993, 2000)                                  | Reep, et al. (1996)                             |
| <b>ORB to ITa</b>      | Rempel-Clower & Barbas (2000)                        | -   |
| <b>ORB to FEF</b>      | Barbas (1992); Carmichael & Price (1995)             | -   |
| <b>ITa to RHIN</b>     | Saleem & Tanaka (1996)                               | Burwell & Amaral (1998)                         |
| <b>GUS to RHIN</b>     | Insausti, Amaral, & Cowan (1987)                     | Burwell & Amaral (1998)                         |
| <b>MORB to RHIN</b>    | Barbas (1993); Barbas, et al. (1999)                 | Reep et al. (1996)                              |
| <b>RHIN to BL AMYG</b> | Rempel-Clower & Barbas (2000)                        | Burwell & Amaral, (1998)                        |
| <b>ITa to BL AMYG</b>  | Ghashghaei & Barbas (2002)                           | McDonald (1998)                                 |
| <b>BLAMYG to MORB</b>  | Ghashghaei & Barbas (2002)                           | McDonald, (1998)                                |
| <b>BL AMYG to ORB</b>  | Amaral & Price (1984);<br>Ghashghaei & Barbas (2002) | Reep et al. (1996); Ongur & Price (2000)        |
| <b>ORB to VS, SD</b>   | Ferry, et al. (2000)                                 | Ongur & Price (2000)                            |
| <b>BL AMYG to VS</b>   | Friedman, et al. (2002)                              | Swanson (2000)                                  |
| <b>GUS to LH</b>       | Sewards & Sewards (2001)                             | Risold, et al. (1997)                           |
| <b>BLAMYG to LH</b>    | Barbas, et al. (2003); Price (2003)                  | DeFalco et al. (2001); Petrovich, et al. (2001) |
| <b>LH to BLAMYG</b>    | Russchen (1986); Ghashghaei & Barbas (2002)          | Peyron et al. (1998); Sah, et al. (2003)        |

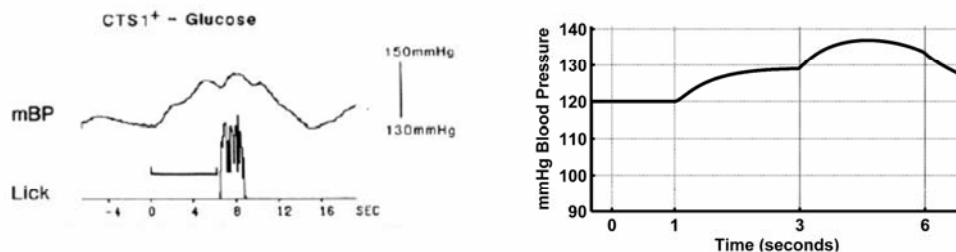


|                            |  |                                     |
|----------------------------|--|-------------------------------------|
| <b>VS to VP/VP to PPTN</b> | Haber, et al. (1993); Spooren, et al. (1996) | Semba & Fibiger (1992); Zahm (2000) |
| <b>VS to VTA/VP to VTA</b> | Haber, et al. (2000)                         | Zahm (2000)                         |
| <b>LH to PPTN</b>          | Veazey, et al. (1982)                        | Semba & Fibiger (1992)              |
| <b>PPTN to VTA</b>         | Lavoie & Parent (1994)                       | Oakman, et al. (1995)               |

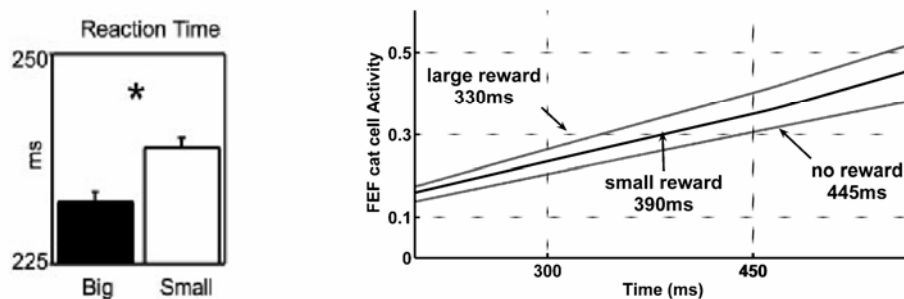
**Table 1**

*Model inputs and outputs.* The neural model in Fig. 1 has four input types: visual, gustatory, drive (e.g., specific hungers and other internal states) and arousal. To illustrate the timing of visual information flow through the system, Fig. 2b gives typical latencies at which cells in different brain regions respond to visual stimuli. Two key outputs of the model are a hypothalamic emotional signal that predicts CS-induced changes in systolic (peak of cardiac cycle) blood pressure (hereafter BP) and an orbitofrontal signal that predicts reaction times (RTs) of voluntary saccadic eye movements. In Fig. 3, experimental results on BP and saccadic RTs are presented alongside model simulations of signals that predict these variables.

**a. Systolic Blood Pressure**



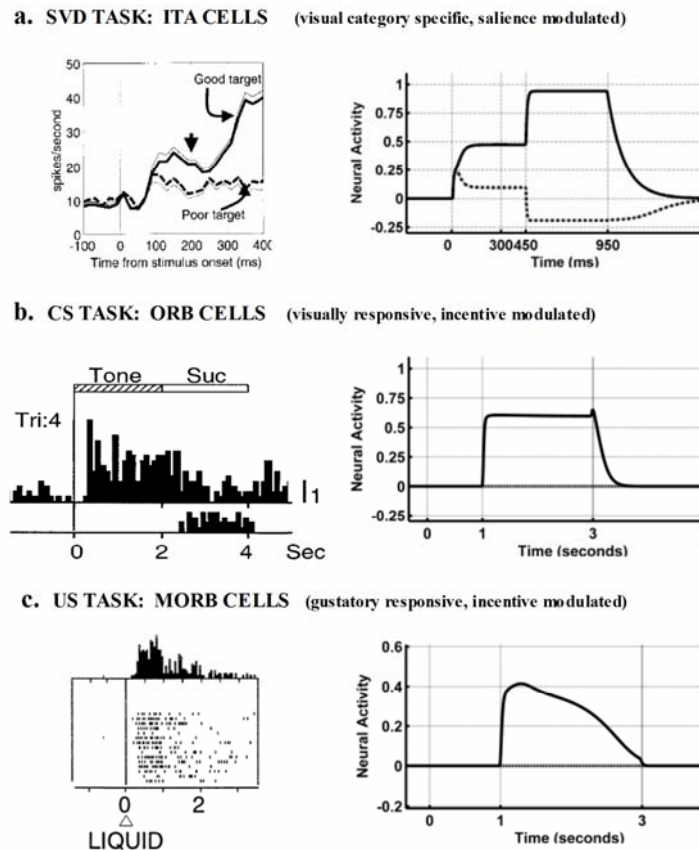
**b. Saccadic Reaction Time**



**Figure 3: Blood pressure and saccade latency:** MOTIVATOR reproduces conditioned and unconditioned changes in systolic blood pressure using drive signals from LH cells (Nakamura et al., 1992; Braesicke et al., 2005). Motivational influences on saccadic reaction time are reproduced using incentive value signals broadcast by orbitofrontal cortex (Lauwereyns et al., 2002; Roesch and Olson, 2003). 3a: Simulated blood pressure output (right) compared with recorded blood pressure (left) during CS Task performance

[Reprinted with permission from Nakamura et al. (1992)]. Small increases in blood pressure follow the presentation of conditioned stimuli, unconditioned stimuli, or the consumption of unconditioned stimuli, but not neutral stimuli (Braesicke, et al 2005; Nakamura et al., 1992). 3b: Simulated saccadic reaction times (right) replicate trends in observed changes in saccadic reaction time [Reprinted with permission from Roesch and Olson (2003)]. For simulated reaction times, large rewards correspond to high hunger drive inputs and small rewards correspond to low hunger drive levels. For experimentally observed reaction times, big and small rewards correspond to the amount of juice given as reinforcement.

*Modeled cell types.* Many neurons have characteristic activation profiles that allow them to be classified as exemplars of a functional cell type (Ono et al., 1986b; Nishijo et al., 1988a; Tremblay and Schultz, 2000b). The activation profiles of eleven neural cell types were simulated by model neurons. Figs. 4 and 5 compare simulated cell activations and electrophysiological discharge profiles for cell types, including ITA (anterior inferotemporal cortex), lateral orbitofrontal cortex (ORB), medial orbitofrontal cortex (MORB), basolateral amygdala (AMYG), lateral hypothalamic output (LH\_out) cells, and lateral hypothalamic gustatory-receptive cells (LH\_gus). Cells in the AMYG discriminate between rewarding and aversive stimuli and are modulated by hunger and satiety (Muramoto et al., 1993; Yan and Scott, 1996). Nishijo, et al (1988a, b) reported that some AMYG cells are multimodal, motivationally-modulated, and respond in a food-specific fashion.



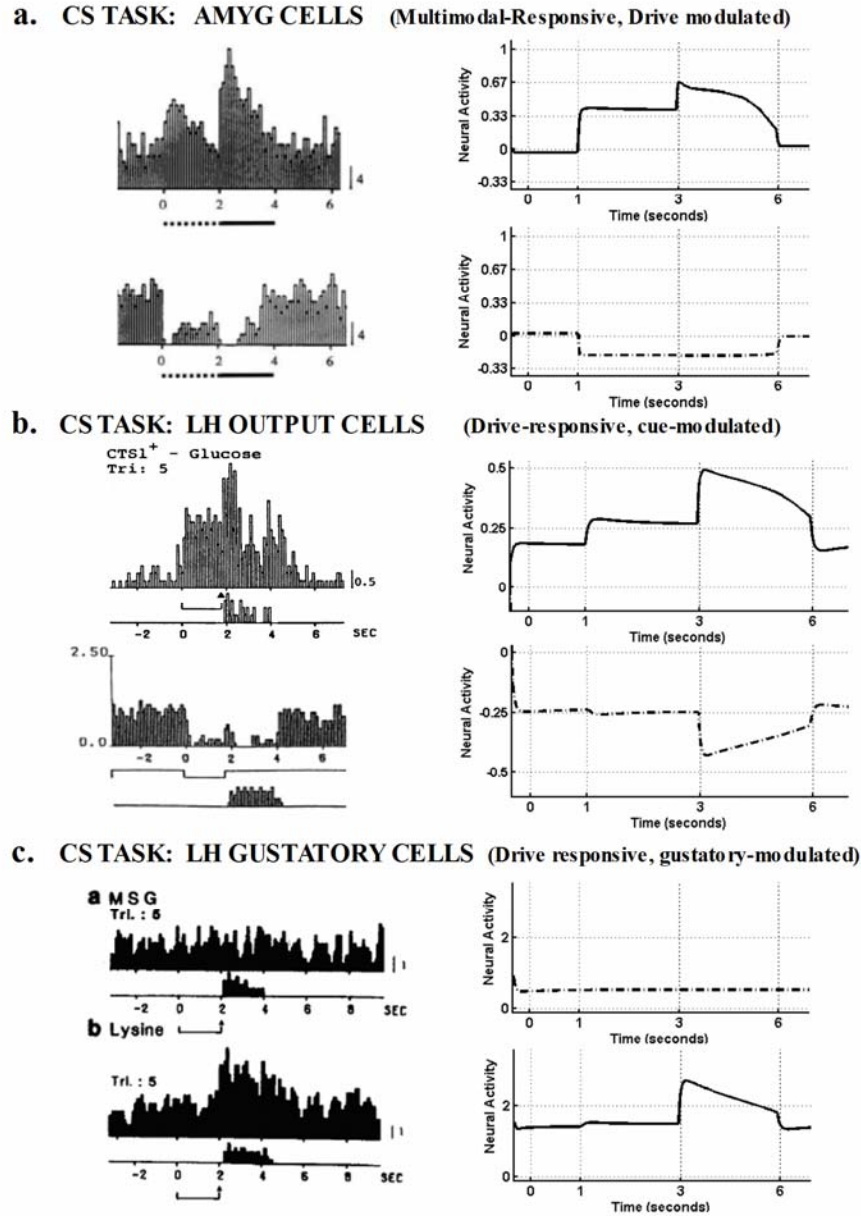
**Figure 4: Simulated and observed responses of cortical cells: ITA, ORB, and MORB:** 4a shows the electrophysiological profile of an ITA cell recorded during the

performance of the SVD task (Jagadeesh et al., 2001), next to the simulated ITA cell activity in the same task (right). Cells in the inferotemporal cortex respond selectively to categories of visual stimuli in a view invariant fashion (Richmond and Sato, 1987). ITA cell types do not discriminate between appetitive or aversive motivational information, but they are modulated by incentive value [Reprinted with permission from Jagadeesh et al. (2001)]. 4b: Orbitofrontal cells distinguish between appetitive and aversive stimuli and respond in proportion to the drive or incentive value of a stimulus (Thorpe et al., 1983; Hikosaka and Watanabe, 2000; Schoenbaum et al., 2003; Roesch and Olson, 2004). Here the electrophysiologically recorded activity from an orbitofrontal neuron during the performance of the CS task is compared with the activity from a simulated ORB neuron performing the same task [Data reprinted with permission from Yonemori et al. (2000)]. 4c: Electrophysiological profile of a reward-responsive orbitofrontal cell recorded during a free reward task (Tremblay and Schultz, 2000a, b) compared with the response from a simulated MORB cell [Reprinted with permission from Tremblay and Schultz (2000a)]. Tremblay and Schultz (Ibid.) report that reward-responsive and cue-responsive cells in the orbitofrontal cortex form distinct populations.

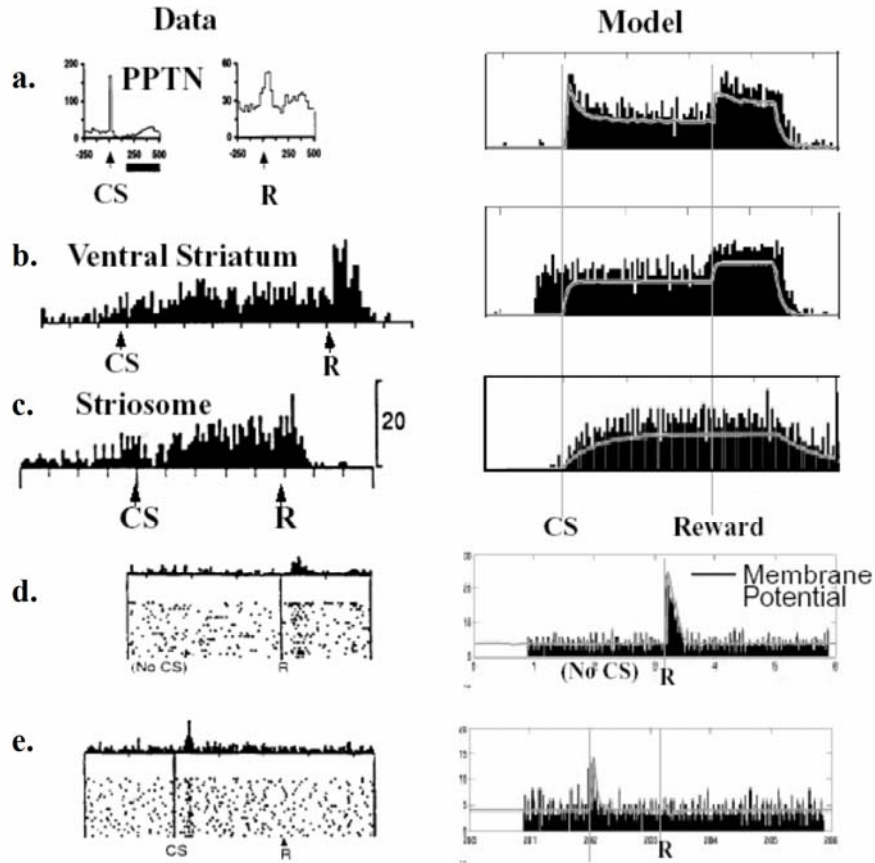
Lateral hypothalamic cells (LH) respond to foods and associated cues (Ono et al., 1986a). Karadi et al. (1992) and Nakano et al. (1986) report that LH cells tend to respond to the deprivation of a given metabolite in the same way they respond to the taste of that metabolite. Ono and associates (1986a) show that LH cells respond in similar fashion to food rewards and the conditioned stimuli that predict them. These responses distinguish between appetitive and aversive inputs and are modulated by hunger (Ono et al., 1986a; Fukuda et al., 1987). These hunger responses are often selective for glucose, specific amino acids, etc. (Torii et al., 1998). Hence an LH cell that was excited by a CS trained to predict glucose would also tend to be excited by glucose deprivation and the taste of glucose. Ono and associates (Nakamura et al., 1987) identified two classes of LH cells whose responses differentiated between appetitive and aversive stimuli: opposite cells and specific cells. Opposite cells respond oppositely to rewarding and aversive stimuli. Specific cells prefer either appetitive or aversive stimuli and do not respond strongly to both.

Four additional cell types were discussed and modeled previously (Brown et al., 1999): pedunculopontine nucleus (PPTN/LDT), substantia nigra (VTA/SNc), matrix medium spiny projection neurons of the ventral striatum (VS), and striosomal delay (SD) cells of the VS. Fig. 6 presents the current model's simulations of these four cell types along with neurophysiological recordings of these basal ganglia cell types. These cell types include: US and CS responsive PPTN/LDT and VS cells, striatal reward-expectant -or SD-cells, and dopaminergic VTA/SNc cells.

In addition to the 11 cell types detailed in Figs. 4-6, the model includes two additional cell types: rhinal (RHIN) cells and lateral hypothalamic input (LH\_in) cells. The LH\_in cells register drive inputs and have activation profiles similar to LH\_gus cells but are separated as a distinct LH class on computational grounds that are described below (cf., Grossberg, 2000a). RHIN cells are included on the basis of evidence from lesion studies which show that the rhinal cortex plays a critical role in the discrimination of food rewards based on flavor or appearance (Parker and Gaffan, 1998). RHIN cells have discharges similar to ITA cells (Liu and Richmond, 2000). Section 4 mathematically describes the dynamics of model cell types.



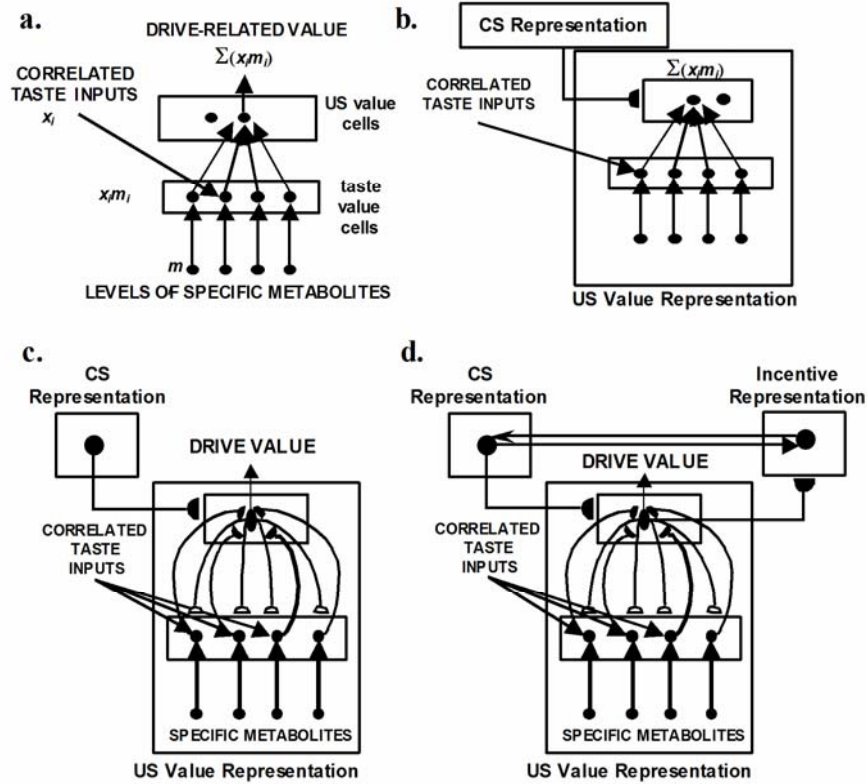
**Figure 5: Simulated and observed responses of subcortical cells during CS task:** 5a: Comparison of the electrophysiological profile of a basolateral amygdala ‘opposite cell’ recorded during the CS task with the activity of simulated AMYG cell in same task (right) [Reprinted with permission from Muramoto et al. (1993)]. 5b: Electrophysiological response profile recorded from an LH ‘opposite cell’ during the performance of the CS Task is shown adjacent to the simulated activity of LH output cells during the same task [Reprinted with permission from Ono et al. (1986a) and Nakamura and Ono (1986)]. 5c: An experimental recording from a LH ‘specific cell’ during performance of the CS is presented along side a simulated LH\_gus cell during the same task [Reprinted with permission from Torii et al. (1998)].



**Figure 6: BG cell types from Brown et al. (1999):** Left: Experimentally recorded activation profiles of neurons (data). Right: Simulated neural activation profiles (model). Fig. 6a: Pedunculopontine tegmental nucleus cells (PPTN cell). 6b: CS and US responsive ventral striatal neuron and simulated VS cell. 6c: Reward expectant striatal cell responds until the delivery of reward (Schultz et al., 2000). SD cells simulate this class of cells. 6d: Electrophysiological profile of dopamine cell recording during reward consumption compared with simulated dopamine neuron. 6e: Electrophysiological profile of CS-responsive dopamine neurons (Ljungberg et al., 1992) compared with simulated dopamine neuron [Reprinted with permission from Brown et al. (1999)].

## 2. RESULTS

**2.1 Model Mechanisms and Processing Stages.** The model integrates and extends previous modeling work treating conditioning, extinction, reversal learning, and cue valuation. MOTIVATOR includes mechanisms that resolve seven issues in succession: (1) the calculation of US value, (2) calculation of CS value, (3) automatic and outcome-specific revaluation, (4) preferential ordering of multiple, simultaneously presented cues, (5) the detection of reward and omission of reward, (6) opponent processing, and (7) segregated pathways for evaluating visual cues and the consummatory value of rewards.



**Figure 7: Stepwise elaboration of an evaluative circuit:** See text for details.

**2.1.1 Calculation of Value.** Fig. 7 illustrates how the model calculates the value of a food reward as it is being tasted, determines the incentive value of a CS in a way that is automatically revalued in an outcome-specific fashion, and disambiguates multiple competing cues while blocking learning of distractors.

Fig. 7a describes a network that calculates the value of an unconditioned stimulus (US) during consumption. Animals form central representations of the drive-related or affective value of a US (Cardinal et al., 2002). The model proposes that one such central representation computes a drive-weighted sum of taste inputs excited during consumption of a food reward. Studies show that humans and animals have specific hungers, henceforth “drives”, that are inversely related to blood levels of metabolites such as sugar, salt, protein, and fat (Davidson et al., 1997; see Section 4.4.3). Similarly, the gustatory system has chemical sensitivities to complementary tastes such as: sweet, salty, umami, and fatty (Rolls et al., 1999; Kondoh et al., 2000; see Section 4.4.2). In general, a food US is complex. It may have several component tastes that correspond to several drives.

Fig. 7a shows a lower layer of cells that perform pairwise multiplications, each involving a taste  $x_i$  and its corresponding drive level  $m_i$ . Thus these cells are called “taste-drive cells”. Taste-drive cells are located in the lateral hypothalamus (LH). LH neurons such as glucose-sensitive neurons provide archetypal examples of LH cells that are both chemo- and taste-sensitive: Glucose-sensitive neurons are excited by low glucose levels, inhibited by high glucose levels, and respond to the taste of glucose with excitation (Shimizu et al., 1984; Karadi et al., 1992). The activation that results from the pairwise multiplication of taste and drive signals in these cells is then projected to a higher cell layer and summed there by a cell that represents the



current value of the US *as a whole*. Thus these cells are called US-value cells. Such US-value representations can emerge from a competitive learning process that associates distributed patterns at the taste-drive cells with compressed representations at the US-value cells that survive the competition at their processing level (see Equations (12-14)). US value cells are located in the amygdala (AMYG) and help explain observations of neurons in the amygdala that selectively respond to specific foods or associated stimuli in a manner that reflects the expected consumption value of the food (e.g. Nishijo et al., 1988a, 1988b).

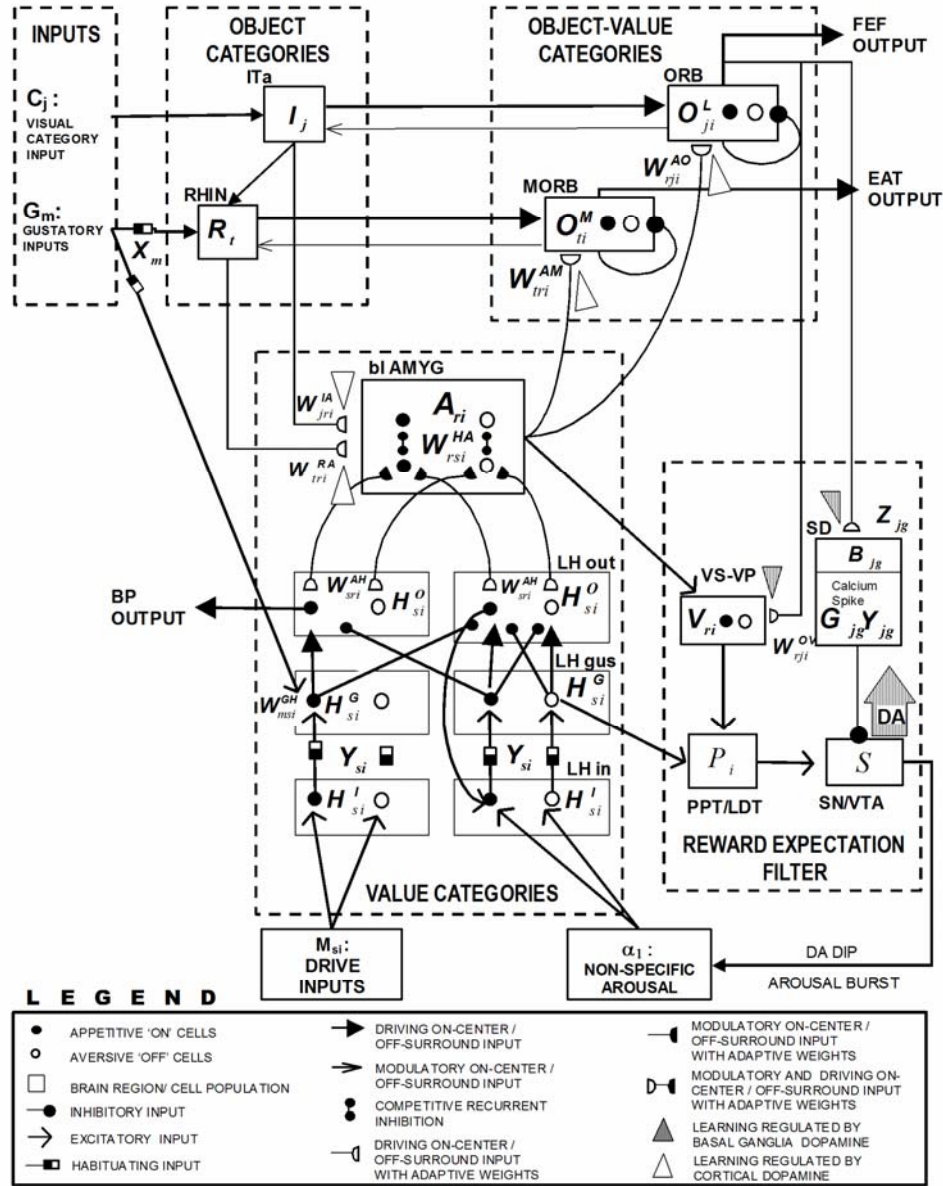
Fig. 7b illustrates the hypothesis that learning can create functional pathways by which a CS becomes a conditioned reinforcer by learning to activate a US-value representation in the AMYG during CS-US pairing protocols (see Equations (10, 38)). Despite the fact the CS generates no gustatory inputs to the taste-drive cells and is not actually consumed, the model is able to use this CS-US association to compute the prospective value of the US, given current drives, during the period between CS onset and US presentation (actual food delivery). The model can do this if the CS-activated US-value representation in the AMYG can, in turn, activate the taste-drive cells in the LH that have activated it in the past, when the US was being consumed.

This is accomplished, as noted in Fig. 7c, by adaptive “top-down” pathways from US-value cells layer in the AMYG to taste-drive cells in the LH. The resultant bidirectional signaling between taste-drive cells and integrative US-value cells can help to stabilize learning and to prime the taste-value combinations that are expected in response to the conditioned reinforcer CS (cf., Carpenter and Grossberg, 1987). The taste-drive cells in the LH multiply the top-down inputs from CS-activated US-value cells by current drive levels, and the resultant activities are projected by convergent bottom-up pathways to US-value cells, which compute a new sum. These interactions introduce nonlinearities of a type that are consistent with Prospect Theory and the principle of diminishing returns: the resultant sigmoid function amplifies values of small rewards while undervaluing large rewards (Kahneman and Tversky, 1979; Grossberg and Gutowski, 1987).

Fig. 7d shows how the circuit can be extended so that the current value of the expected US is used to compute the incentive value of the CS that predicts it. These cells compute object-value properties: They fire when the object or event that they represent has sufficient motivational support (see Equation (11)). As in the Cog-EM model circuit proposed by Grossberg (1972b, 1975; see review in Grossberg et al., 1987), these incentivized CS object-value representations compete, and project modulatory signals back to the sensory stages of CS representation. Such competition and modulatory feedback allow the model to choose among multiple competing cues and to reduce sensory activations of, and block learning with non-predictive CS representations. In addition to sending modulatory feedback that enhances the relative salience of the sensory CS representation, the object-value representations of the CS send output to motor areas to trigger actions towards valued goal objects. Fig. 8 summarizes an anatomical interpretation of these processes in terms of cells in LH, AMYG, ITA, RHIN, and ORBL (see Equations (10-18)).

Lacking from the Fig. 7d circuit are processes that enable the detection of reward omission. Papini (2003) argued that mammals have two systems for detecting nonreward: an allocentric mechanism that resets expectations regarding environmental contingencies, and an egocentric mechanism that estimates the motivational, homeostatic and emotional cost of nonreward. The model proposes that adaptive timing mechanisms (Grossberg and Merrill, 1992) and habituated opponent processing via gated dipole circuits (Grossberg, 1972a, 1972b) fulfill

the allocentric and egocentric components of reward-omission learning (Dickinson and Balleine, 2001; Papini, 2003). The adaptive timing and gated dipole models are discussed below.

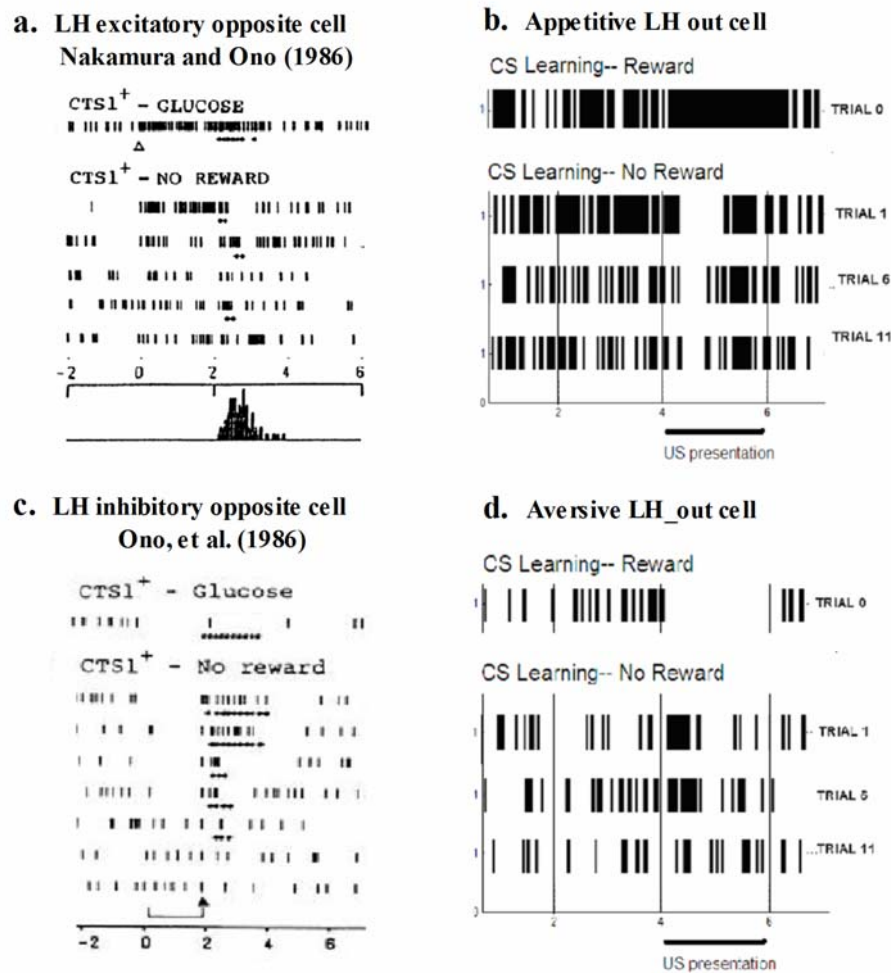


**Figure 8: Model detail with equation variables:** Distinct cell types are represented with different labeled compartments. Cells may or may not show a selective response to affective information of different valences. Cells which selectively respond to appetitively valenced information are indicated by filled circles. Cells which selectively respond to aversive affective information are indicated by open circles. Activation is transmitted between cells along specific pathways. Pathways are indicated by edges with arrowheads (fixed excitatory connection), semicircles (learned excitatory connection), or filled circles (fixed inhibitory connection). Filled arrowheads carry driving excitatory inputs. Open arrowheads carry excitatory signals that modulate or multiply driving



inputs. Similarly, filled semicircles carry adaptively gated driving inputs while open semicircles carry adaptively gated signals that modulate driving inputs. Half-filled rectangle: Pathways which show activity driven habituation. Weights ('W') have a pathway specific superscript. For variable names, subscript and superscript definitions, and other equation details see Section 4.

**2.1.2 Adaptive Timing Mechanisms.** Adaptive timing circuits can learn temporal expectations that help an animal to balance planned vs. reactive, or consummatory vs. exploratory, behaviors (Grossberg, 1982; Grossberg and Schmajuk, 1989; Grossberg and Merrill, 1992; Brown et al., 2004). More recently, Cohen et al. (2007) have discussed this as the balance between exploitation vs. exploration. The model incorporates adaptive timing to detect the unexpected omission or presentation of rewards and cues that predict rewards. Brown et al. (1999) described a basal ganglia (BG) model in which an unexpected US or CS triggers a dopamine (DA) burst while the omission of an expected US elicits a DA dip (Fig. 9; see Equations (20-33)). The DA burst is a reinforcement signal that speeds learning of cue-reward associations (in VS, ITA, ORB AND AMYG) and cue-response associations (in dorsal striatum), whereas the DA dip speeds the learning of associations that mediate extinction.



**Figure 9: Antagonistic rebounds in simulated and observed lateral hypothalamic cells:** Left: 9a and 9c electrophysiological recordings from lateral hypothalamic

“opposite cells” in rats during learning and extinction trials. Right: Simulated neural activation profiles (model). 9a: Response of an excitatory lateral hypothalamic opposite cell. During rewarded trials the cell responds strongly to both CS and US presentation. After several extinction trials activation dissipates [Reprinted with permission from Ono et al., (1986a)]. 9b: Simulated appetitive LH\_out cells show a transient suppression following reward omission. Trial 0 indicates the last rewarded CS task. Trials 1 through 11 indicate successive extinction trials. 9c: Inhibitory opposite cell recorded from a rat during the CS task. The cell is normally inhibited by CS and US presentation and shows a strong transient excitatory response when reward is omitted [Reprinted with permission from Nakamura and Ono (1986)]. 9d: Simulated LH\_out aversive cells show a transient excitation following the omission of reward. Spikes are generated from appetitive and aversive LH\_out cell activity during extinction trial simulations of the CS task.

The BG adaptive timing circuit (Figs. 1 and 8) includes a model ventral striatum (VS) in which convergent inputs from US-specific value cells and CS-specific ORB cells help condition cue-reward associations, the pedunculo-pontine nucleus (PPTN) which relays CS- and US-related excitations to the SNc/VTA, and striosomal delay cells (SD) which compute a CS-cued and adaptively-timed inhibition of DA cells in the SNc/VTA. Dips and bursts in dopamine cell activity are the result of the balance of phasic inhibitory and excitatory inputs to the SNc/VTA. SD cells issue an inhibitory signal only if they have been activated by a cue that reliably predicts reward. Unexpected rewards elicit dopamine bursts because the occurrence of reward excites PPTN/LDT cells and, in turn, SNc/VTA cells while anticipatory signals from SD cells to the SNc/VTA never materialize. Dopamine dips are generated when an expected reward is omitted. In this case, a predictive cue is presented and SD cells send an adaptively timed inhibitory input to the SNc/VTA at the expected time of reward. However, at the expected time of reward, the reward is omitted and dopamine cells in the SNc/VTA receive no excitatory input from the PPTN/LDT to offset inhibitory inputs from SD cells. The result is a transient suppression or dip in the activity of dopamine cells at the expected time of reward.

*2.1.3 Gated Dipole Opponent Drive Processing.* The gated dipole opponent processing model was developed to address the interactions of appetitive and aversive motivational processes (Grossberg, 1972b, 1975, 2000a; Solomon, 1980). Behavioral studies show that appetitive and aversive stimuli are not simply processed in independent parallel circuits, but interact as though processed in opponent systems yielding such important phenomena as summation, excitatory and inhibitory conditioning, the relief associated with the offset of pain, or frustration following omission of reward (Amsel, 1968; Denny, 1970; Dickinson and Dearing, 1979; Weiss et al., 1996; Dickinson and Balleine, 2001).

The gated dipole model (embedded in the LH of Fig. 8; see Equations (15-17, 34)) describes an opponent mechanism by which the cessation of appetitive or aversive stimuli can generate an antagonistic rebound signal of the opposite valence. A feedforward gated dipole circuit obeys five constraints: (1) separate ON- and OFF-channels that process appetitive and aversive information; (2) input cells that summate two types of excitatory inputs: *phasic inputs* that result from the presentation of appetitive or aversive stimuli, and *tonic inputs* that affect both channels equally, reflecting baseline arousal levels; (3) slowly habituating and recovering chemical transmitter levels that gate, or multiply, the phasic-tonic signals emerging from input cells; (4) second-stage cells that receive the gated signal from the phasic-tonic cells and relay this gated signal with excitatory sign to their own channel but inhibitory sign to the opponent

channel; and (5) a competitive or opponent, output stage, at which cells in either channel can fire only if the excitation they receive from their channel exceeds the inhibition they receive from the opponent channel.

The antagonistic rebound properties of a gated dipole arise from the difference in reaction rates between the slowly habituating transmitter gates and fast changes in the phasic or tonic inputs. The offset of phasic inputs may generate rebounds (e.g., the offset of a fearful stimulus triggers relief; Denny, 1970). The level of tonic arousal inputs provides the energy, and controls the sensitivity, of the rebound to cue offset. When tonic arousal suddenly increases (unexpected events are arousing) the gated dipole may again generate a rebound. Both sets of rebound disconfirm ongoing affective processing. The model (Fig. 8) assumes that a dip in the activity of dopaminergic SNc/VTA cells generates a fast arousal increment (see Section 5 4.4.4), which generates a rebound activation in the output cells of an OFF channel that were inactive before the DA dip. Such a rebound in the OFF channel that opposes an appetitive ON channel rapidly and selectively shuts off formerly active channels and activates a negative affective state (e.g., rebound from hunger to frustration; Amsel, 1968).

Both the phasic offset and tonic onset types of antagonistic rebound can speed extinction and preserve outcome-specific information, as shown below. Moreover, the antagonistic rebounds and opponent properties predicted by gated dipole output cells allow the model to explain electrophysiological responses of lateral hypothalamic *opposite cells* reported by Ono and associates (Nakamura and Ono, 1986; Ono et al., 1986a; Nakamura et al., 1987). These cells appear to be organized in ON and OFF (appetitive and aversive) channels, such that cells that are excited by an appetitive CS and its associated US (such as glucose) are also inhibited by the presentation of an aversive stimulus and its associated US (such as electric shock). Also, these cells show rebounds following the omission of reward: appetitive cells are rapidly shut off while aversive cells are transiently activated. Figs. 9a and 9c show two such cells as reported by Ono and associates (Nakamura and Ono, 1986; Ono et al., 1986a), whereas Figs. 9b and 9d show model simulations of these responses.

**2.1.4 Recurrent Dipoles and Rebounds to CS Offsets.** The hypothalamic dipole circuits in Fig. 8 are recurrent, or feedback, networks because they have excitatory connections that carry signals from the output cells back to the same channel's input cells (see Equations (15, 17)). Recurrent dipoles can maintain a motivational baseline in the presence of small distracting inputs, prevent learned synaptic weights from saturating, and allow secondary inhibitory conditioning (Grossberg and Schmajuk, 1987). In the present model, without recurrence, the dipole could generate rebounds to arousal increments or US offsets, but not to CS offsets. With recurrence, the habituating transmitter levels in the circuit are capable of adapting to CS-related inputs, thereby enabling rebounds to be generated in response to CS offsets. Thus termination of an appetitive CS generates an aversive affective reaction which is necessary for secondary inhibitory conditioning.

**2.1.5 Parallel cortical incentive processing channels for US and CS.** Fig. 8 includes separate pathways for processing visual CS (involving learned and unlearned connections between ITA, ORB, and AMYG; see Equations (10-12, 36, 38) and gustatory US information (involving learned and unlearned connections between RHIN, MORB, and AMYG; see Equations (12, 18, 19, 39, 41)). Electrophysiological studies suggest that US and CS processing in the orbital prefrontal cortex are distinct (Tremblay and Schultz, 2000b). Lesion studies also support this analysis, as RHIN lesions disrupt US but not CS processing (Parker and Gaffan, 1998). The model explains these observations by segregating the US-processing RHIN-to-

MORB (medial orbitofrontal) stream from the CS-processing ITA-to-ORB (lateral orbitofrontal) stream. Separate streams for US and CS processing supports specific feedback enhancement by the respective orbitofrontal regions of sensory representations at the stages (RHIN vs. IT) that feed them. Such feedback speeds learning that involves highly active representations in the affected loop while blocking learning in competing representations.

**2.2 Model Parameters and Task Protocols.** The training sequence for the model mirrors the training regime for the SVD-FSS (Simultaneous Visual Discrimination – Food Specific Satiety) task described by Murray and colleagues (Baxter et al., 2000). The US learning task was simulated first, followed by the CS learning task and the SVD task. Training the model in this order allows the AMYG to form a US-specific representation with which visual stimuli in the CS or SVD tasks can later be associated.

*2.2.1 System Inputs, Outputs, and Equations.* There were four inputs that could be varied during any trial: a phasic CS signal, a phasic US signal, a drive input, and an arousal input (see Section 4.2). For all neural elements, the model computes time-varying activations, two of which (Fig. 3) can be regarded as output signals: a lateral hypothalamic output related to blood pressure and a ORB-generated object-value output used to compute the reaction time of a saccade (see Equations (44-46)). In the SVD task, saccades are instrumental, whereas in the CS task saccades are not needed to gain reward.

The differential equations that specify model interactions and dynamics fall into three general categories: short term memory (STM) equations that describe changes in neuronal activation levels (see Equations (1-4)), medium term memory (MTM) equations that describe experience-dependent but short-term effects such as facilitation, short-term depression, and transmitter habituation and recovery (see Equation (5)), and long term memory (LTM) equations that describe experience-dependent changes that can be long-lasting, most notably long term potentiation or depression of synaptic efficacy (see Equations (6-9)). Parameters, initial values, inputs, and outputs for these equations are discussed in the Section 4.

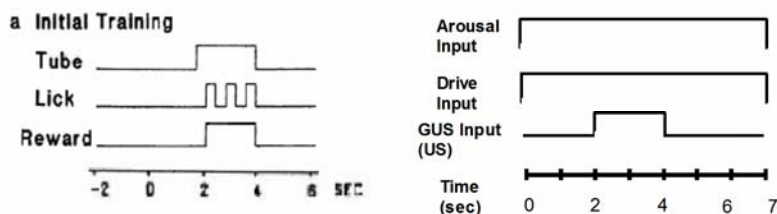
## **2.3 Simulations**

*2.3.1 US Learning Task.* The simulated US task follows the timing of the Ono et al. (1986a) version of the reward presentation task (Fig. 10a). The US task trains the model to recognize drive and GUS representations of two rewarding food stimuli, US1 and US2.

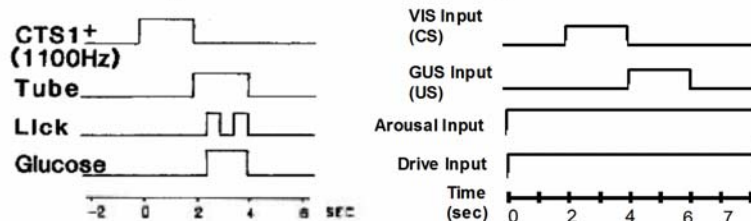
*Satiety Devaluation of US Task.* The devaluation task was simulated after the model was fully trained to recognize US1 and US2. The satiety devaluation simulation examined the response of the system to US1 as GUS and hunger inputs were systematically decreased from high initial values across 20 trials. In the first trial of the devaluation sequence, there is no GUS habituation and the initial values of the opponent hunger and satiety inputs were set to 4 and 0, respectively; by the end of the devaluation sequence, the hunger and satiety inputs had each reached a value of 2. To isolate the effects of habituation and satiety, LTM values (synaptic weights) were fixed across these twenty trials. This simulation demonstrated that the model could register continuous changes in US value.

*Food Specific Satiety US Task.* The FSS simulation tested the specificity of US1 devaluation by satiety. This simulation demonstrated an FSS effect; namely, the model's response to US1 was devalued while its response to US2 was spared. GUS habituation and hunger drive levels were changed to reflect FSS such that US1-specific tastes were habituated to 0.4 and US1-specific drives were satiated, from 4 to 2.

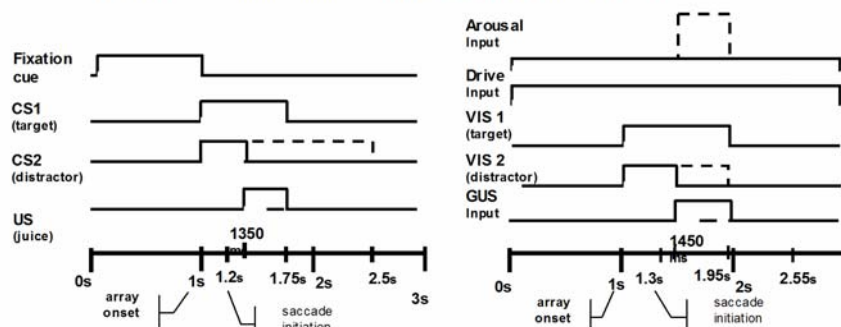
**a. US Learning Task: Food Familiarization or Free Reward Task**



**b. CS Learning Task: Pavlovian Conditioning**



**c. SVD Task: Simultaneous Visual Discrimination: CS1+ vs. CS2-**



**Figure 10: Timing and input sequence for the US, CS, and SVD tasks:** 10a, left: Timing of free reward task reported by Ono, et al (1986) [Reprinted with permission from Ono et al. (1986)]; 10a, right: Simulated US task utilizes un-cued delivery of gustatory inputs and two endogenous inputs: drive and arousal. 10b, left: CS Learning Task as reported by Nakamura et al. (1987) [Reprinted with permission from Nakamura et al. (1987)]; 10b, right: Simulated version of CS Task. 10c, left: Simultaneous visual discrimination task (SVD) as reported by Jagadeesh et al. (2001). Solid line: Time course of inputs during rewarded trials. Dashed line: Time course of inputs during unrewarded trials. During unrewarded trials there was a 2 second time out rather than a juice reward. 10c, right: Simulated version of SVD Task. Simulation experiments were based on task reported by Jagadeesh et al. (2001) except a fixation cue was omitted. Solid lines: Input sequence during rewarded trials. Dashed lines: Input sequence during unrewarded trials. The increase in arousal during unrewarded trials is not provided as an input but is contingent on a dopamine dip signal that follows reward expectation learning in the basal ganglia circuit.

**2.3.2 CS Learning Task.** There were four simulations of the CS task: CS learning, extinction, satiety and devaluation, and outcome-specific devaluation. Simulations of the CS task demonstrate how sensory cues associate with object-value and drive representations in the ORB and AMYG, replicate the devaluation of a CS that occurs when its associated US is specifically sated, match reported acceptance and rejection curves, and demonstrate the specificity of the automatic devaluation of a CS. These model simulations also replicate the dynamics of experimentally observed neurophysiological cell types in the AMYG, LH, ORB, SNc/VTA, VS, PPTN, and striosomal striatum (SD).

The simulated CS task mimicked the timing of the CS task reported by Nakamura et al. (1987) (Fig. 10b). The CS learning task demonstrates that the model can learn two stimulus-reward associations with distinct outcomes: CS1 was rewarded by a subsequent presentation of US1 (CS1+US1) and CS2 was paired with US2 (CS2+US2). The task also allowed the model to demonstrate that motivation impacts saccadic latency. The model uses a cumulative spike counter and a ‘race to threshold’ rule for generating saccades. In particular, the model FEF receives and sums stimulus-related activity, effectively acting as a cumulative spike counter for each stimulus (see Equation (46)). When the cumulative activity associated with a stimulus reaches a fixed threshold of 0.3, that stimulus is selected as the target for a saccade and a response is generated. When multiple stimuli are present, the target is determined by a ‘race to threshold’.

When a CS activates the corresponding CS-selective ORB cell, ORB cell activity is sent to a CS-selective FEF cell. The CS-selective FEF cell integrates ORB cell inputs, biased by arousal and ITA inputs, across time until the cumulative activity exceeds a fixed threshold of 0.3. After the threshold is exceeded, a saccade is made to the stimulus associated with the winning ORB cell. Mechanisms detailed by Brown et al. (2004) in the TELOS model explain how saccadic targets are selected by the frontal eye fields (FEF). In TELOS, object category-selective FEF cells receive ITA inputs and can be modulated by motivational signals from the ORB, thereby biasing the selection of saccadic targets and influencing saccadic reaction times.

**CS Extinction Trials.** The extinction simulations consisted of 20 conditioning trials (rewarded) followed immediately by twenty extinction trials (unrewarded). The details and timing of the CS extinction task are as described in Fig. 10b except that, during an extinction trial, the US is not presented, and thus the CS is unrewarded (CS-).

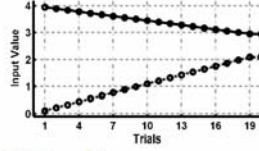
**CS Devaluation Trials.** CS devaluation trials demonstrate that the system decreases its responses to CS1 and US1 as GUS and hunger inputs are systematically decreased. LTM adaptive weight values were fixed across trials.

**Outcome-Specific CS Devaluation.** These simulations demonstrate the specificity of the CS devaluation effect. To examine how neuronal responses to CS presentation would differ at various stages of satiation, a sequence of twenty trials of CS presentation were simulated, each successive trial using a progressively smaller drive input (Fig. 11a-e(1)). In the first trial, the opponent hunger and satiety inputs were set to 4 and 0, respectively. By the end of the devaluation sequence, the hunger and satiety inputs had each reached values of 2 and 3. To isolate the CS effect, no US was presented. Therefore, these trials are extinction trials. To isolate the effect of drive satiety alone, GUS habituation and (un)learning were turned off during these 20 trials. Comparisons of final and initial simulated AMYG, LH, and ORB responses are also shown in Fig. 11. These simulations demonstrated that the automatic and outcome-specific devaluation of a CS1 that was previously paired with a satiated US1 coexists in the model with a maintained strong response to a CS2 that was previously paired with a non-satiated US2.

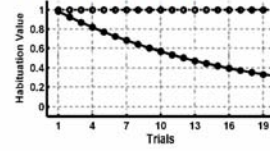


### a. Satiation Sequence Inputs

#### (1) DRIVE INPUT Satiety

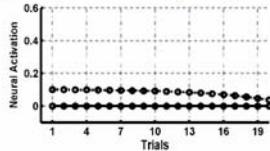


#### (2) GUS INPUT Habituation

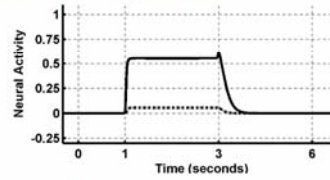


### b. ORB cells

#### (1) Response to Satiety Sequence

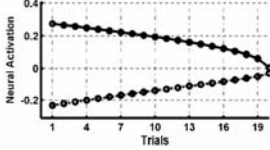


#### (2) Differential Satiety Response

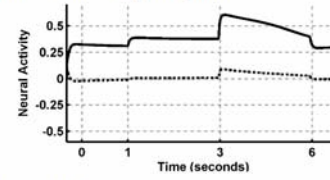


### c. LH output cells

#### (1) Response to Satiety Sequence

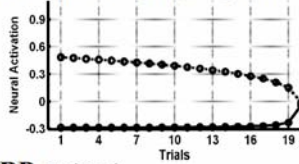


#### (2) Differential Satiety Response

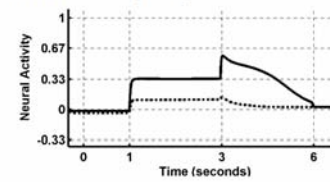


### d. AMYG cells

#### (1) Response to Satiety Sequence

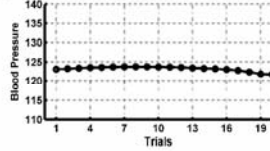


#### (2) Differential Satiety Response

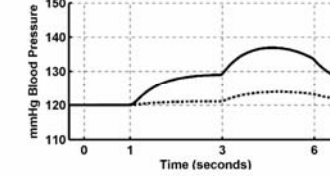


### e. BP output

#### (1) Response to Satiety Sequence

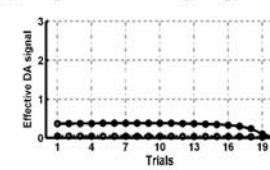


#### (2) Differential Satiety Response

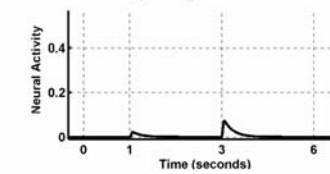


### f. Cortical Dopamine Output

#### (1) Response to Satiety Sequence



#### (2) Differential Satiety Response



**Figure 11: Food-specific satiety and CS devaluation during CS task:** 11a-11e, left column and 11a(2): Effects of 20 trials of progressively increasing satiety on neural activities related to CS presentation. No reward was presented during these trials. Data points in figures represent the average activation for the indicated neural variable during CS presentation ( $t=1$  until  $t=3$  seconds). 11b-11e, right column except (a.2): Plots present an overlay of traces showing the activation of simulated variables during automatic revaluation of Pavlovian stimuli following food-specific satiety. Solid lines: Response to CS2 presentation, CS2 is paired with the unsated reward US2. Dashed lines: Response to CS1 presentation, CS1 is paired to the devalued reward US1. US1-related

GUS signals were set to a habituation level of 0.35 while US1-related hunger inputs were set to 3 and satiety inputs nonspecifically set to 2. LTM values were fixed for all trials. **Key:** 11a shows inputs during devaluation sequence: (1): Drive inputs, hunger and satiety; (2): Consumption-related gustatory habituation. 11b(1): US1-specific ORB response (solid lines) and US2-specific ORB response (dashed lines). 11b(2): Differential response of CS1 (dashed) and CS2-specific (solid) ORB cells during the automatic revaluation simulation. 11c(1): Response of appetitive (solid lines) and aversive (dashed lines) glucose-sensitive LH\_out cells. 11c(2): Response of the glucose-sensitive LH\_out cell to CS1 when US1 is sated (dashed lines) vs. response to CS2 when US2 unsated (solid lines). 11d(1): US1-specific AMYG response (solid lines) and US2-specific AMYG response (dashed lines). 11d(2): Differential response of US1-specific (dashed) and US2-specific (solid) AMYG cells to CS1 or CS2 presentation. 11e(1): CS1-generated blood pressure response. 11e(2): Differential blood pressure response when CS1 is presented (dashed lines) vs. when CS2 is presented (solid lines). 11f(1): Average dopamine burst in response to CS1 presentation. 11f(2): Differential dopamine response to CS1 presentation (dashed) vs. CS2 presentation (solid).

**2.3.3 SVD Task.** Five SVD tasks were simulated: SVD learning, extinction, reversal, alternation, and the SVD task with specific food reinforcer devaluation by satiation (SVD-FSS). The first simulation demonstrates dopamine-driven changes in cue preference with the learning and reversal of a visual discrimination problem, and shows qualitative matches to the behavioral curves reported by (Jagadeesh et al., 2001). A second simulation demonstrates that food specific satiety (FSS) can also alter cue preference in a fairly dopamine-independent fashion. These simulations of the SVD-FSS task show qualitative matches to the behavioral results described in the experiments by Murray and colleagues (Malkova et al., 1997). The SVD-FSS simulation highlights how the AMYG-LH system influences behavior and preferences. While performing the SVD task, model neuron discharges resemble those of cell types experimentally observed in the AMYG, ITA, and ORB.

The SVD task trial structure was based on the SVD experiments performed by Jagadeesh et al. (2001), as illustrated in Fig. 10c. In the SVD learning task, two stimuli are simultaneously presented; saccadic choice of the stimulus designated as the target is rewarded (CS1+), whereas saccadic choice of the other stimulus, designated the distractor, is unrewarded (CS2-). Successful trials reinforce the CS1+US1 association. As in the CS task, during the SVD task the model generates saccadic responses using a cumulative spike counter in the FEF and a ‘race to threshold’ rule (see Equation (46)). Whichever CS-selective FEF cell is first to exceed the fixed threshold of 0.3 determines the stimulus selected for a saccade. Inputs from the ORB to the FEF ensure that saccadic reaction time is modulated by motivational state.

*SVD Task: Extinction.* These simulations demonstrate the extinction of the association between CS1 and US1. Trials in the extinction protocol differed from trials in the SVD training protocol, shown in Fig. 10c, in only one regard: the US presentation was omitted. When the CS1+US1 association is extinguished, discrimination behavior (saccadic preference) drops to chance.

*SVD Task: Reversal Learning.* For the reversal experiment, one US is used throughout the simulation. After a block of trials in which only saccadic choice of CS1 leads to US (reward) delivery, the contingency reverses for the next block, and only choice of CS2 leads to reward. Within each trial, the input sequence was the same as described in the normal SVD task (Fig.



10c). The reversal experiment demonstrates how reinforcement learning and dopaminergic value systems can change cue preferences in response to contingency reversals.

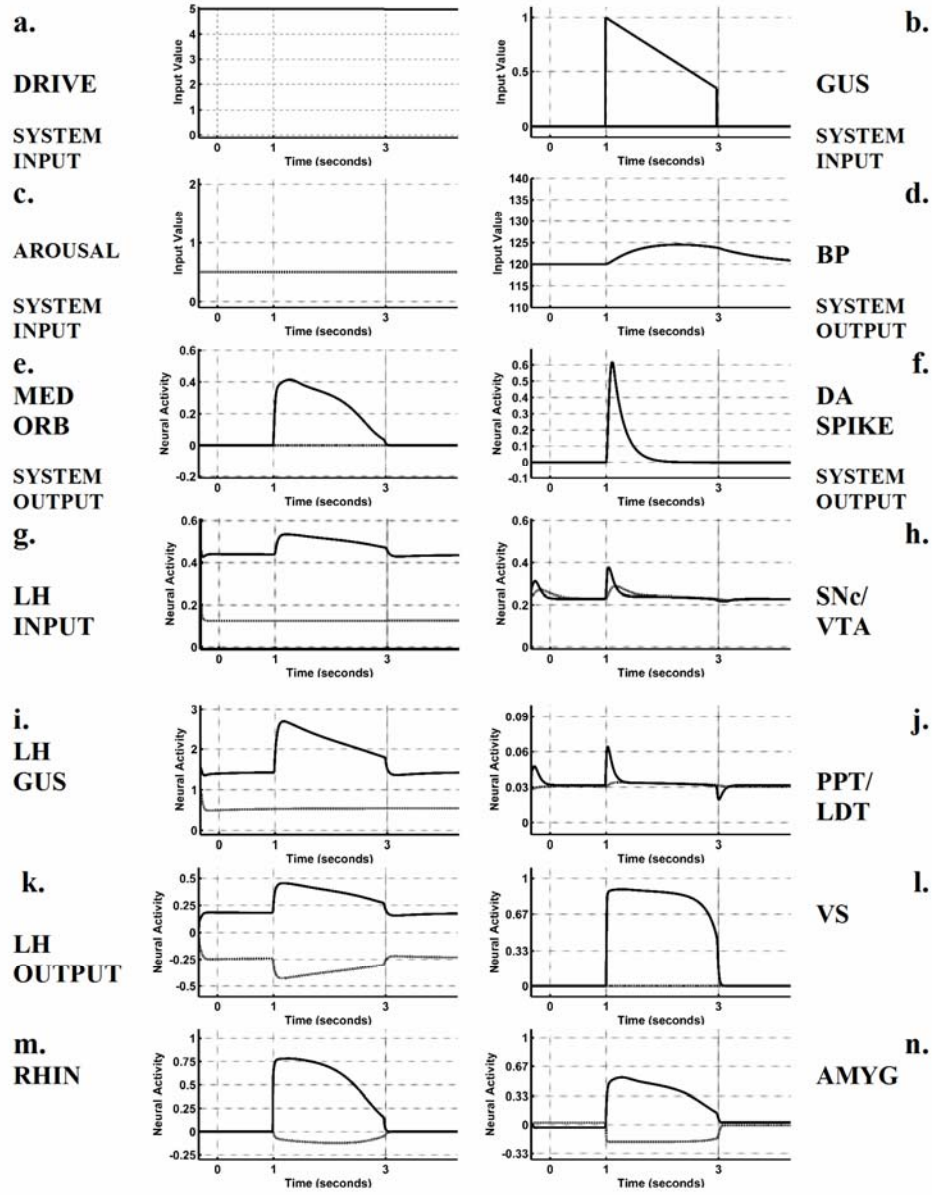
*SVD Task: Alternation.* The alternation task incorporates some elements also found in the SVD-FSS task (below). Two associations, CS1+US1 and CS2+US2, were pre-trained prior to alternation simulations. Then both CS1 and CS2 were presented simultaneously while the model maintained equally high drive for both US1 and US2. In the first block of 15 trials, saccadic choice of CS1 was rewarded (CS1+US1) but CS2 was unrewarded (CS2-). Then, to create the alternation, for trials 16-30 saccadic choice of CS2 was rewarded (CS2+US2) but CS1 was unrewarded (CS1-); e.g., ORB recording in serial alternation example from Thorpe et al. (1983). Otherwise, the alternation protocol used the same presentation timings as the SVD task (Fig. 10c). In the model, the beginning of the alternation engenders a violation of a timed expectation of reward. The resultant dopamine dip, which acts directly as a (negative) reinforcement learning signal on cortical and striatal sites, also disinhibits a nonspecific source of arousal to the hypothalamus, engendering an arousal burst that leads to an antagonistic rebound in the AMYG. Together, these signals can bring about quick changes in preference for a cue, despite a substantial recent history of reward for that cue.

*SVD-FSS Task.* As with the alternation experiment, two visual discriminations were trained prior to the SVD-FSS experiments, resulting in the learning of CS1+US1 and CS2+US2 associations. In the SVD-FSS task, saccadic choice of either stimulus is always rewarded, CS1 with US1 and CS2 with US2. Hence the visual discrimination is not between a rewarded target and an unrewarded distractor, but between two reward-predicting stimuli (CS1+ vs. CS2+) on the basis of the animal's current preference for the expected outcomes, US1 or US2. For the first five trials, hunger and GUS inputs were high for both outcomes. For the last five trials, these inputs for outcome US1 were set to values indicative of satiety. But CS-US synaptic connections were left unmodified. The input sequence of the SVD-FSS trial was similar to that described for SVD trials (Fig. 10c). The SVD-FSS task affords a view of the impact of food-specific satiety on decision-making. The SVD-FSS task tests whether the model's cue (CS) preference can shift solely as a result of selective satiation, without any additional reinforcement learning.

## **2.4 US Learning Task**

*2.4.1 Normal Performance of the US Task.* The model's performance of the US Task can be parsed into five phases: (1) Equilibration to drive inputs, (2) US presentation, (3) Calculation of US metabolic value, (4) Incentive value and response generation, and (5) Dopamine responses. Fig. 12 details these stages and presents the results from the simulation of one US task trial.

*2.4.2 Normal Learning of the US Task.* Simulations of the US task demonstrate how outcome-specific representations of US1 and US2 food rewards form. The model requires approximately 40-50 trials before the LTM weights become asymptotically stable. Some aspects of US processing are learned while others are not. Thus, connections between basal ganglia, hypothalamic, and gustatory regions are assumed to have been learned prior to the US task, so US presentation elicits a large dopamine, blood pressure, and LH response prior to any training (Ono et al., 1986a; Nakamura et al., 1993). Connections involving the ITA, RHIN, AMYG, ORBM, and ORBL, in contrast, are learned during trial simulations, so training with the US task is essential for cortically represented stimuli to gain access to outcome-specific information.



**Figure 12: Performance of the US learning task:** 12a-12n: Activity of system variables during US task. US turns on at  $t=1$  second. US1 activates GUS input (12b) corresponding to ‘sweet’ taste. GUS inputs activate RHIN (12m) and LH\_gus (12i) cells. RHIN cells (12m) categorically recognize each US by taste features and activate AMYG cells (12n). LH\_gus cells (12i) receive taste inputs that correlate with the specific metabolite or drive processed in that LH\_gus cell. LH\_gus activity projects to LH\_in (12g), LH\_out (12k), and PPTN (12j) cells. The affective value of a US is calculated by AMYG cells (12n) which cluster and recognize US-specific drive features. Affective value is calculated by summing the taste-modulated drive inputs from LH\_out cells (12k) during US consumption. US incentive value is measured by MORB cells (12e) which receive contemporaneous RHIN (12m) and AMYG (12n) cell inputs. Dopamine bursts at US onset, and dips at US offset, (12f), are generated by SNc/VTA cells (12h) which are excited by PPTN inputs (12j). **Key:** 12a: Glucose driven hunger input. 12b: GUS input

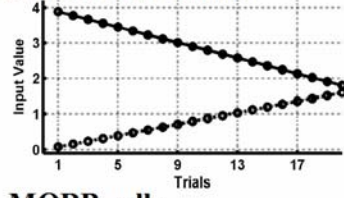
for sweet taste. 12c: Nonspecific arousal input. 12d: Blood pressure output of the model. 12e: US1-specific (solid line) and US2-specific (dashed line) MORB cells. 12f: Effective cortical dopamine burst or dip (solid line). 12g: Glucose hunger driven LH\_in cell (solid line); opponent LH\_in satiety cell (dashed line). 12h: Phasic output of SNc/VTA cells (solid line); time-averaged SNc/VTA output (dashed line). 12i: LH\_gus cell sensitive to sweet taste and glucose drive inputs (solid line); dashed lines trace the opponent LH\_gus cell. 12j: Phasic output of PPTN cell (solid line); lasting hyperpolarization of PPTN (dashed line). 12k: Sweet-taste and glucose-specific LH\_out cell (solid line); dashed line traces activity in its aversive opponent cell. 12l: VS cell responsive to US1. 12m: US1-specific RHIN cell category (solid line); US2-specific RHIN cell (dashed line). 12n: US1-specific AMYG cell.

*2.4.3 US Devaluation and Satiety.* Fig. 13 shows the results from model simulations in which hunger was systematically reduced and GUS input systematically habituated across twenty trials (see Equation (35); for input regime see Methods, Section 2.3.1). The responses of LH output (Fig. 13c(1)), MORB (Fig. 13b(1)), and AMYG cells are clearly diminished by increasing satiety and the habituation of GUS inputs (Rolls et al., 1986; Nishijo et al., 1988a; Yan and Scott, 1996). While not shown, RHIN cell responses also decrease because of the habituation of GUS inputs (see Equations (18)). In these simulations, the majority of the devaluation of cell responses is due to gustatory habituation, but comparison with CS devaluation simulations demonstrates that satiety inputs alone are capable of significantly diminishing cell responses. Sensory-specific gustatory signal habituation and drive-specific satiation are each integral mechanisms of food specific satiety (FSS). Blood pressure and dopamine output are also attenuated by increases in satiety level (see Figs. 13d(1), 13e(1)).

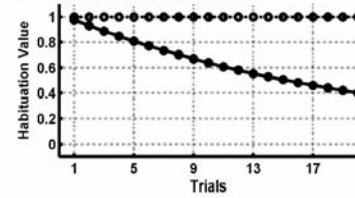
*2.4.4 Simulations of Food Specific Satiety.* The second column of Fig. 13 shows that the US devaluation is food-specific. The results demonstrate that US1 is devalued relative to US2 in the response of blood pressure (BP) and dopamine, MORB, and LH cells (Figs. 13e(2), 13d(2), 13b(2), 13c(2)). The differential valuation of US1 and US2 occurs because they are composed of different nutrients (see Section 4.2.2). Decreasing the hunger drive for US1 leaves the hunger drives for US2-associated nutrients elevated. The high drive for US2-associated nutrients means the AMYG calculates a high value for US2. The RHIN and other LH cells (not shown) are similarly differentially modulated by the food-specific satiation (Rolls et al., 1986).

### a. Satiation Sequence Inputs

#### (1) DRIVE INPUT Satiaty

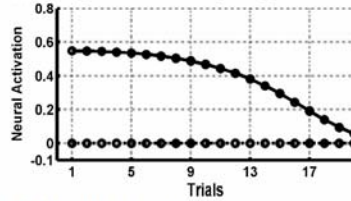


#### (2) GUS INPUT Habituation

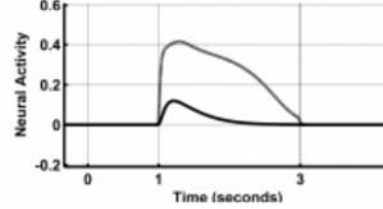


### b. MORB cells

#### (1) Response to Satiety Sequence

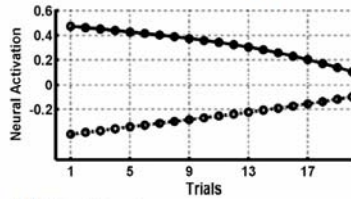


#### (2) Differential Satiety Response

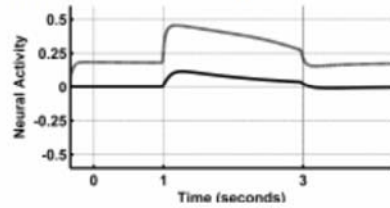


### c. LH output cells

#### (1) Response to Satiety Sequence

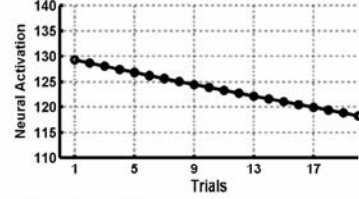


#### (2) Differential Satiety Response

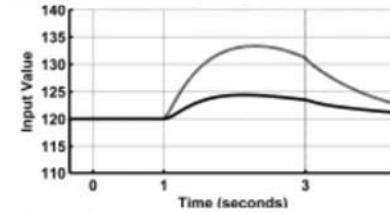


### d. BP output

#### (1) Response to Satiety Sequence

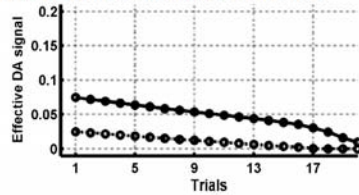


#### (2) Differential Satiety Response

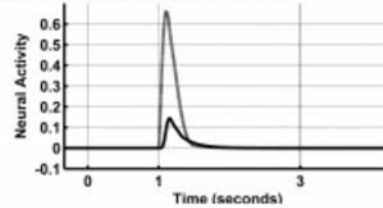


### e. Cortical Dopamine Output

#### (1) Response to Satiety Sequence



#### (2) Differential Satiety Response



**Figure 13: Food-specific satiety related devaluation of STM activity:** 13a-13e, left column and 13a(2): Effects of 20 trials of progressive satiation on neural activation related to US consumption. Data points in figures represent the average activation for the indicated neural variable during US presentation ( $t=1$  until  $t=3$  seconds). 13b-13e, right column except 13a(2): Plots present an overlay of traces showing the activation of simulated variables during the consumption of different food rewards. Dashed lines: The sated reward, US1, was presented. Solid lines: An unsated reward, US2, was presented. For FSS trials, US1-specific hunger and satiety inputs were set to 2 and 1.5, respectively; US1-specific GUS tastes had been habituated to 0.43. **Key:** 13a(1): Glucose-sensitive

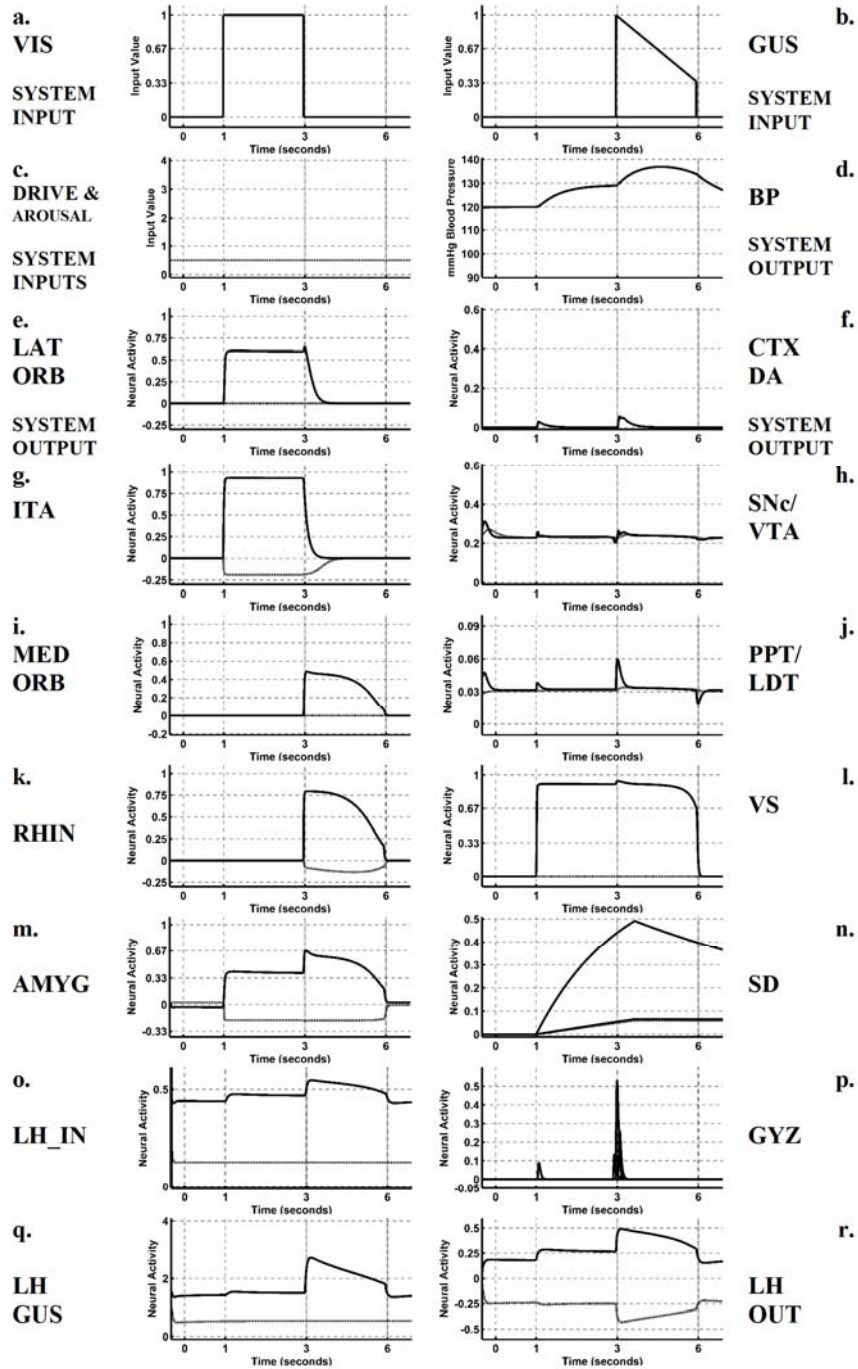
hunger inputs (upper, solid line) and satiety inputs (lower, dashed line) across the 20 devaluation trials. 13a(2): GUS habituation for sweet taste (lower, solid line) and lack of habituation for the unencountered salty taste (upper, dashed line). 13b(1): The response of US1-specific MORB cells to increasing satiety (solid line) and US2-specific MORB cells (dashed line). 13b(2): The differential response of US1 (solid) and US2-specific (dashed) MORB cells during the FSS simulation. 13c(1): The response of appetitive (solid line) and aversive (dashed line) glucose-sensitive LH\_out cells. 13c(2): The response of the glucose-sensitive LH\_out cell when US1 is presented (solid line) vs. when US2 is presented (dashed line). 13d(1): The decrease of the US-generated blood pressure response with increasing satiety. 13d(2): Blood pressure response to US1 (solid line) vs. US2 (dashed line). 13e(1): The average dopamine burst in response to US presentation across 20 satiation trials. 13e(2): Differential effective cortical dopamine response to US1 vs. US2.

## 2.5 CS Learning Task

*2.5.1 Normal Performance of the CS Task.* The model's performance of the CS Task can be parsed into nine phases. Fig. 14 details the responses of simulated neurons during the performance of the CS learning task. The first phase of the CS task is the initial equilibration of model variables prior to CS onset at  $t=1$  sec. In the second phase, the VIS input, CS1, is presented (Fig. 14a) to the CS1-specific ITA cell (Fig. 14g; see Equation (10)). The ITA cell activates CS1-specific lateral ORB cell (Fig. 14e; see Equation (11)) and, using previously learned connections, activates the US1-specific AMYG cell (Fig. 14m; see Equations (12-14, 38)). In the third phase, the CS expresses its conditioned reinforcer properties and is evaluated by AMYG-LH interactions. In particular, the US1-specific AMYG cells activate US1-prototypical taste and drive feature LH\_out cells (see Equations (17)). LH\_out cells (Fig. 14r) multiply the AMYG inputs by the current drive levels and relay this information back to the AMYG (Fig. 14m). The AMYG sums this information and generates an estimate of the momentary affective value of US1 and, consequently, CS1 (Fig. 14m). AMYG cell activity then drives the CS1-specific ORB cell via incentive motivational signals (Fig. 14e), which represents the object-value that controls approach to CS1.

In the fourth phase of the CS task, ORB activity modulates FEF and ITA activity (Fig. 14g) and thereby helps to control CS1-oriented eye movements and CS1-compatible visual attention (see Equation (46)). In addition the LH elicits a CS1-related blood pressure response (Fig. 14d; Equations (44-45)) which prepares the system for the action and its consequences. In the fifth phase, the activity elicited by CS1 in the AMYG, ORB, and LH is relayed to the BG where dopamine cells in the SNc/VTA respond to the appetitive value of the CS with a burst of activity. CS1-related value information is carried to the SNc/VTA via the PPTN/LDT. There are two paths through the BG leading to the PPTN/LDT, one monosynaptic and projecting directly, the other disynaptic and projecting indirectly.





**Figure 14: CS learning task, after ten conditioning trials: Key:** 14a: The CS1 visual stimulus. 14b: A sweet taste driven by US1-related GUS input. 14c: A glucose-deprivation driven hunger input (solid line) and a nonspecific arousal input (dashed line). 14d: Model blood pressure output. 14e: The CS1-specific ORB cell (solid line) and a CS2-specific ORB cell (dashed line). 14f: Effective cortical dopamine burst (solid line) and dip (dashed line). 14g: CS1-selective ITA cell (solid line) and CS2-selective ITA cell (dashed line). 14h: Phasic output of SNc/VTA cells (solid line) and time-averaged

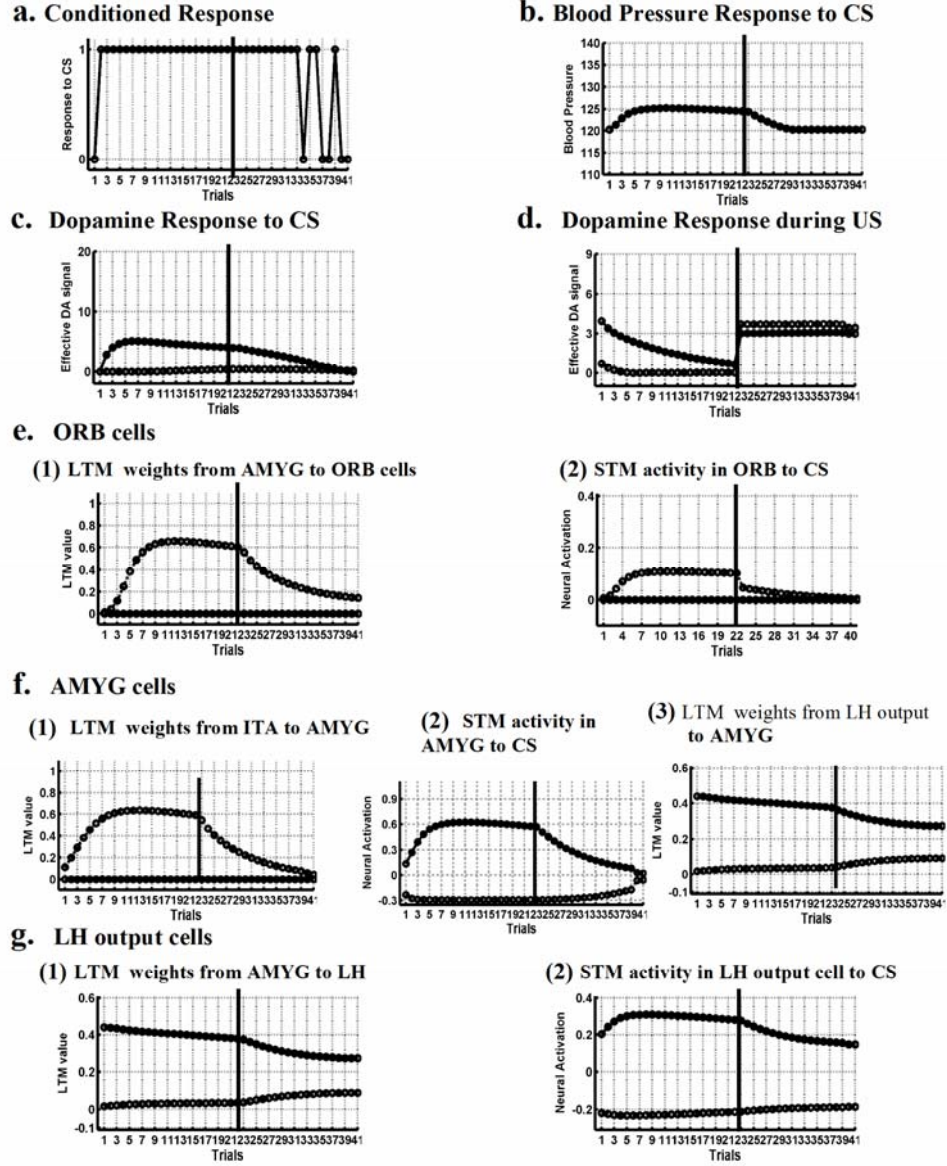
SNC/VTA output (dashed line). 14i: US1-specific MORB cell (solid line) and US2-specific MORB cell (dashed line). 14j: Solid lines trace the phasic output of the PPTN cell; dashed lines the hyperpolarization (suppression) of tonic PPTN inputs. 14k: Activation of the US1-selective RHIN cell (solid line) and US2-selective RHIN cell (dashed line). 14l: VS cell responsive to CS1 and US1 (solid line) and VS cell responsive to US2 (dashed line). 14m: US1-specific AMYG cell (solid line) and US2-specific AMYG cell (dashed line). 14n: CS1-related activation of SD cells with different decay rates. 14o: Glucose-deprivation driven LH\_in cell (solid line) and opponent LH\_in cell (dashed line). 14p: Adaptively timed output, GYZ, from SD cells to SNC/VTA cells. 14q: Activity of LH\_gus cell sensitive to sweet taste and glucose drive inputs (solid line) and opponent LH\_gus cell (dashed line). 14r: Sweet-taste and glucose-specific LH\_out cell (solid line) and aversive opponent LH\_out cell (dashed line).

The monosynaptic path carries mainly US-related value information from LH\_gus cells to the PPTN/LDT (Figure 14j; see Equation (21)). The first branch of the disynaptic path to the PPTN/LDT issues from the AMYG and ORB to the PPTN/LDT via the VS (Figure 14l) and carries the majority of CS-related value signals. The path from the ORB to the VS is learned, while the path from the AMYG to the VS is unlearned (see Equation (20)). AMYG inputs to the VS reflect the value of specific outcomes. In the second branch of this disynaptic projection, the VS inhibits tonic ventral pallidal (VP) activity, which has a net excitatory (disinhibitory) effect on the PPTN/LDT.

The ORB also projects to the SNC/VTA via a second, net inhibitory path. This second path relies on CS-related inputs from the ORB to activate SD cells (Figure 14n; see Equation (23)), which issue an adaptively-timed inhibitory input (Figure 14p) to the SNC/VTA (Figure 14h; see Equation (25-28, 44)).

In the sixth phase of the task, CS1 presentation terminates and the US is presented for two seconds. As in the Pavlovian task described by Ono et al. (1986a) the CS1 VIS input turns off and the US1 GUS input turns on for 2 seconds (Figure 14b). In the seventh phase, the US1 is evaluated as described in the US task (Figure 12). US-related dopamine signals are processed differently in the eighth phase of the task from that seen in the US task. In the eighth phase of the CS task, the inhibitory path from the ORB to the SNC/VTA plays a prominent role. This inhibitory path learns to generate an adaptively timed inhibition that suppresses US1-generated dopamine responses. The ability of SD cells to suppress US-related dopamine responses reflects the number of training trials and the strength of the predictive relationship between CS and US. After forty training trials, the model is typically fully trained. Even after ten trials dopamine spikes elicited during US presentation are greatly reduced (Figure 14f).

*2.5.2 Normal Learning of the CS Task.* The learning of the CS1+US1 association is depicted in the first twenty trials of each graph in Fig. 15. Subsequent trials show effects of extinction. If the previously model has been trained with the US Task (as here), the model learns most aspects of the CS Task within ~20 trials, during which most LTM adaptive weights reach a plateau. US-related weights change little during the CS Task (Fig. 14d(3), 14g(1)). An additional 20-30 trials are required for the plateauing of the adaptively timed weights,  $Z_{gi}$ , which gate the inhibitory inputs from SD cells to dopamine cells in the SNC/VTA (Equation 43).



**Figure 15: Trial level changes in learning and extinction of CS task:** Simulations involved 20 rewarded trials followed by 20 extinction trials. Data points in the plots represent the average activation of model variables either during CS presentation (time = 1-3 seconds) or US consumption (time = 3-5 seconds). For LTM weights, data points represent average LTM values across the trial (time = 1-6 seconds). During extinction trials, a strong dopamine dip (above a threshold  $D_2 = 0.2$ ) elicits an arousal burst. 15a: A saccadic response was (“1”) or was not (“0”) elicited by CS1 during each trial. 15b: Average blood pressure response to CS1. 15c: Average CS1-generated dopamine responses: Dopamine bursts (solid lines) and dips (dashed lines). 15d: Average US1-generated dopamine responses: Dopamine bursts (solid lines) and dopamine dips (dashed lines). 15e(1): LTM weights linking the US1-specific AMYG cell to the CS1-specific ORB cell (solid line), and (dashed lines) LTM weights linking US1-specific AMYG cells with CS2-specific ORB cell. 15e(2): Average CS1-specific ORB response (solid lines),



and (dashed lines) the response of the CS2-specific ORB cell . 15f(1): LTM weights (solid line) linking the CS1-selective ITA cells to US1-specific AMYG cells, and LTM weights (dashed line) from the CS2-selective ITA cell to the US1-specific AMYG cell. 15f(2): Average response of US1-specific AMYG cells during CS1 presentation (solid lines), and (dashed lines) the response of US2-specific AMYG cells to CS1 presentation. 15f(3): LTM weights (solid line) linking the glucose-specific LH output cells with the US1-specific AMYG cells, and LTM weights (dashed line) from the sodium-specific LH\_out cell to the US1-specific AMYG cell. 15g(1): LTM weights from the US1-specific AMYG cells to the glucose-specific LH output cells (solid lines), and (dashed line) LTM weights from the US1-specific AMYG cell to the sodium-specific LH output cells. 15g(2): Average glucose-specific LH output cell response to CS1 (solid line), and opponent LH cell response to CS1 (dashed line).

CS1 forms a specific association with US1 using conditioned reinforcer adaptive weights from the ITA to AMYG (Fig. 15f(1); see Equation (38)), with incentive motivational adaptive weights from the AMYG to ORB allowing needed rewards to modulate available stimuli (Fig. 15e(1); see Equation (36)), with Now Print adaptive weights from the ORB to VS linking motivationally relevant cues to dopamine cells (see Equation (42)), and with adaptively timed weights from the ORB to SD enabling previously attended cues to establish timed expectations of reward (see Equation (43)). The growth of these LTM weights boosts STM activity over the course of training, elevating AMYG, ORB, LH\_out activity, and via connections with the VS and SNc/VTA, dopamine activity (Fig. 15f(2), 15e(2), 15g(2), 15c). Rises in LH activity boost blood pressure (Fig. 14b). A strong adaptively-timed inhibitory signal from the SD to the SNc/VTA assures that a dopamine dip will be generated if a reward is omitted (e.g. Fig. 16f). Complete suppression of US-generated dopamine signals happens after about thirty trials. Further training, on the other hand, does not suppress CS-generated dopamine signals (Fig. 15c). Inhibition of US1-generated dopamine responses by SD cells, and small decrements in drive and GUS inputs (Fig. 15d) slightly decrease LTM values after the plateau value has been reached.

*2.5.3 Extinction of Pavlovian Stimuli.* During an extinction trial, the CS is presented but the US is omitted. Fig. 15 illustrates the decay of LTM weights. The vertical line in the panels of Fig. 15 separates learning trials from extinction trials. Learning at incentive motivational (see Equation (36)) and LH to AMYG (see Equation (37)) LTM weights is gated by postsynaptic activity and samples presynaptic activity. Learning in conditioned reinforcer (see Equation (38)), basal ganglia (see Equations (42, 43)), and AMYG to LH (see Equation (40)) LTM weights is gated by presynaptic activity and samples postsynaptic activity. Extinction trials diminish the predictive significance of the CS and lead to the decay of learned CS1+US1 associations but not to the decay of associations relating to the US itself. LTM weights decrease according to two processes: when the activity they sample decreases (see Equations (6, 7)) or when dopamine dips actively trigger weight decay (see Equations (8, 9)).

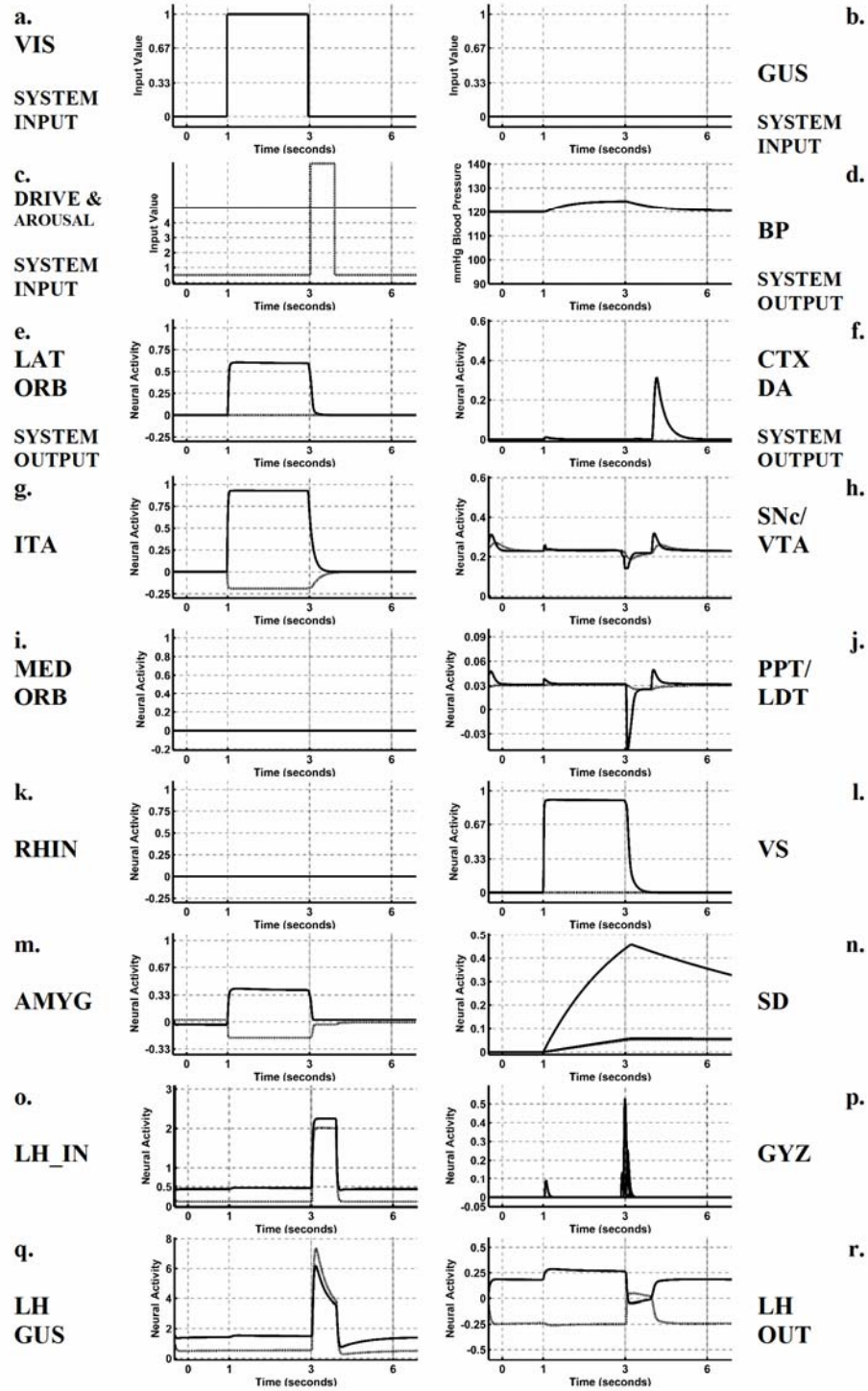
CS-related LTM weights decay to differing degrees during extinction trials. Conditioned reinforcer weights linking the ITA and AMYG extinguish completely (Fig. 15f(1)). Incentive motivational weights linking the AMYG and ORB decay incompletely (Fig. 15e(1)). Conditioned reinforcer weights extinguish completely because learning along this path is gated by ITA activity (Equation 38) and ITA activity remains strong throughout the CS extinction task, ensuring dopamine dips can drive weight decay on every extinction trial. The decay of incentive motivational LTM weights reaches an asymptote because ORB activity gates learning and ORB

activity decreases across extinction trials (Figs. 15e(2)). This decrease diminishes LTM weight plasticity, sparing incentive motivational weights from complete decay (Fig. 15e(1)). This helps explain some savings observed when an extinguished association is relearned (Rescorla, 2001).

LTM weights from the RHIN to the AMYG and AMYG to MORB carry US-related information and are spared from decay because these synapses are not active during CS extinction trials (see Equations (39, 41)). LTM weights between the AMYG and LH (Fig. 15f(3), 15g(1)) do not extinguish because recurrent connections between the LH and AMYG afford a persistent low level of activation that sustains the basic pattern of weights. In addition, when an arousal burst is unleashed by a dopamine dip in the SNc/VTA, the LH gated dipole circuit is reset, terminating activity in the AMYG, thereby preventing weight decay.

*2.5.4 Normal Performance of CS Task Extinction Trial.* The normal nine-phase process of CS task performance is cut short during the CS extinction task. Normally, the US is presented immediately after the CS input is terminated. A trace of the CS remains active in the ITA (Figure 16g) while the US is presented, allowing CS-selective cells in the ITA to sample US-related activity in the AMYG. Omitting the presentation of the US eliminates US-related activity in the RHIN, MORB, AMYG, LH\_out, and eliminates US-generated blood pressure (BP) changes (Fig. 16i, 16k, 16m, 16o, 16q, 16r). Hence, during extinction trials, ITA cells have no US-related AMYG activity to sample. Furthermore, when the US is omitted, reward-related inhibition from SD cells causes SNc/VTA cells to generate a large dopamine dip signal that is broadcast to other brain areas (Fig. 15d). When the magnitude of this dopamine dip surpasses a threshold value, it triggers an arousal burst in the LH (Fig. 16d; see Section 4.4.4). The dopamine dip accelerates the decay of AMYG, ORB, and ITA connections, while the arousal burst causes an antagonistic rebound in the AMYG that resets the activity in the AMYG and LH. The arousal burst resets AMYG activity to zero, limiting the decay of weights between the AMYG and LH and speeding the decay of conditioned reinforcer weights (Fig. 15f).

*2.5.5 CS Satiety-Dependent Devaluation.* Fig. 11 shows model simulations where hunger inputs were systematically reduced and GUS input systematically habituated (Figs. 11a and 11b). Reducing the hunger drive inputs to LH\_in cells lowers the tonic or ‘resting’ activity of these, LH\_gus, and LH\_out cells. As a consequence of feedforward opponent inhibition from LH\_gus to LH\_out cells, an increase in satiety inputs to LH\_in cells also suppresses hunger-sensitive LH\_out cells (Fig. 8; see Equation (17)). Lower LH\_out cell activity means lower AMYG cell activity, and ultimately ORB cell activity. Hence, responses of LH\_out, ORB, and AMYG cells decrease with increasing satiety, showing that CS value is sensitive to satiety (Figs. 11b(1), 11c(1), 11d(1)). Diminished LH\_gus and LH\_out cell responses to CS presentation also translate into decreased dopamine and blood pressure responses to CS presentation (see Equations (21, 28, 44-45); Figs. 11e(1), 11f(1)). CS-elicited activity in the ITA is little changed by satiety or hunger because no competing stimuli are presented as part of the CS task and motivation-related inputs from the ORB primarily act to suppress competing visual stimuli (see Equation (10)).



**Figure 16: CS learning task, extinction trial with arousal burst:** Model responses during an extinction trial. See Fig. 14 for description of variables and graphs.

**2.5.6 Simulations of Outcome-Specific Devaluation of CS.** The second column of Fig. 11 shows simulations of food-specific and outcome-specific CS devaluation experiments. Two

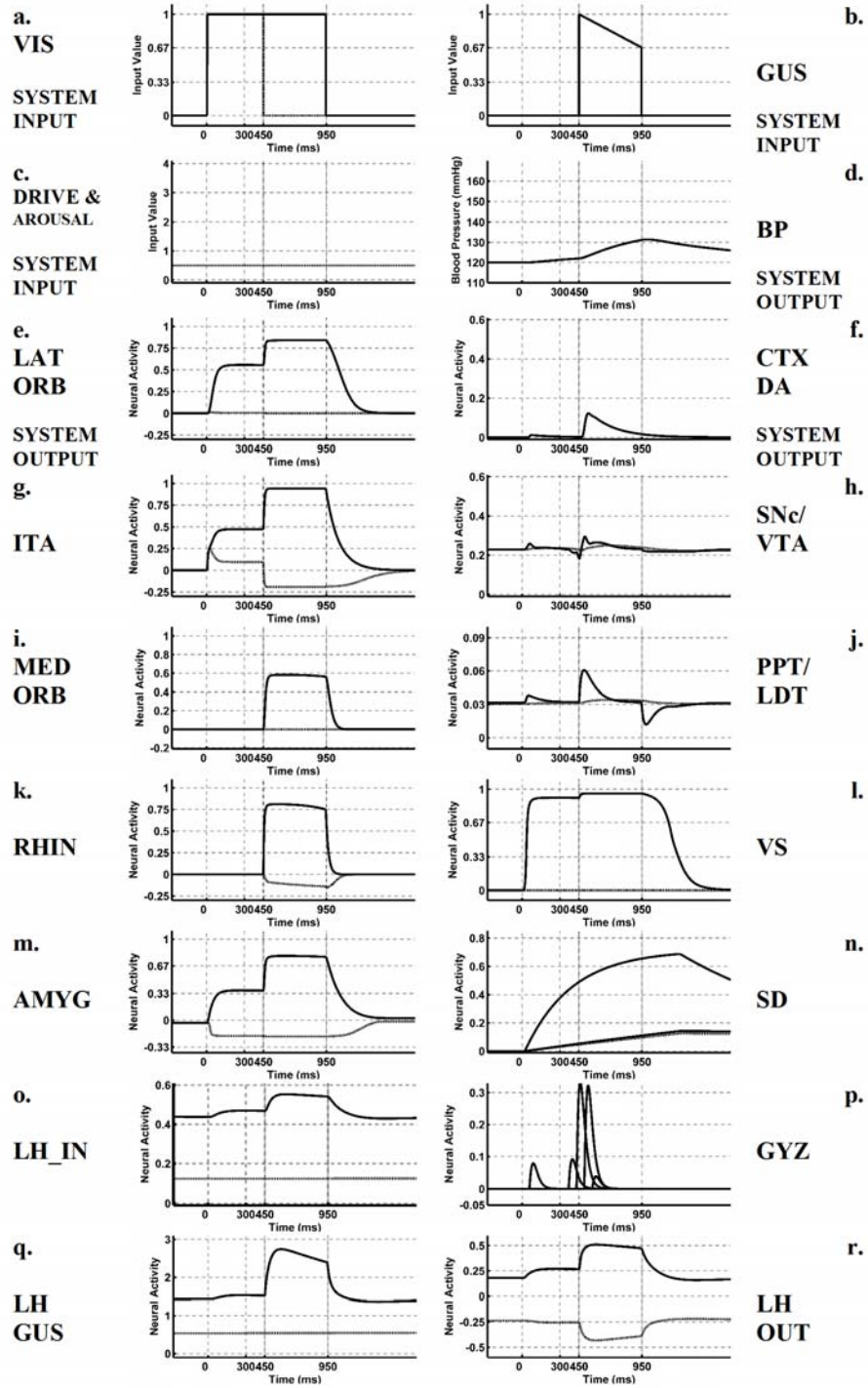
experiments were simulated: the first experiment (Fig. 11a(2)-f(2)—dashed lines) showed the model response to the stimulus pair CS2+US2 after US2 had been specifically devalued by changing the hunger and satiety inputs to the LH to responses to the taste-drive properties of US2. The second experiment (Fig. 11a(2)-f(2)—solid lines) used these same hunger and satiety inputs only this time testing the response of the model to a different stimulus-reward pair, CS1+US1, where US1 has one taste drive feature the same and one taste-drive feature different from US1 (see Section 4.2.2). Fig. 11b(2) shows that the ORB response to CS2 was automatically and specifically devalued, while the ORB cell response to CS1 was relatively spared by the change in satiety.

The difference in ORB cell responses reflects the different values of US1 and US2 predicted by drive-value category cells in the AMYG (Fig. 11d(2)). Drive-value category cells in the AMYG are activated whenever a CS is presented and calculate the value of the CS by modulating and summing the taste-drive feature inputs from LH\_out cells. These cells sum LH\_out activity whenever the CS or US is presented. The hunger and satiety inputs with which the simulations were run result in a reduced response of LH\_out cells to taste-drive features associated with US2, but not for all of the taste-drive features associated with US1 (Figs. 11c(2)). As a result of the different levels of activation of taste-drive feature cells in the LH during CS or US presentation, the activation of drive value category cells in the AMYG to either CS1 or US1 is much greater than the activation of these cells to CS2 or US2 (Fig. 11d(2)). As a result of this, MORB cells also respond more to US1 than US2 during reward consumption. Owing to the habituation of US2-related GUS taste features, RHIN cells, which respond to GUS inputs, respond more strongly to US1 than US2.

## 2.6 SVD Task

**2.6.1 Mechanisms in Normal Performance.** The model executes the SVD (simultaneous visual discrimination) task in 8 stages: (1) Initialization; (2) dual CS presentation; (3) dual CS valuation; (4) response generation and attentional modulation; (5) CS-related dopamine processing; (6) US presentation; (7) US valuation and response generation; (8) US-related dopamine processing. Fig. 17 details these stages and the responses of simulated neurons during the performance of the SVD task.

**2.6.2 Normal Learning of SVD Task.** Fig. 18 shows the learning and extinction of the SVD task. During the first 15 trials, the model learns the CS1+ vs. CS2- discrimination and strengthens the LTM weights from ORB to SD that inhibit US-elicited dopamine responses (Figs. 18(a), 18(d)). Prior to this simulation, the model was trained to recognize US1 and US2 as part of the US Task. After 5 SVD trials, ITA and ORB responses reliably discriminate between a target stimulus (CS1+) and a distractor (CS2-) (Fig. 18e(2)). Selection of the target (CS1+) over the distractor (CS2-) relies on changes in the values of LTM weights from the ITA to AMYG (see Equation (38)) and AMYG to ORB (Figs. 18e(1), 18f(1); see Equation (36)). CS1+US1 associations are reinforced by reward-related dopamine bursts. With every pairing of CS and US, weights elicit stronger responses to CS1+ in the AMYG, ORB, and LH output cells. After 4-5 trials, systemic blood pressure responses occur in response to the presentation of CS1+ (Fig. 18b). These stimulus-elicited changes in blood pressure are driven by inputs from the LH to cardiovascular regions in the medulla. CS1+-elicited dopamine responses occur after 5 trials (Fig. 18c). Over many trials, US-dependent dopamine bursts are suppressed by learned inhibitory inputs from the SD to SNc/VTa (Fig. 18d).



**Figure 17: STM activity during the performance of the SVD task:** During the SVD task, two stimuli were presented: CS1+ and CS2-. CS1+ is the discrimination target and the CS1+US1 association was reinforced. Saccades are elicited from the FEF using a threshold rule (see Equation (47)). At  $t=0$  sec., CS1+ and CS2- are simultaneously presented as VIS inputs (17a) to ITA cells (17g). ITA cells then project to the AMYG (17m) and ORB (17e). ITA representations of CS1+ and CS2- activate the AMYG via

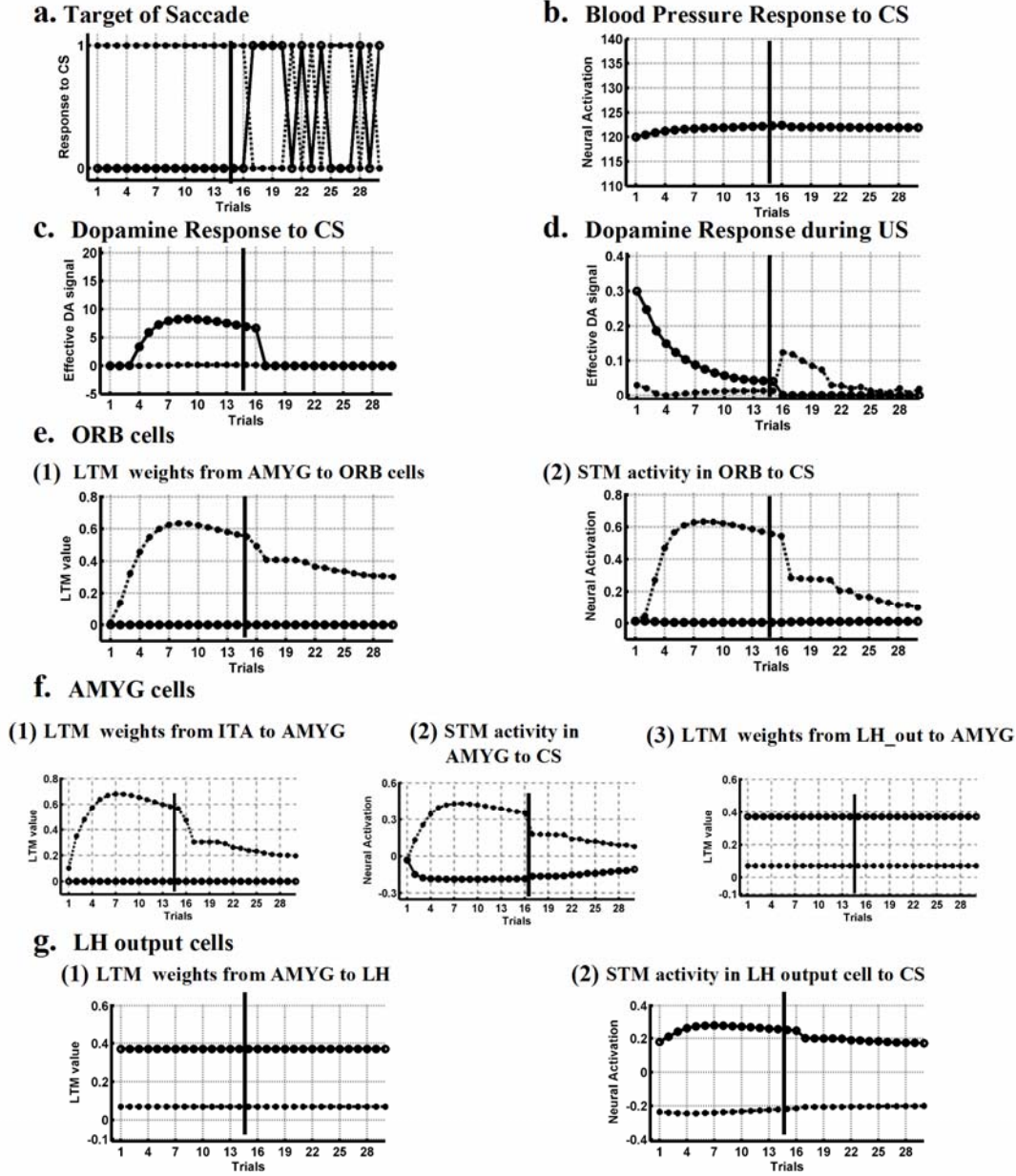
conditioned reinforcer weights. A strong CS1-US1 association activates the US1-specific AMYG cell (17m) and AMYG-LH interactions (17o, 17q, 17r) calculate the expected value of US1. This drive value estimate is relayed to the ORB (17e). Incentive motivational weights connecting the AMYG to the ORB activate the CS1-specific ORB cell (17e). CS1-specific ORB cells (17e) modulate the salience of ITA (17g) and FEF cells. CS1 drives changes in blood pressure via LH activation (17d). Between 300-450ms a saccade is elicited and CS2- terminates. CS1-related AMYG and ORB activity also elicits dopamine spikes (17f) from the SNc/VTA (17h) via the VS (17p) and PPT/LDT (17j). CS1-specific ORB cell also activates SD cells (17n). At t=450ms, US1 is presented as a GUS input (17b) for 500ms. US1 presentation overlaps with CS1 (17b, 17c). US1-related GUS inputs activate LH\_gus (17q) and RHIN (17k) cells. The consumption value of US1 is calculated by AMYG-LH interactions (17m, 17o, 17q, 17r) and registered as incentive value by ORB (17e) and MORB (17i) cells. ORB cells affix this motivational value to CS1+ (17e). US-generated signals activate blood pressure via the LH (17d). Lastly, the AMYG and LH generate a US1-related dopamine burst (17f) from the SNc/VTA (17h) via the VS (17p) and PPTN (17j). US-generated dopamine spike (17f) is partially suppressed by inhibitory inputs from SD cells (17n, 17p). Descriptions of STM activities (17a-17r) are as Fig. 14.

*2.6.3 Extinction in the SVD Task.* SD cells learn a temporal expectation of reward and inhibit the SNc/VTA during the usual time of reward. On extinction trials where the reward is omitted, the result is a dip in dopamine activity. Fig. 18d shows that the first extinction trial elicits a dopamine dip. This dopamine dip gates weight decay resulting in a substantial reduction in the strength of LTM weights from the ITA to AMYG and from the AMYG to ORB (Figs. 18e(1), 18f(1); see Equation (36, 38)). As with the CS extinction task, the LTM weights from LH to AMYG, and AMYG to LH also slightly decay during SVD extinction trials (Figs. 18f(3), 18g(1); see Equation (37, 40)). These decreases in LTM weights between the AMYG and LH lead to a reduction in reward and CS-related blood pressure signals, dopamine output, and STM activity in the ORB, LH\_out, and AMYG cells (Figs. 18b, 18c, 18e(2), 18f(2), 18g(2)). As a consequence of reduced ORB activity, the selection of targets for saccades by the FEF becomes more variable (Fig. 18a; see Equation (46)).

As with CS extinction trials, SVD extinctions spare the LTM weights from RHIN to AMYG and AMYG to MORB from decay. These LTM weights are spared from decay because reward-related GUS signals gate the plasticity of these weights and no US-related signals are presented during SVD extinction trials.

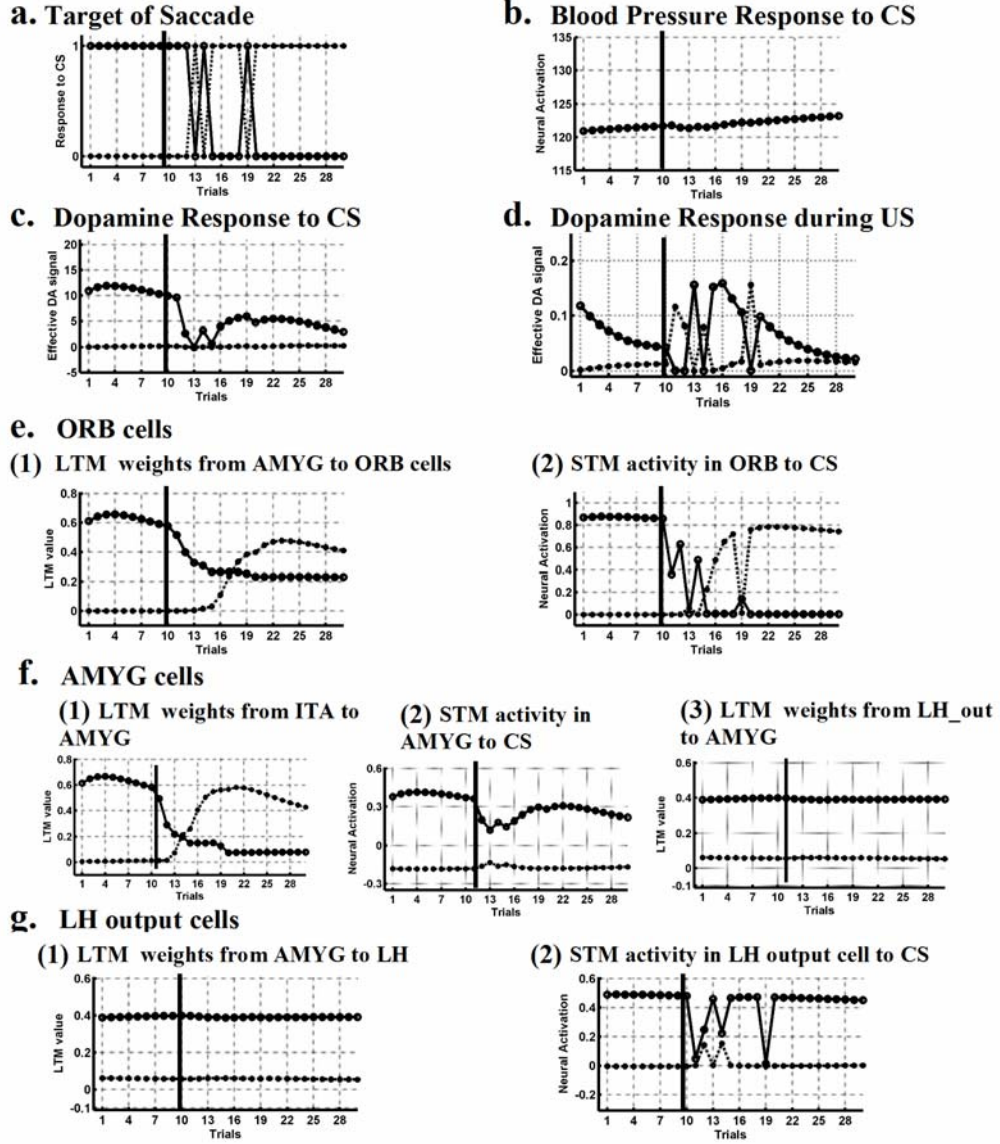
*2.6.4 Extinction Trial Mechanisms.* As with CS extinction trials, SVD extinction trials rely on three mechanisms for the decay of learned associations: (1) a decrease of neural activity in the AMYG and LH as a result of reward omission, (2) dopamine dips elicited from the SNc/VTA during nonreward and broadcast throughout cortex, speeding extinction by gating the decay of weights, and (3) an arousal burst input the LH that resets activity in the AMYG-LH system, zeroing out AMYG activity. Arousal burst and dopamine dip signals are triggered by the adaptively timed inhibitory inputs from the SD to the SNc/VTA. The net result is the decay of dopamine-gated weights during the 500ms period of nonreward.





**Figure 18: Learning and extinction of SVD task:** During this SVD simulation, two stimuli were presented for 30 trials: CS1+ and CS3-. The CS1+US1 association is reinforced for the first 15 trials. After trial 15, no rewards are given. Saccades were elicited from the FEF using a threshold rule that measured ORB activity (see Equation (47)). With the exception of 18d, data points represent the average activation of model variables during stimulus presentation (time = 0–450 ms). Data points in 18d represent average activity during the reinforcement period (time = 450–950 ms). For LTM weights, data points represent average LTM values across the interval from  $t=0$  until  $t=950$ ms. As with the CS task, an arousal burst can be triggered by a dopamine dip ( $D_2$  threshold of 0.2). The model could generate a dopamine dip that would elicit an arousal burst within 10–15 trials. After 5 extinction trials the CS1+US1 association appears extinguished and

discrimination behavior drops to chance. Detailed descriptions of traces and variables in graphs (18a-18f) are listed in Fig. 15.



**Figure 19: Reversal of SVD task:** The SVD reversal task involves reversing the learned stimulus-reward contingencies of two stimuli, CS1 and CS2. 30 trials are simulated. For the first 10 trials, the CS1+US1 association was trained. In the last 20 trials reinforced a CS2+US1 association. This sequence constituted one reversal. For LTM weights, data points represent the average value of LTM weights from time = 0–950 ms. For STM activity, data points represent the average activity of cells during target CS presentation (time = 0–450 ms) or US consumption (time = 450–950 ms). Arousal bursts are triggered by a dopamine dip ( $D_2=0.2$ ). Saccades were elicited from the FEF by ORB activity (Equation (46)). See Fig. 15 for a description of the graphs (19a-19g).

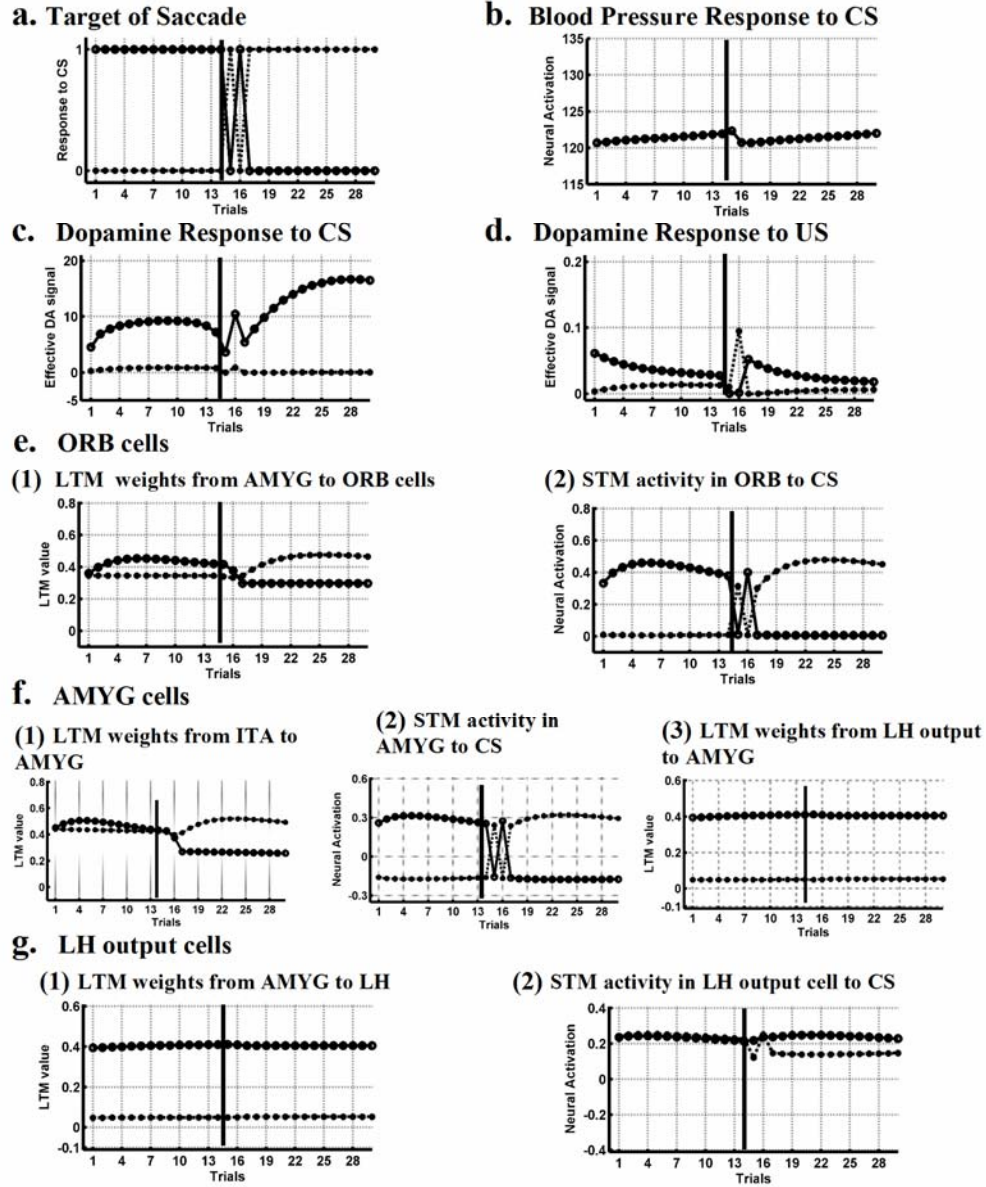


**2.6.5 Reversal Trials.** Fig. 19 shows simulation results from the SVD reversal experiment. Thirty trials were simulated. The first ten trials reinforced a CS1+US1 association. As is noted in the section on the SVD learning task, LTM weights between the AMYG, ITA, and ORB plateau quickly, within 5-10 trials (Figs. 19e(1), 19f(1), 19g(1)). The rate at which dopamine bursts are suppressed across training trials reflects the ability of LTM weights between the ORB and SD to learn the expected time of reward (Fig. 19d). US-elicited dopamine responses require 5-10 trials before they are completely suppressed (Fig. 19d).

The discrimination problem was between CS1 and CS2. In the last 20 trials of the reversal experiment, the CS2+US1 contingency was reinforced. During trials 11, 12, 14, and 19 the model selects the unrewarded stimulus, CS1-. Thus no reward is given during trials 11, 12, 14, and 19 and dopamine dips and arousal bursts are evident (Fig. 19d). These dips and bursts drive a decrease in the LTM weights from the ITA to AMYG and AMYG to ORB (Figs. 19e(1), 19f(1)). In tandem, there were decreases in CS-related blood pressure, dopamine and STM activity in the ORB, LH\_out, and AMYG (Figs. 19b, 19c, 19e(2), 19f(2), 19g(2)). From trial 20 on, the CS1- vs. CS2+ decision is reliably discriminated and the CS2+US1 association reinforced. These trials enhance the representation of CS2 in the AMYG, ORB, and LH\_out, and enhance CS2-related dopamine activity (Figs. 19c, 19e(2), 19f(2), 19g(2)). A differential response to CS2+ over CS1- also emerges in the ITA that is driven by ORB top-down modulatory inputs. In contrast, LTM weights related to US1 processing change little after reversal (Fig. 19f(3), 19g(1)).

**2.6.6 SVD Task: Alternation.** In the alternation experiment, thirty trials were simulated. The first fifteen trials reinforced a CS1+US1 association while the last fifteen trials reinforced a CS2+US2 association. Fourteen of fifteen CS1+ vs. CS2 discrimination trials were correctly performed (Fig. 20a). On trial 16, the reward contingencies altered so that CS1 was no longer rewarded and CS2 was reinforced by US2. One perseveration error was committed after the change in contingency. This perseveration error is accompanied by a dopamine dip, which causes the decay of dopamine-gated LTM weights linking the ITA to the AMYG and AMYG to ORB (see Equations (8, 9, 36, 37)). After the decay of these LTM weights, behavior alternated and CS2+ was reliably chosen over CS1- (Figs. 20d, 20a). New learning by the LTM weights linking ITA to AMYG and AMYG to ORB (Figs. 20e(1), 20f(1)) enables the model to reliably select the target (CS2+) over the distractor (CS1-). These new LTM values drive changes in blood pressure, dopamine activity, and STM responses in the AMYG, ORB, and LH output cells (Figs. 20b, 20c), 20e(2), 20f(2), 20g(2)).

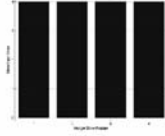
**2.6.7 SVD-FSS Task: Cue Preference and Revaluation.** In the SVD-FSS task, 10 trials were simulated. All the trials were run in extinction, in that no rewards were given and all stimuli were trained prior to the task. There were 5 ‘before’ and 5 ‘after’ trials simulated. In the before trials, drive inputs were unsated. In the after trials, drives were sated.



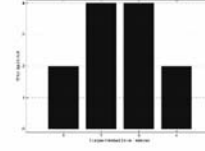
**Figure 20: Alternation of SVD task:** In the SVD alternation task, CS1+US1 and CS2+US2 associations were pre-trained prior to alternation simulations. 30 alternation trials were simulated. The first 15 trials reinforced a CS1+US1 association while the last 15 trials reinforced a CS2+US2 association. The shift from the CS1+ vs. CS2- to the CS2+ vs. CS1- discrimination constituted one alternation. For LTM weights, data points represent the average value from  $t = 0-950$  ms; for CS-related STM activity data points represent the average activity of cells activated by the target CS from  $t = 0-450$ ms; for US-related neural activity data points represent the mean neural activity from  $t = 450-950$ ms. As in other tasks, an arousal burst can be triggered by a dopamine dip and saccades were elicited from the FEF by ORB activity (Equation (46)). See Fig. 15 for a description of the graphs (20a-20g).

### a. Satiation Sequence Inputs

#### (1) DRIVE Inputs Pre-Satiety

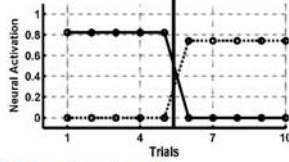


#### (2) DRIVE Inputs Post-Satiety

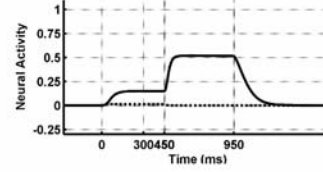


### b. ORB cells

#### (1) Preference between CS1 and CS2

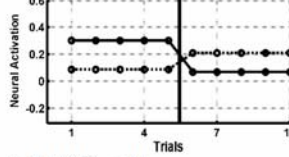


#### (2) Differential Satiety Response

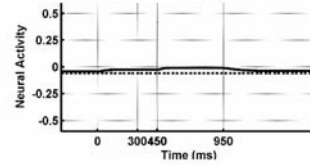


### c. LH output cells

#### (1) Preference between CS1 and CS2

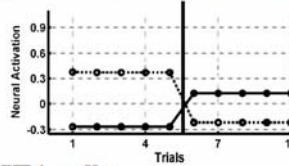


#### (2) Differential Satiety Response

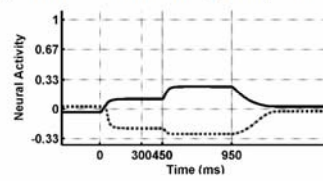


### d. AMYG cells

#### (1) Preference between CS1 and CS2

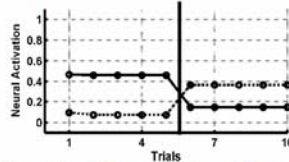


#### (2) Differential Satiety Response

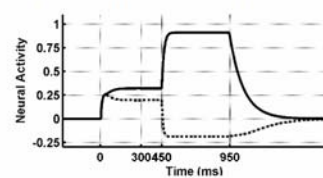


### e. ITA cells

#### (1) Preference between CS1 and CS2

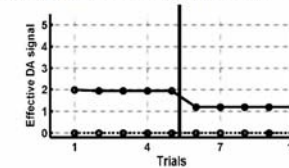


#### (2) Differential Satiety Response

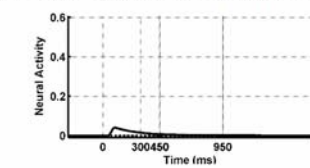


### f. Cortical Dopamine Output

#### (1) Response to Satiety Sequence



#### (2) Differential Satiety Response



**Figure 21: Satiety-related modulation of CS preference in SVD-FSS task:** The SVD-FSS task demonstrates cue preference can shift as a result of selective satiation without any additional reinforcement learning. Two visual discriminations were trained prior to the SVD-FSS task, resulting in the learning of CS1+US1 and CS2+US2 associations. Data points represent the average target CS-related STM activity during the interval from  $t = 0$ -450ms. 5 ‘before’ and 5 ‘after’ trials simulated. For the first 5 trials,

appropriate rewards were given on each trial and the DRIVE and GUS inputs were high (set to 4 and 1, respectively); for the last 5 trials, no rewards were given and the model inputs reflected US1-specific satiety (US1-related GUS inputs were set to a habituation level of 0.4 and US1-related DRIVE inputs were set to 2). LTM weights were held fixed for all ten trials. Except for graphs in parts 21a and 21e traces and figures are as in Fig. 16. Traces in the right hand column were sampled from the last 5 trials. 21a(1): Represents the drive level before satiation. 21a(2): Represents the drive levels after satiation—the sugar and protein inputs were sated. 21e(1): Shows the response of the model to the simultaneous presentation of CS1 (solid lines) and CS2 (dashed lines). 21e(2): The differential response of CS1-specific (dashed) and CS2-specific (solid) ITA cells.

For the first 5 ‘before’ trials, there was no initial GUS habituation, and the specific hunger drive inputs to drive-sensing LH\_in cells were uniformly set to an initial value of 4 while satiety drives were set to 0 (Fig. 21a(1)). In response to these initial drive levels, the model selected the CS2+US2 stimulus pair (Fig. 21b(1)). These trials demonstrated that the initial conditions of the model favored CS2 over CS1. Thus, CS-related dopamine signals and the STM response of the LH, AMYG, ORB and ITA were greater to CS2 than CS1 (Figs. 21d(2), 21e(2), 21f(2), 21c(2), 21b(1)).

For the last 5 ‘after’ trials, a food-specific satiety was simulated by reducing two hunger drives while leaving the other hunger drives elevated. In particular, the hunger drives representing protein levels (umami taste) and sugar levels (sweet taste) were reduced to a value of 2.2 while the drive inputs representing fat and salt remained high, coded by a value of 4. These hunger inputs were typically coded as the input vector [2.2, 4, 4, 2.2] (Fig. 21a(2)). Satiety drive inputs were uniformly set to 2.2 during the last 5 ‘after’ trials. The result of this shift in drive levels is a shift in model CS preference.

The model now prefers CS1 over CS2, as indicated by ORB cell activity in Figs. 21b(1) and 21b(2)). The ORB cell responses to each CS reflects the drive value of the rewards associated with each of the stimuli. The drive value of these rewards is calculated by drive value category cells in the AMYG. Figs. 21d(1) and 21d(2) show the responses of US1 selective drive value category cells (dashed lines) and US2 selective drive value category cells (solid lines). The activation of drive-value category cells in the AMYG reflects the amount of reward-compatible drive-related activity in the LH. Uniform drive inputs favor the selection of the AMYG cell selective for US1 (Fig. 21d(1)—first five trials). Substantially reducing drive inputs associated with US1 results in less activity in the LH that is compatible with US1 and a stronger activation for the AMYG cell selective for US2 (Figs. 21c(1), 21d(1)—last five trials). A smaller activation for US1-related drives means a smaller activation for CS1-related dopamine signals and a greater STM response to CS2 than CS1 in the ORB and ITA (Figs. 21f(1), 21b(1-2), 21e(1-2)).

### 3. DISCUSSION

This article explicates neural information processing and learning principles and circuits that guide the reactive and predictive processing of food and food-related stimuli. It does so by showing how the MOTIVATOR model simulates observed changes in blood pressure, saccadic reaction times, learning, extinction, reversal, choice behaviors, and electrophysiological responses of identified cell types in multiple brain regions that are devoted to cognitive-

emotional processes. The consequences of lesions to model AMYG and ORB components are detailed in a companion article (Grossberg, Bullock & Dranias, 2008).

### **3.1 Discussion of Results**

**3.1.1 Blood Pressure Response.** The model successfully reproduced dynamic neural antecedents to systolic blood pressure responses (BPRs) during CS and US tasks (Nakamura et al., 1992). The model also replicated observations showing that the blood pressure response is conditionable (Braesicke et al., 2005). In addition, the model predicts that the magnitude of BPRs tracks reward magnitude. The model further predicts that the expression, but not the learning, of BPRs should be DA-independent. BPR expression is DA-independent because glutamatergic pathways are predicted to be sufficient to carry signals from the ITA or AMYG to the LH and then on to cardiovascular regulatory regions in the medulla (Fig. 8). This makes blood pressure a candidate for measuring affect following ventral pallidal (VP) lesions or 6-OHDA lesions (Smith and Berridge, 2005). If BPRs persist following VP and DA manipulations, this may serve to verify that the DA-independent ‘liking’ system extends into midbrain and amygdala structures.

**3.1.2 Influence of Satiety on Value.** The model’s response to devaluation, in the CS-FSS (food-specific satiety) and US-FSS tasks, replicates reported behavioral patterns and electrophysiological neural responses (Rolls et al., 1986; Scott et al., 1995; Critchley and Rolls, 1996; Rolls et al., 1999). Experiments that assay behavioral responses to increasing satiety reveal that animals quickly switch from acceptance to rejection of food rewards such as juice. In particular, experiments that assay neuronal responses during trials where juice is consumed with increasing satiety have found that neurons have a logistically decreasing response to US or CS presentation as satiety increases (Rolls et al., 1986; Critchley and Rolls, 1996). The model exhibits the FSS property in that both simulated and electrophysiologically recorded neurons in the AMYG, LH, ORB have diminished responses to a sated food, but preserved responses to an unsated food (Rolls et al., 1986; Nishijo et al., 1988a). At a behavioral level, animals show depressed Pavlovian responding to a CS that has been selectively devalued (Hatfield et al., 1996). In the model, simulated responses and neural activity also decrease in an outcome-specific fashion (Fig. 11).

In the SVD-FSS task, automatic shifts in CS preference are reported in experiments (Baxter et al., 2000) and replicated here by simulations (Fig. 21). No data are available on neural responses or BPRs during performance of the SVD-FSS task, but the model makes predictions of neural activity based on the same parameters that allowed the model to fit neural activity during the CS, US, and SVD tasks. In particular, the model predicts different AMYG cells respond to stimuli on the basis of the drive value of expected rewards (Fig. 11d), that the selection of a saccade target leads to the suppression of competing cell activity in the ITA (Fig. 17g), that the degree of suppression in ITA reflects the drive value of the chosen stimulus (Fig. 21e), that dopamine burst magnitude will automatically reflect drive value of stimuli and rewards (Fig. 21f), and that ORB cells will show an automatic, cue-selective, and outcome-specific reduction in neural responses following food-specific satiety (Fig. 21b).

**3.1.3 Saccade Latency.** Visually selective neurons in the FEF respond to targets for saccades and are reliable indicators of a forthcoming saccade (Bichot et al., 2001; Schall and Thompson, 1999). Saccadic response time (SRT) is correlated with the time it takes a visually-selective FEF cell to respond to a target in its receptive field during visual discrimination (Sato et al., 2001). Numerous studies have demonstrated that target salience is one of the attributes that can speed the discrimination of targets from distractors (Lauwereyns et al., 2002; Roesch and Olson, 2004). Incentive motivation also impacts SRT, but the timing of instrumental saccades

involves dorsal striatal mechanisms (Corbit et al., 2001; Brown et al., 2004). The model predicts that part of the difference in reaction time (RT) observed during responses toward stimuli associated with large and small rewards is due to stimulus-outcome associations that are established by Pavlovian conditioning. Simulations of the SVD-FSS task show how these Pavlovian associations can increase the salience of target stimuli, speeding the time to decision for saccade tasks, replicating RT data (Fig. 3b). Section 2.3.3 and the Section 4.6.2 detail how the model generates saccades using ORB cell activity, a cumulative spike counter, and a ‘race to threshold’ rule. Saccadic reaction times are sensitive to motivational level because ORB cell activity reflects motivational value, thereby motivational value directly influences when cumulative ORB activity exceeds saccade threshold.

*3.1.4 Learning and Memory Results.* Simulations of the CS task qualitatively replicate the exponential growth and decay curves of learning and extinction (Fig. 15). The model matches properties of the discrete time Rescorla-Wagner model of conditioning (Wagner et al., 1980); e.g., larger rewards speed learning with diminishing returns. In addition, the model accounts for features of conditioning that the Rescorla-Wagner model of conditioning cannot, such as the ability of more salient cues to overshadow less salient cues and the effects of interstimulus interval on conditioning (Grossberg, 1972b).

In the SVD task, the model qualitatively matches the behavioral results reported by Jagadeesh et al. (2001). Both the model and the data show that a discrimination problem is learned faster than a reversal (Fig. 19). During the SVD task, both the data by Jagadeesh and the model show that ITA responses to the CS- and CS+ become differentiated within the first ten learning trials (Fig. 19). During extinction or reversal, the model produces a perseverative pattern of errors, similar to the pattern of errors reported during reversal trials by Jagadeesh et al. (2001). In particular, the data shows a perseveration of responding for the previously rewarded stimulus during the first 5 trials of the reversal task. In the simulation there is a similar initial perseveration. Perseverative errors are also observed in reversal studies by Rolls (e.g., Rolls et al., 1996). In the MOTIVATOR model, this initial stage of perseveration reflects past learning of the CS-US association. The perseveration continues until the LTM weights that support it have been eroded by dopamine dip-gated weight decay (Fig. 19f). Once a new behavioral response is made, it can be reinforced by dopamine bursts (Fig. 19).

In addition to replicating SVD learning curves (Jagadeesh et al., 2001), the model can replicate data reported by Malkova et al., (1997) where FSS leads to an immediate shift in cue preference (Fig. 21). This change in cue preference occurs as a result of devaluation of the US and the automatic revaluation of the associated CS.

*3.1.5 Neurophysiological Results.* Figs. 3-6 illustrate some of the cell temporal profiles that the model simulated. Because published recordings made during the SVD task are sparse, the cell profiles generated by the model in the SVD task are predictions (Figs. 17, 21), based on parameters set for the US and CS task (Figs. 12, 14). The little data available on RHIN electrophysiology, especially for the region identified by Parker and Gaffan (1998), reveal that RHIN cells respond much like ITA cells, but are more strongly modulated by reward level (Xiang and Brown, 1999; Mogami and Tanaka, 2006).

A recent study by Padoa-Schioppa and Assad (2006) reported orbitofrontal neuronal responses in a task similar to SVD. Despite some differences in the stimuli used, their observations support the object-value interpretation of ORB and MORB cell activity profiles. They observed responses associated with a CS, and reported that taste-value cells show an



increasing response to rewards of increasing magnitudes. The increase obeyed a principle of diminishing returns similar to that demonstrated by MORB and ORB cells in Figs. 13 and 11.

### **3.2 Theoretical and Experimental Comparisons**

**3.2.1 Opponent Properties of the LH.** Ono et al. (1986a) identified two functional classes of neurons in the lateral hypothalamus: ‘opposite’ neurons and ‘specific’ neurons. Like model LH\_out cells, opposite cells appear to be affectively valenced, conditionable, respond similarly to both the CS and US, reveal inhibition and excitation patterns characteristic of opponent processing when presented with stimuli of the opposite valence, and can show antagonistic rebounds during extinction trials (Nakamura et al., 1987). In contrast, Ono and colleagues do not provide a systematic analysis of ‘specific cells’, which we predict to correspond to model LH\_gus and LH\_in cells (Figs. 12, 14). In the few data reported, specific cells were shown only to respond to specific metabolites such as sugar or protein, and not to stimuli of the opposite valence such as electric shock. This corresponds to drive-sensing LH\_in cells which respond specifically to either appetitive or aversive drive information. The LH\_gus cells similarly respond only to either appetitive or aversive information (opponent interactions occur in the subsequent stage), from GUS and LH\_in cells. Additional research into the response properties of LH specific cells during learning and extinction would be desirable to clarify their role in reversal.

**3.2.2 Approach and Avoidance Regulation.** In addition to opponent coding in the LH, cells in the AMYG, ORB and LH are also coded into appetitive and aversive channels (Thorpe et al., 1983; Ono et al., 1986a; Muramoto et al., 1993; Paton, et al. 2006). These ON and OFF channels appear to link hunger and satiety circuits directly into approach and avoidance behavior circuits. This links the opponency of homeostatic regulatory and affective systems with the opponency of approach and avoidance behavior as a control mechanism for restoring homeostatic balance.

**3.2.3 Ventral Striatum and Reward Valuation.** Appetitive and aversive information is processed throughout an evaluative neuraxis including the LH, AMYG, ORB, and VS. Considering the constraints reprised above, it now seems unlikely that the VS-VP is the source of affective valuation, as Berridge and others have proposed. Although Berridge demonstrated that hedonic behavioral responses can be blocked or elicited by manipulations of the ventral pallidum (Reynolds and Berridge, 2002), the weight of the evidence makes it more likely that those protocols were altering hedonically expressive behaviors rather than affective states. Lesions of the VS or VP result in the disinhibition of, or loss of the ability to disinhibit, pontine motor generators that normally trigger the fixed action patterns that make up hedonic behaviors such as appetitive lip-smacking, aversive head shaking, and grooming behaviors (Zahm, 1999; Pecina and Berridge, 1996). The model predicts that blood pressure changes indicative of affective state will prove to be dissociable from hedonic behaviors that can be affected by VS/VP lesions.

**3.3 Alternative Models of Satiety and Revaluation.** Dayan (2001) addressed Pavlovian conditioning and the automatic revaluation of stimuli using a temporal difference (TD) model. His model was not used to replicate results from a specific set of tasks. Nor is it a real-time model of neural processing. It did, however, address how Pavlovian and habit learning mechanisms differentially evaluate stimuli and responses in tasks similar to the US task, US-FSS task, the CS task and CS-FSS task.

Dayan proposed that the AMYG, ORB, SNc/VTA, and ventral striatum compose a TD circuit that learns to compute “reward prediction errors”. These reward prediction error signals are represented in phasic components of the DA signal (Schultz, 2007) and calculated in the



SNc/VTA by subtracting the expected value of a reward (signaled in the AMYG/ORB) from the actual reward value. In order to address FSS phenomena, Dayan augmented the core circuit with: (1) a stimulus representation area where CS and US can form stimulus-stimulus associations; (2) a “hard-wired evaluator” (HE) that modulates reward inputs to the SNc/VTA by hunger and satiety levels; (3) separate areas to plan Pavlovian and instrumental responses; and (4) a separate ventral striatum, which translates dopaminergic error signals into response-altering motivational signals. During consumption, a US activates: (1) the HE, allowing US signals to be translated into reward signals that are weighted by hunger and satiety, (2) the Pavlovian response planning area, allowing the US to elicit a UR, and (3) the AMYG/ORB, to help establish reward expectations. During CS presentation, a CS uses learned pathways: (1) to elicit a CR via the Pavlovian response planning area, (2) activate a hunger-weighted reward signal via a learned pathway from the sensory representation to the HE; and (3) activate the AMYG/ORB and calculate the value of an expected outcome.

The current model differs in several important ways from Dayan’s model. First, the brain areas responsible for Dayan’s hard-wired evaluator were left unspecified. Second, the model is based on the TD algorithm, which incorrectly predicts that during learning of a CS-US association, DA bursts will slowly migrate through intermediate times as the DA response is transferred from the time of primary reward to the time of a predictive CS’s onset. Instead, the data show that dopamine bursts occur only at the time of CS onset or primary reward – not at intermediate times. This predictive failure of TD models was overcome in the model of Brown et al. (1999) for the learning of reward prediction error signals on which MOTIVATOR builds. Brown et al. (1999) showed how to realize predictive error calculations without using TD operations. Instead, it used spectral timing circuitry that had previously been used to model adaptively-timed dynamics in the hippocampus (Grossberg and Merrill, 1992, 1996; Grossberg and Schmajuk, 1999) and the cerebellum (Fiala et al., 1996; Grossberg and Merrill, 1996).

MOTIVATOR incorporates the Brown et al. (1999) model of basal ganglia interactions (see lower right circuit in Fig. 1). In it and in the data, the early CS-induced burst grows as learning progresses, whereas the later US reward-induced burst shrinks during learning (Schultz, 2007). Because proponents of the TD model had not noticed the incorrectness of its prediction of the appearance of bursts at a full range of intermediate times between CS and US during learning, they (Niv et al., 2005) recently used the same TD property to argue that the dopaminergic “uncertainty response” reported by Fiorillo et al. (2003) must be an averaging artifact. Strong empirical evidence against this TD-based interpretation was recently reported in Fiorillo et al. (2005). Tan and Bullock (2008a, 2008b) have recently shown how an extension of the Brown et al. (1999) model can explain the DA “uncertainty response” reported in Fiorillo et al. (2003, 2005).

More generally, MOTIVATOR can describe neural activity in real time, is much more specific in its anatomical and neurophysiological hypotheses than the model sketched in Dayan (2001), and has a significantly broader explanatory range. To exemplify the latter, in the augmented TD model of Dayan (2001), CS revaluation only modulates response vigor, not the salience of perceptual cues, as seems necessary to explain electrophysiological results akin to those reported by Jagadeesh et al. (2001). In addition, the augmented TD model omits any treatment of AMYG-LH interactions, and it collapses the AMYG-ORB distinction. Thus it cannot explain many key observations.

*3.3.1 Lateral hypothalamic Models of Valuation and Satiety.* Conover and Shizgal (Conover and Shizgal, 1994; Conover et al., 1994) experimentally and theoretically examined

the issues of valuation and satiety using sodium and sugar appetite as model drives. They proposed that the LH plays a critical role in the calculation of reward value following satiety. They summarized their observations into two possible models: a series and a convergence model (Conover and Shizgal, 1994). These architectures presume that the LH is divided into separate appetitive and aversive channels that process hunger and satiety for sodium and sugar. Their experimental results led them to propose that these distinct drive inputs are summed in separate appetitive and aversive channels and the net activation is then registered at a central site. Their ultimate proposal is similar to Fig. 2a and supports the CogEM prediction that the LH includes a gated dipole opponent drive circuit (Grossberg, 1972b; Section 2.1.3).

*3.3.2 Instrumental Models of Satiety.* Busemeyer et al. (2002) proposed a dynamical performance model that is capable of simultaneously comparing and deciding between different items in a manner compatible with the SVD task. The dynamics of the core model were designed to replicate observations about the deliberation time associated with different binary choices. The model differs from the MOTIVATOR model in that it does not include synaptic weight learning, and it makes decisions solely on the basis of the utilities of response alternatives, which limits its applicability to Pavlovian phenomena. In order to perform the decision-making task, the model is presented with a set of cues, a set of response options associated with those cues, the current needs of the model, and a set of values corresponding to the utility of those options in satisfying those needs.

The utility of response options is presumed to be encoded in the ORB and LH. The need state of the model is presumed to be encoded in the LH. When a set of cues is presented for a decision, the corresponding option sets are primed and the utility of these response options is compared with the current needs of the model. The response options with the largest expected gains are forwarded on to the amygdala which then compares this information with attentional inputs. The ORB then takes this value information from the amygdala and decides which cue is to be the goal of action. Behavioral output is governed by the VS. The model has a number of features similar to MOTIVATOR in terms of its basic outline of the role of different brain regions. However, little argument is presented to justify the formal model's biological correspondences, and the model's nodes' dynamics are not systematically compared with neuronal homologues.

Frank and Claus (2006) simulated the SVD-FSS task as an application of their decision-making and instrumental learning model. In their model, motivational information arises from the AMYG and projects to the ORB. On the basis of fMRI studies, they suggest the lateral ORB favors aversive information and the medial ORB favors appetitive information. In order to generate a SVD-FSS effect with their model, they had to introduce an additional training session following the food-specific satiation of the US. During this additional training session, the model is presented with both the CS and the devalued reward and learns to associate a new value with the CS. The need for such an additional training session is contrary to the purpose and protocol of devaluation experiments (Malkova et al., 1997) and the need to retrain shows that their model makes predictions that are incompatible with the phenomenon of automatic devaluation. In addition, their model differs from the MOTIVATOR model in that it does not simulate a real-time neurophysiological output.

There are no other computational models that simulate the FSS task using data from neurophysiological cell types. There are models that use neurophysiological cell types to treat tasks such as spatial alternation or reversal. There are also numerous models of US, CS, and SVD learning. A subset of recent models is reviewed below.

*3.3.3 Neurophysiological Orbitofrontal Models.* Deco and Rolls (2005) proposed a performance model of the ORB in a serial reward reversal task. In this task an animal is presented with a single CS and must decide whether it is associated with reward or punishment. After several trials, the contingency reverses. Deco and Rolls hypothesize that once all contingencies have been learned, the reversal of this task relies on five functional classes of orbitofrontal cells: sensory cells, reward cells, motivationally valenced sensory-intermediate cells, error cells, and rule cells. Sensory cells receive object identity information. Reward cells signal both the expectation and presentation of rewards or punishments. Motivationally valenced sensory-intermediate cells are used to represent all possible stimulus-outcome associations—different cells respond to different objects depending on which outcome is expected. These cells determine whether an object will elicit approach or avoidance. Error cells are activated whenever an error in the serial reversal task is made. Rule cells compete with one another to indicate whether the current trial is expected to be a forward or reversal trial. Each rule cell primes the set of sensory-motivational cells that are consistent with whether a given CS is expected to be rewarded or punished. The authors propose that when an error is made, an error signal is calculated from an extra-orbital source and, in the subsequent trial, the alternate rule cell becomes dominant. If no error is made, the dominant rule cells from the previous trial remain active.

A problem with this proposal is that it over-predicts. The model always reverses in one trial whereas animals studied by Rolls and others in earlier studies (e.g., Rolls and Baylis, 1994) typically take as many as ten trials to reverse their behavior, with more trials required for olfactory stimuli (Rolls et al., 1996). In addition, the model's attribution of reversal learning entirely to ORB is at odds with data showing that lesions to the basal ganglia impair reversal, often severely (Ferry et al., 1999).

Deco and Rolls' proposal also suffers because it does not explain how task-irrelevant learning is prevented and task-relevant learning is permitted. The MOTIVATOR model incorporates attentional and learning mechanisms that answer these questions: attentional feedback suppresses learning to distractors (Fig. 8), adaptable weights allows model responses to reflect the predictive strength of stimulus-reward associations (see Equations (6, 7)), and basal ganglia structures regulate learning by generating predictive error and reinforcement signals (Fig. 8, see Equation (20 - 33)).

In the MOTIVATOR model, reversals are not attributable to “reversal rule” cells. As detailed in Fig. 20, the MOTIVATOR model relies on rapid learning of stimulus outcome associations. As a result reversals require a period of 10-15 trials to unlearn an association and reverse contingency. Both the shape of the learning and reversal curves reported by Jagadeesh et al. (2001) and the number of trials required for behavioral reversal suggest that a learning based mechanism plays a dominant role in reversal learning, at least for the first few reversals. The slow learning of new contingencies many figures from Critchley and Rolls (1996).

**3.4 Predictions and Discussion.** The model makes a number of novel electrophysiological, physiological, and behavioral predictions. The model reproduces discharge dynamics of known cell types in the CS, US, and SVD tasks, including signals that predict saccadic reaction times and CS-dependent changes in systolic blood pressure (Figs. 12, 14, 17). The model uses these circuits to predict the existence of neural cell types and the response profiles of these neurons in a number of brain areas for two important tasks that have not been experimentally treated in the primate: the SVD-FSS task and the CS-FSS task. The model also predicts that antagonistic rebounds observed in LH opposite cells are caused by a dopamine dip

that disinhibits an arousal burst to the LH (Figs. 16, 18). The model predicts that when antagonistic rebounds are observed in the LH during the CS and SVD extinction trials, they will also be observed in the activity profiles of AMYG cells (Fig. 16). The model predicts that dopamine cells from the SNc/VTA are important for learning in the cortex and basal ganglia, but not for AMYG-LH learning (Fig. 8). The model predicts that interactions between the AMYG, LH and cardiovascular medulla bring about conditioned changes in systolic blood pressure and that the magnitude of the BPR will reflect the motivational support of the cue, i.e. hunger level (Fig. 11). The model predicts that US value, reflected in the AMYG is calculated by the nonlinear sum of taste-drive cells in the LH\_out (Fig. 7). Another electrophysiological result predicted is that during extinction trials the DA dip engendered by omission of reward will release an arousal burst response in both types of specific cells of the LH, namely appetitive and aversive cells (Fig. 16). The model also predicts that the specific enhancement seen in the ITA during the SVD and SVD-FSS task is due to lateral ORB projections and could be disrupted by suitable ORB lesions (Fig. 8).

#### 4. MATHEMATICAL EQUATIONS AND PARAMETERS

**4.1 Modeling Framework.** The model is defined by three types of differential equations that represent dynamics operating on different time scales: short term memory (STM) equations represent the activities of simulated neurons, medium term memory (MTM) equations represent habituated transmitter processes, and long term memory (LTM) equations represent long term changes in learned adaptive weights. For each trial, task-specific input signals were provided and equations were numerically integrated using MATLAB and a fourth-order Runge-Kutta algorithm. The time step was 0.00033.

*4.1.1 Short Term Memory Equations.* The activity, or short term memory (STM), dynamics of simulated neurons was modeled using a single voltage compartment whose membrane potential,  $V(t)$ , was determined by an equation of the form:

$$C_m \frac{dV(t)}{dt} = -(V(t) - E_{leak})g_{leak} - (V(t) - E_{excite})g_{excite} - (V(t) - E_{inhibit})g_{inhibit} . \quad (1)$$

In (1),  $C_m$  is the membrane capacitance, conductances  $g_{excite}$  and  $g_{inhibit}$  represent the excitatory and inhibitory inputs to the cell, and terms  $E_{excite}$  and  $E_{inhibit}$  represent the reversal potentials for excitatory ions and inhibitory ions, respectively (Grossberg, 1973; Hodgkin and Huxley, 1952; van den Pol and Trombley, 1993; Bertrand and Changeux, 1995). Term  $g_{leak}$  is a constant leakage conductance and  $E_{leak}$  is the reversal potential of leaked ions.

In the equations that follow, the explicit reference to time, (t), is dropped and the capacitance is represented as the reciprocal of the time constant:  $1/\tau$ . Also, letters other than V are used for the membrane voltages; e.g. I for the voltage of a model ITA cell. The equations were scaled so that the resting, or leakage reversal, potential is 0 and the excitatory reversal potential is 1. The inhibitory reversal potential varied in a neuron-specific fashion (Owens and Kriegstein, 2002) that calibrates how sensitive to inhibition the baseline firing rate is. Three forms of rescaled STM equations were used in simulations: shunting STM equations, additive STM equations, and attentive shunting STM equations (Gove et al., 1995; Grossberg, 1973, 1988).

Shunting equations are of the form:

$$\frac{1}{\tau} \frac{d}{dt} V_i = -V_i + (1 - V_i) E_i - (C + V_i) I_i. \quad (2)$$

Term  $V_i$  is the activity of the  $i$ -th model neuron or neuron population. In (2),  $-V_i$  represents the passive decay of neural activity toward the leakage reversal potential of zero,  $E_i$  denotes the total excitatory input that can drive  $V_i$  to its excitatory saturation point of 1, and  $I_i$  denotes the total inhibitory input that can drive  $V_i$  to its inhibitory saturation point  $-C$ , where  $C$  is chosen as 0, 0.2, or 1. For all STM equations except ITA and SD cells,  $\tau = 50$ . For ITA cells,  $\tau = 10$ , and for SD cells it is variable, set according to the parameter  $\beta_g$ , defined in (37).

Additive equations are leaky integrator equations that approximate (2) when  $E_i$  and  $I_i$  do not drive  $V_i$  too close to 1 or  $-C$ :

$$\frac{1}{\tau} \frac{d}{dt} V_i = -V_i + E_i - I_i. \quad (3)$$

Attentive shunting networks link top-down and bottom-up inputs in an excitatory on-center projection and an inhibitory off-surround projection. Excitatory on-center projections are indexed by  $i$  and inhibitory off-surround inputs are indexed by  $k \neq i$ :

$$\frac{1}{\tau} \dot{V}_i = -V_i + (1 - V_i) \left( D_i (1 + M_i) \right) - (C + V_i) \left( \sum_{k \neq i} D_k + \sum_{k \neq i} M_k \right). \quad (4)$$

The bottom-up inputs  $D_i$  in (4) provide driving on-center excitatory inputs  $D_i$  (Chance et al., 2002) and driving off-surround inhibitory inputs  $D_{k \neq i}$ . Top-down signals provide modulatory, excitatory on-center inputs  $M_i$  that multiplicatively control the gain of driving inputs (Gove et al., 1995; Grossberg, 1999, 2003; Salinas and Thier, 2000; Chance et al., 2002), and driving inhibitory off-surround inputs  $M_k$ ,  $k \neq i$ . This type of anatomy has been predicted to embody attentional processes, to realize contrast gain control, and to enable cortical learning to proceed in a stable manner (Carpenter and Grossberg, 1987, 1991; Grossberg, 1980, 1999, 2003).

**4.1.2 Medium Term Memory Equations.** Medium term memory (MTM) equations describe how recent use of synapses affects the gain of signals transmitted through these synapses from one brain region to another. To capture these habituation (inactivation, depressing) effects, MTM equations describe fluctuations in the accumulation and depletion of a chemical transmitter  $G$ :

$$\dot{G} = (1 - G) - G[V]^+ \quad (5)$$

In (5),  $(1 - G)$  is the accumulation rate and  $G[V]^+$  is the activity-dependent habituation rate, where  $[V]^+$  is an output signal from the activity  $V$  of the presynaptic cell (Grossberg, 1972b, 1980). A similar equation can be used to describe the activation and inactivation of postsynaptic sites. Variations of this equation were used to model the habituation of appetitive and aversive signals.

**4.1.3 Long Term Memory Equations.** Long term memory (LTM) is formed by lasting changes in the conductive efficacy of synapses (Brown et al., 1990). The model employs two basic LTM equations: (1) activity-gated steepest descent learning; and (2) dopamine-gated steepest descent learning. Gated steepest descent learning describes how the synaptic strengths change as a function of presynaptic and postsynaptic activity levels. The outstar variant of the

gated steepest descent learning involves a rule wherein presynaptic activity gates learning and LTM weights approach postsynaptic cell activations (Grossberg, 1968). These changes in synaptic efficacy are associated with the mechanisms of long term potentiation (LTP) and long term depression (LTD). In (6), term  $W_{jk}$  is the LTM trace in the pathway from cell j to cell k, term  $S_j$  is a non-negative presynaptic gating signal from cell j, and term  $T_k$  is a non-negative postsynaptic cell signal from cell k to which weight  $W_{jk}$  is attracted by steepest descent when  $S_j > 0$ :

$$\varepsilon \dot{W}_{jk} = S_j (T_k - W_{jk}) \quad (6)$$

The weights  $W_{jk}$  change at a rate scaled by  $\varepsilon$ .

The instar equation describes a similar process of gated steepest descent except that learning is gated by a postsynaptic signal  $T_k$  and LTM weights are attracted to a presynaptic signal,  $S_j$  (Grossberg, 1976):

$$\varepsilon \dot{W}_{jk} = T_k (S_j - W_{jk}) \quad (7)$$

Figs. 1 and 8 indicate that many LTM weights are affected by dopaminergic inputs,  $D_i$ . Learning by these weights is governed by a dopamine-modulated steepest descent rule. For the outstar case of dopamine-gated steepest descent, the equation is:

$$\varepsilon \dot{W}_{jk} = S_k \left[ (A + D_1)(T_j - W_{jk}) - (B + D_2)W_{jk} \right] \quad (8)$$

The first term,  $S_k$  gates learning and decay by presynaptic activity. The term  $(A + D_1)(T_j - W_{jk})$  shows that steepest descent learning is modulated by a dopamine burst  $D_1$  and occurs at a rate A in the absence of dopamine. Term  $-(B + D_2)W_{jk}$  describes weight decay that is modulated by a dopamine-dip ( $D_2$ ) and which occurs at a rate B in the absence of dopamine. Dopamine bursts and dips signal unexpected rewards or unexpected non-rewards (Brown et al., 1999, 2004; Schultz, 2007).

The dopamine-modulated instar equation describes a process of gated steepest descent that is similar to the dopamine-modulated outstar equation:

$$\varepsilon \dot{W}_{jk} = T_k \left[ (A + D_1)(S_j - W_{jk}) - (B + D_2)W_{jk} \right] \quad (9)$$

In (9), learning is gated by the postsynaptic signal,  $T_k$  and LTM weights are attracted to presynaptic signal,  $S_j$ .

## 4.2 Inputs.

**4.2.1 Visual Inputs.** Visual (VIS) inputs,  $C_j$ , provide binary inputs to ITA cells that represent the outcome of visual preprocessing (see Equation (10)). Three visual stimuli were used: CS1, CS2, and CS3. The default association of CS1 is with US1 and of CS2 with US2. CS3 was used with either US.

**4.2.2 Gustatory Inputs.** Gustatory (GUS) inputs  $G_m$  provide a binary taste input to RHIN and LH\_gus cells (see Equations (16, 18)). The GUS signals code four basic tastes: fatty, salty, sweet, and umami (Scott et al., 1995; Torii et al., 1998). These tastes were presented as food reward vectors. There were two food rewards US1: [1, 0, 0, 1] and US2 [1, 1, 0, 0], representing



a sugar and protein reward and a sugar and salt reward (e.g. milk or sports drink), respectively. GUS inputs habituate during food consumption and completely habituate after either 25 or 150 trials, depending on the type of trial being simulated (for specification of the habituation of GUS inputs, see Equation (35)).

**4.2.3 Drive Inputs.** Drive or metabolic inputs  $M_{si}$  to LH\_in cells determine hunger and satiety level (see Equation (16)). Appetitive metabolic signals correspond to sugar, protein, salt, and fat (Ono et al., 1986a; Bergen et al., 1996; Davidson et al., 1997; Torii et al., 1998), which were selected because behavioral studies show that depriving one these nutrients leads animals to search them out at the expense of others (Cabanac and Lafrance, 1992; Davidson et al., 1997; Sclafani, 2004). These homeostatic drives correlate with taste inputs so that the taste and drive profile of a food reward have similar components. After each trial in which a food reward was consumed, drive inputs decayed in a food-specific and linear manner, so that during devaluation trials, satiety occurred after 25 “ingestions” of a given food. For all other trial types, 150 “ingestions” of a given food were required for satiety to occur.

**4.2.4 Arousal Inputs.** Arousal inputs,  $\alpha_1$ , to LH\_in cells were normally set to 0.5 (see Equation (16)). Arousal inputs increase sixteen fold when there is an arousal burst. An arousal burst is triggered whenever the cortical dopamine dip ( $D_2$ ) signal falls below a threshold value of 0.2. While MOTIVATOR does not incorporate a detailed model of the pathway by which cortical dopamine dips elicit arousal bursts in the lateral hypothalamus, evidence suggests the amygdala plays a role in this process. The amygdala is involved in extinction, frustration and nonreward (Papini, 2003) and transient inactivation of the central nucleus nonspecifically excites lateral hypothalamic cells (Nakamura et al., 1987). Hence the burst of hypothalamic activity associated with surprising nonreward may be a result of amygdala inactivation by the dopamine dip or the disinhibition of other indirect sources of arousal to the LH (e.g. Zahm, 2000).

### 4.3 Short Term Memory Equations.

**4.3.1 ITA STM Equations.** ITA cell activities,  $I_j$ , model visually-sensitive, cue-selective cells in the anterior inferotemporal cortex. The subscript  $j$  indexes the different visual object categories coded by cells in IT4. ITA cells are modeled by an attentive shunting network (4) that receives bottom-up driving VIS inputs  $C_j$  and top-down attentive ORB inputs,  $O_{ji}^L$ :

$$\frac{1}{10} \dot{I}_j = -I_j + (1 - I_j) \left( 6C_j \left( 1 + 2 \sum_i O_{ji}^L \right) \right) - (0.2 + I_j) \left( 9 \sum_{k \neq j} C_k + 17 \sum_{k \neq j, i} O_{ki}^L \right). \quad (10)$$

By (10),  $I_j$  passively decays at rate 1 via term  $-I_j$ . The excitatory and inhibitory shunting terms  $(1 - I_j)$  and  $-(0.2 + I_j)$ , respectively, bound cell activity within the interval  $[-0.2, 1]$ . The driving excitatory input is  $6C_j$ . The top-down modulatory excitatory ORB inputs are  $2 \sum_i O_{ji}^L$ .

The summation over index  $i$  enables appetitive and aversive ORB inputs to enhance salience equally. The inhibitory driving off-surround terms are  $9 \sum_{k \neq j} C_k$  and  $17 \sum_{k \neq j} \sum_i O_{ki}^L$ .

**4.3.2 ORB STM Equations.** ORB cell activities,  $O_{ji}^L$ , model object-value orbitofrontal cells that respond selectively to motivationally supported visuosensory cells (Hosokawa et al., 2004). Subscript  $i$  indicates whether the cell is appetitive or aversive and subscript  $j$  indicates the ITA visual category that the ORB cell prefers. ORB cells are modeled by an attentive shunting

network that receives bottom-up driving ITA inputs,  $I_j$ , and top-down attentive modulatory AMYG inputs,  $A_{ri}$  (Amaral and Price, 1984):

$$\begin{aligned} \frac{1}{50} \dot{O}_{ji}^L = & -O_{ji}^L + (1 - O_{ji}^L) \left( 0.1 [I_j]^+ \left( 1 + 160 \sum_r W_{jri}^{AO} [A_{ri}]^+ \right) \right) \\ & - (O_{ji}^L) \left( 2 \sum_{k \neq j} \sum_r W_{kri}^{AO} [A_{ri}]^+ + 0.2 \sum_{k \neq j} [I_k]^+ + 50 \sum_{l \neq i} [O_{jl}^L]^+ \right). \end{aligned} \quad (11)$$

By (11),  $O_{ji}^L$  passively decays at rate 1 via term  $-O_{ji}^L$ . The excitatory and inhibitory shunting terms  $(1 - O_{ji}^L)$  and  $-(O_{ji}^L)$ , respectively, bound cell activity within the interval  $[0, 1]$ . The bottom-up excitatory driving ITA input is  $0.1 [I_j]^+$ . The output signal function  $[I_j]^+$  rectifies  $I_j$  at a zero threshold, passing only non-negative values of  $I_j$ . The top-down modulatory excitatory attentive input is  $160 \sum_r W_{jri}^{AO} [A_{ri}]^+$ , where the  $W_{jri}^{AO}$  are LTM weights that learn to map US-specific AMYG cells,  $(r, i)$ , to CS-specific ORB cells,  $j$ , where  $r$  denotes the  $r$ -th AMYG category and  $i$  whether it is appetitive ( $i = 1$ ) or aversive ( $i = 2$ ). The driving inhibitory off-surround inputs from ITA and AMYG are  $0.2 \sum_{k \neq j} [I_k]^+$  and  $\sum_{k \neq j} \sum_r W_{kri}^{AO} [A_{ri}]^+$ , respectively.

In addition, there is opponent inhibition between appetitive and aversive ORB cells,  $50 \sum_{l \neq i} [O_{jl}^L]^+$ .

**4.3.3 AMYG STM Equations.** AMYG cell activities,  $A_{ri}$ , model multimodal, US-selective cells found in the basolateral amygdala (Nishijo et al., 1988a, 1988b). The subscript  $r$  identifies the  $r$ -th US-specific affective category and the subscript  $i$  indicates whether the cell is appetitive or aversive. AMYG cells are modeled by attentive shunting networks that receive bottom-up excitatory driving LH\_out cell inputs,  $H_{si}^O$ , and top-down attentive modulatory excitatory inputs from ITA cells,  $I_j$ , and RHIN cells,  $R_t$ :

$$\begin{aligned} \frac{1}{50} \dot{A}_{ri} = & -A_{ri} + (1 - A_{ri}) \left( \sum_s W_{sri}^{HA} [H_{si}^O]^+ \left( 0.3 + 13 \sum_j W_{jri}^{IA} [I_j]^+ + 13 \sum_t W_{tri}^{RA} [R_t]^+ \right) + 150 f(A_{ri}) \right) \\ & - (0.2 + A_{ri}) \left( 0.25 \sum_{u \neq r} \sum_s W_{usi}^{HA} [H_{si}^O]^+ + 250 \sum_{u \neq r} g(A_{ui}) \right). \end{aligned} \quad (12)$$

By (12),  $A_{ri}$  passively decays at rate 1 via term  $-A_{ri}$ . The excitatory and inhibitory shunting terms  $(1 - A_{ri})$  and  $-(0.2 + A_{ri})$ , respectively, bound cell activity within the interval  $[-0.2, 1]$ .

The driving excitatory LH\_out input is  $\sum_s W_{sri}^{HA} [H_{si}^O]^+$ . The excitatory modulatory inputs are from the ITA,  $13 \sum_j W_{jri}^{IA} [I_j]^+$ , and the RHIN,  $13 \sum_t W_{tri}^{RA} [R_t]^+$ . The  $W_{rsi}^{HA}$  are LTM weights that associate LH drive features with US-specific AMYG categories. Weights  $W_{jri}^{IA}$  and  $W_{tri}^{RA}$  are

LTM weights that map CS-specific ITA cells and US-specific RHIN cells to US-specific AMYG cells, respectively. The driving inhibitory off-surround LH\_out input is  $0.25 \sum_{u \neq r} \sum_s W_{usi}^{HA} [H_{si}^O]^+$ .

Equation (12) also includes recurrent excitation  $150f([A_{ri}]^+)$  and inhibition  $250 \sum_{u \neq r} g([A_{ui}]^+)$ .

The recurrent excitatory signal function,  $f$ , ensures strong feedback amplification for small values of  $A_{ri}$  between  $[0, 0.025]$ , but progressively less amplification above 0.025:

$$f(A_{ri}) = \begin{cases} 0 & \text{if } A_{ri} \leq 0 \\ (A_{ri})^2 & \text{if } 0 < A_{ri} \leq 0.025 \\ (A_{ri})^2 e^{-1000(A_{ri}-0.025)} & \text{if } A_{ri} > 0.025. \end{cases} \quad (13)$$

The recurrent inhibitory signal function,  $g$ , is:

$$g(A_{ui}) = \begin{cases} 0 & \text{if } A_{ui} \leq 0 \\ (A_{ui})^2 & \text{if } A_{ui} > 0. \end{cases} \quad (14)$$

For values of  $A_{ri}$  in the interval  $[0, 0.025]$ , the recurrent signal functions (13) and (14) support fast contrast enhancement among competing AMYG cells. As a result, a single winner is chosen. Outside this interval, recurrent inhibition  $g([A_{ui}]^+)$  continues to grow quadratically to maintain suppression of the losers, whereas recurrent excitation is negligible. This allows choice without a loss of analog sensitivity so that the activity level of the AMYG winner closely tracks the amplitude of its non-recurrent inputs.

**4.3.4 LH\_in STM Equations.** LH cells, taken together, form an array of recurrent gated dipole opponent processing circuits wherein metabolic, sensory reinforcer, and amygdala category signals converge. LH input cell activities,  $H_{si}^I$ , model LH cells that reflect the identity and concentrations of specific nutrients (Karadi et al., 1992); see Fig. 8. The subscript  $s$  indexes the metabolic features processed by the LH\_in cell. The subscript  $i = 1, 2$  indicates whether the cell is appetitive or aversive. LH\_in cells receive metabolic inputs  $M_{si}$ , arousal inputs  $\alpha_1$ , and LH\_out cell feedback signals,  $H_{si}^O$  (see Equation 17):

$$\frac{1}{50} \dot{H}_{si}^I = -H_{si}^I + 0.05M_{si} + 0.25\alpha_1 + 0.35[H_{si}^O]^+ \quad (15)$$

Activity  $H_{si}^I$  passively decays at rate 1 via term  $-H_{si}^I$ . Additive equations are used and remain bounded because the inputs are bounded.

**4.3.5 LH\_gus STM Equations.** LH\_gus cell activities,  $H_{si}^G$ , model taste-responsive LH drive-specific cells (Karadi et al., 1992). Each cell is indexed by homeostatic features,  $s$ , and affective valence,  $i$ . LH\_gus cells receive driving inputs from LH\_in cells,  $H_{si}^I$ , and modulatory GUS inputs,  $G_m$ :

$$\frac{1}{50} \dot{H}_{si}^G = -H_{si}^G + 5Y_{si}([H_{si}^I]^+)(1 + 0.6W_{msi}^{GH} X_m G_m). \quad (16)$$

Activity  $H_{si}^G$  decays at rate 1 via term  $-H_{si}^G$ . The excitatory driving inputs are  $4Y_{si}[H_{si}^I]^+$ .  $Y_{si}$  is a habituating transmitter that gates  $H_{si}^I$  input signals; see Equation (34). In the excitatory

modulatory gustatory inputs  $0.6W_{msi}^{GH} X_m G_m$ , the taste-specific habituated term  $X_m$  gates gustatory inputs,  $G_m$  (see Equation 35).  $W_{msi}^{GH}$  are fixed weights that map tastes to drive features in a one-to-one fashion.

**4.3.6 LH\_out STM Equations.** LH output cell activities,  $H_{si}^O$ , model LH cells that generate opponent responses to appetitive and aversive cues and rewards (Ono et al., 1986a). Each cell activity is indexed by homeostatic features,  $s$  and affective valence,  $i$ :

$$\begin{aligned} \frac{1}{50} \dot{H}_{si}^O = & -H_{si}^O + (1 - H_{si}^O) \left( 1.75 [H_{si}^G]^+ \left( 1 + 1.75 \sum_r W_{sri}^{AH} [A_{ri}]^+ \right) \right) \\ & - (1 + H_{si}^O) \left( [H_s^G]^+ + 0.25 \sum_{u \neq s} [H_{ui}^G]^+ + 0.5 \sum_{u \neq s} \sum_r W_{rui}^{AH} [A_{ri}]^+ \right). \end{aligned} \quad (17)$$

The excitatory and inhibitory shunting terms  $(1 - H_{si}^O)$  and  $-(1 + H_{si}^O)$  in (17) bound cell activity within the interval  $[-1, 1]$ . LH\_gus cells provide bottom-up excitatory driving inputs  $1.75 [H_{si}^G]^+$ . The top-down attentive excitatory modulatory input from AMYG categories is  $1.75 \sum_r W_{sri}^{AH} [A_{ri}]^+$ . The  $W_{sri}^{AH}$  are LTM weights that map US-specific AMYG category cell outputs to US-related metabolic features in LH\_out cells. Summation over  $r$  occurs across US-specific AMYG category cells. Inhibitory driving off-surround terms are from AMYG categories  $0.5 \sum_{u \neq s} \sum_r W_{rui}^{AH} [A_{ri}]^+$  and from LH\_gus cells  $0.25 \sum_{u \neq s} [H_{ui}^G]^+$ . These inhibitory inputs suppress LH\_out cells at which there is a poor match between top-down attentive signals from AMYG and bottom-up driving input signals from LH\_gus cells. LH\_gus cells provide an additional driving inhibitory input  $H_s^G$  that introduces opponent inhibition between an opponent pair,  $H_{si}^O$  and  $H_s^O$ , of drive channels. In all, the two types of inhibition link the various LH opponent circuits into a gated dipole field (Grossberg 1972b, 1982, 1984; Olson and Grossberg, 1998).

**4.3.7 RHIN STM Equations.** RHIN cells activities,  $R_t$ , model reward-selective, multimodal neurons that are proposed to exist in the rhinal cortex (Parker and Gaffan, 1998). Subscript  $t$  indexes the multimodal US category favored by the  $t$ -th RHIN cell. RHIN cells are modeled using attentive shunting networks and receive bottom-up excitatory driving inputs from ITA cell activities,  $I_j$  and GUS cell activities,  $G_m$ . Top-down modulatory excitatory inputs arise from MORB inputs,  $O_{ti}^M$ :

$$\begin{aligned} \frac{1}{10} \dot{R}_t = & -R_t + (1 - R_t) \left( \left( 0.4 \sum_j W_{jt}^{IR} [I_j]^+ + 5 \sum_k W_{mt}^{GR} G_m \right) \left( 1 + 2 \sum_i [O_{ti}^M]^+ \right) \right) \\ & - (0.2 + R_t) \left( 3 \sum_{u \neq t} \sum_j W_{ju}^{IR} [I_j]^+ + 5 \sum_{u \neq t} \sum_m W_{mu}^{GR} X_m G_m + 6 \sum_{u \neq t} \sum_i [O_{ui}^M]^+ \right). \end{aligned} \quad (18)$$

The excitatory and inhibitory shunting terms  $(1 - R_t)$  and  $-(0.2 + R_t)$ , respectively, bound cell activity within the interval  $[-0.2, 1]$ . The driving excitatory inputs are  $0.5 \sum_j W_{jt}^{IR} [I_j]^+$  and  $5 \sum_{u \neq t} \sum_m W_{mu}^{GR} X_m G_m$ . The terms  $W_{jt}^{IR}$  and  $W_{mu}^{GR}$  are weights that allow RHIN cells to respond selectively to gustatory and visual features of particular unconditioned stimuli. Learning of the  $W_{jt}^{IR}$  and  $W_{mu}^{GR}$  LTM weights is not treated in this paper, but could be done, for example, using mechanisms in Carpenter and Grossberg (1991). The modulatory MORB excitatory input is  $2 \sum_i [O_{ti}^M]^+$ . Summation over the subscript,  $i$ , indicates that both appetitive and aversive inputs from the MORB enhance salience. The driving inhibitory GUS, ITA, and MORB off-surround inputs are  $5 \sum_{u \neq t} \sum_m W_{mu}^{GR} X_m G_m$ ,  $3 \sum_{u \neq t} \sum_j W_{ju}^{IR} [I_j]^+$ , and  $6 \sum_{u \neq t} \sum_i [O_{ui}^M]^+$ , respectively.

**4.3.8 MORB STM Equations.** MORB cell activities,  $O_{ti}^M$ , model reward-selective, motivationally-sensitive neurons in the orbitofrontal cortex (Tremblay and Schultz, 2000a, 2000b). As a consequence of their anatomical connections, ORB cells represent the approach or avoidance value of visual stimuli while MORB cells represent the consumption value of rewards (Small et al., 2007); see Fig. 1. Subscript  $t$  labels the US category in the RHIN to which the MORB cell responds and the subscript  $i$  labels whether the cell is appetitive or aversive. MORB cells are modeled using attentive shunting networks that receive bottom-up driving excitatory RHIN signals,  $R_t$  and attentive top-down modulatory excitatory AMYG inputs,  $A_{ri}$  (Amaral and Price, 1984):

$$\begin{aligned} \frac{1}{50} \dot{O}_{ti}^M = & -O_{ti}^M + (1 - O_{ti}^M) \left( 0.02 [R_t]^+ \left( 1 + 200 \sum_r W_{tri}^{AM} [A_{ri}]^+ \right) \right) \\ & - (O_{ti}^M) \left( 6 \sum_{u \neq t} \sum_r W_{ru}^{AM} [A_{ri}]^+ + 0.2 \sum_{u \neq t} [R_u]^+ + 15 \sum_{j \neq i} O_{tj}^M \right). \end{aligned} \quad (19)$$

The excitatory and inhibitory shunting terms  $(1 - O_{ti}^M)$  and  $-(O_{ti}^M)$ , respectively, bound cell activity within the interval  $[0, 1]$ . The bottom-up excitatory driving RHIN input is  $0.02 [R_t]^+$ . The top-down attentive excitatory modulatory AMYG input is  $200 \sum_r W_{tri}^{AM} [A_{ri}]^+$ . The  $W_{tri}^{AM}$  are LTM weights that map US-specific AMYG categories to US-specific MORB cells which code reward value. The driving inhibitory off-surround RHIN and AMYG inputs are  $0.2 \sum_{u \neq t} [R_u]^+$  and  $4 \sum_{u \neq t} \sum_r W_{ru}^{AM} [A_{ri}]^+$ , respectively. Inhibitory inputs include a recurrent inhibition between appetitive and aversive ORB cells,  $15 \sum_{j \neq i} O_{tj}^M$ .

**4.3.9 BG STM Equations.** Brown et al. (1999) modeled four brain regions that cooperate to generate dopaminergic bursts and dips in response to unexpected rewards and nonreward: ventral striatal matrix cells (VS), ventral striatal striosomal cells (SD), pedunculopontine cells

(PPT/LDT), and dopaminergic cells (SNc/VTA); see Fig. 1. When unexpected rewards or conditioned stimuli are presented, SNc/VTA Cells show a transient burst of activity. Cells in the PPTN/LDT drive this bursting response. Conditioned stimuli excite the PPTN/LDT via VS cells while food rewards drive PPTN/LDT cells via both LH\_gus and VS inputs. CS-related inputs from the ORB activate SD cells. An adaptively-timed inhibitory pathway from SD cells to the SNc/VTA suppresses dopamine bursts at the expected time of reward. If an expected reward is omitted, this adaptively timed signal from SD cells to the SNc/VTA inhibits the dopamine cells, resulting in a transient dip in dopamine. Together, excitatory and inhibitory conditioning pathways through the basal ganglia act to ensure dopamine bursts and dips signal reward prediction errors.

*VS STM Equations.* Ventral striatal matrix cell activities are represented by activities,  $V_{ri}$ . VS cells are sensitive to US-specific value category inputs from the AMYG and learn to respond to object-value inputs from the ORB. Subscript  $r$  indicates that the  $r$ -th VS cell favors the  $r$ -th US-specific AMYG category; subscript  $i$  indicates affective valence. VS cells receive excitatory driving inputs from the AMYG,  $A_{ri}$  and ORB,  $O_{ji}^L$ :

$$\frac{1}{50} \dot{V}_{ri} = -V_{ri} + (1 - V_{ri}) \left( 20 \left( \sum_j W_{rji}^{ov} [O_{ji}^L]^+ + [A_{ri} - 0.1]^+ \right) \right) \quad (20)$$

The excitatory shunting term  $(1 - V_{ri})$  limits the activity of the cell to the interval  $[0, 1]$ . ORB inputs  $20 \sum_j W_{rji}^{ov} [O_{ji}^L]^+$  and AMYG inputs  $20[A_{ri} - 0.1]^+$  are the excitatory driving ORB and AMYG inputs, respectively. The  $W_{rji}^{ov}$  are LTM weights that map CS-specific ORB cells to US-specific VS cells, while preserving affective valence,  $i$ .

*LDT/PPTN STM Equations.* PPTN/LDT activities model the pedunculopontine and laterodorsal tegmental brainstem nuclei (Kobayashi et al., 2002). PPTN/LDT activity is described using a pair of variables,  $P_1$  and  $P_2$ , with coupled equations. PPTN/LDT cells receive opponent, bottom-up, driving inputs from LH\_gus cells,  $H_{s1}^G$  and driving inputs from VS cells (Fig. 8). The effect of VS activity  $V_{ri}$  on PPTN/LDT is modeled as a net excitatory driving input because the VS inhibits the ventral pallidum and thus disinhibits the PPTN/LDT:

$$\frac{1}{50} \dot{P}_1 = -P_1 + (1 - P_1) \left( 2.5 \left[ \sum_s (H_{s1}^G - H_{s2}^G) \right]^+ + \sum_i \sum_r V_{ri} + 40 \right) - 1500(1 + P_1)[P_2]^+ \quad (21)$$

and

$$\dot{P}_2 = -P_2 + (1 - P_2)[P_1]^+ \quad (22)$$

In (21), excitatory and inhibitory shunting terms  $(1 - P_1)$  and  $(1 + P_1)$ , respectively, bound PPTN/LDT activity within the interval  $[-1, 1]$ . The excitatory driving input  $2.5 \left[ \sum_s (H_{s1}^G - H_{s2}^G) \right]^+$  gives the net appetitive signal from LH\_gus cells to the PPTN/LDT. Excitatory driving inputs  $\sum_i \sum_r V_{ri}$  sum VS cell activity, enabling stimuli represented in the ORB to excite the PPTN/LDT.



An arousal input set to 40 biases PPTN cell activity. Term  $1500[P_2]^+$  is a driving, inhibitory off-surround input. Term  $P_2$  approximates the strength of a slow after-hyperpolarization process.

In (22), the after-hyperpolarization process,  $P_2$ , is driven by the “input”  $(1 - P_2)[P_1]^+$ , where  $(1 - P_2)$  bounds  $P_2$  activation to the interval  $[0, 1]$  and  $P_1$  represents excitatory PPTN/LDT output. The strong driving inhibition arising from the afterhyperpolarizing signal  $P_2$  is responsible for generating the phasic profile of PPTN/LDT signals  $P_1$  (see Fig. 12).

*SD STM Equations.* Striosomal cells (SD) carry out an adaptive timing function that suppresses dopaminergic bursts to expected rewards. SD cells can fire at an adaptively timed delay after the onset of a CS, thereby inhibiting dopamine cells in the SNc/VTA. This delayed burst of inhibitory firing is the result of a  $\text{Ca}^{2+}$ -dependent second-messenger process governed by metabotropic glutamate receptors (mGluR) (Fiala et al., 1996; Brown et al., 1999). Activation of mGluRs causes a spike in  $\text{Ca}^{2+}$  currents that can lead to the depolarization of SD cells after a timed delay.

SD cell activity,  $B_{jg}$ , is driven by excitatory signals,  $O_{ji}^L$ , from the ORB. Dendritic spines that receive their driving excitatory inputs from the  $j$ -th ORB cell are indexed with the subscript,  $j$ . The subscript  $g$  indicates the  $g$ -th SD cell responds to ORB inputs at a rate characterized by the  $g$ -th rate parameter,  $\beta_g$ . Note that the subscripts  $j$  and  $g$  do not necessarily index single neurons, but pools of synapses across one or many neurons (Brown et al., 1999):

$$\left(\frac{1}{\beta_g}\right)\dot{B}_{jg} = -B_{jg} + (1 - B_{jg})\left(2.5\left[O_{j1}^L - 0.03\right]^+\right), \quad (23)$$

$$\text{Where } \beta_g = \frac{10}{30g - 23}, \quad g = \{1, \dots, 60\}. \quad (24)$$

In (23), the excitatory shunting term  $(1 - B_{jg})$  bounds activity within the interval  $[0, 1]$ . ORB inputs  $2.5\left[O_{j1}^L - 0.03\right]^+$  provide the excitatory driving inputs. The rate parameters,  $\beta_g$ , span the range of values specified in Equation (24), providing the basis for a range of delayed cell activations and  $\text{Ca}^{2+}$  spikes.

Equations (25)-(27) describe how the activation of mGluR channels on SD cells gives rise to a rapid spike in cytosolic  $\text{Ca}^{2+}$ ,  $G_{jg}Y_{jg}$ , driving adaptively timed SD-to-SNc/VTA inhibitory responses. Equation (25) describes how SD cell activity,  $B_{jg}$  alters the conductance of  $\text{Ca}^{2+}$  current,  $G_{jg}$ :

$$\frac{1}{5}\dot{G}_{jg} = -4G_{jg} + (5 - G_{jg})\left(h(B_{jg})\right), \quad (25)$$

$$\text{Where } h(B_{jg}) = \begin{cases} 1 & \text{if } B_{jg} - 0.2 > 0 \\ 0 & \text{if } B_{jg} - 0.2 \leq 0 \end{cases}. \quad (26)$$

The rate at which  $\text{Ca}^{2+}$  passes to the cytosol from endoplasmic stores is bound by the term  $(5 - G_{jg})$  to the interval  $[0, 5]$ . The function  $h$ , defined in Equation (26), is a step function that indicates SD cell activity must exceed a threshold of 0.2 in order to trigger the rapid buildup of cytosolic  $\text{Ca}^{2+}$ ,  $G_{jg}Y_{jg}$ . As the  $\text{Ca}^{2+}$  concentration builds to maximal level, the available  $\text{Ca}^{2+}$  in

both endoplasmic and local cytosolic stores,  $Y_{jg}$ , rapidly depletes. Equation (27) describes how the level of available  $\text{Ca}^{2+}$ ,  $Y_{jg}$ , is decreased by a  $\text{Ca}^{2+}$  spike,  $G_{jg}Y_{jg}$ :

$$\frac{1}{1}\dot{Y}_{jg} = (1 - Y_{jg}) - 40[G_{jg}Y_{jg} - 0.2]^+ \quad (27)$$

Available  $\text{Ca}^{2+}$  passively accumulates at a rate given by the term  $(1 - Y_{jg})$ . Term  $-40[G_{jg}Y_{jg} - 0.2]^+$  describes the depletion of intracellular  $\text{Ca}^{2+}$  as a consequence of the calcium spike. Once depleted, endoplasmic and other available  $\text{Ca}^{2+}$  stores remain low for as long as the SD cell continues to receive a tonic input. Subsequent calcium spikes occur only after a recovery period has passed.

SD cells are capable of generating  $\text{Ca}^{2+}$  spikes,  $G_{jg}Y_{jg}$ , at a spectrum of delays after CS onset. During the peak phase of a  $\text{Ca}^{2+}$  spike,  $[G_{jg}Y_{jg} - 0.2]^+$ , inhibitory transmission from SD cells to the SNc/VTA is boosted. LTM weights,  $Z_{gj}$ , (see Equation (46)) strengthen the influence of these delayed spikes on neural transmission, allowing SD cells to generate an adaptively timed output,  $Z_{gj}[G_{jg}Y_{jg} - 0.2]^+$ .

*SNc/VTA output STM Equations.* Activities of model neurons in the dopaminergic midbrain are represented by  $S_1$ . Neurons in the SNc/VTA receive excitatory driving inputs,  $P_1$ , from PPTN/LDT cells and an adaptively timed, inhibitory input from SD cells,  $Z_{gj}[G_{jg}Y_{jg} - 0.2]^+$ :

$$\frac{1}{50}\dot{S}_1 = -S_1 + (1 - S_1)(10[P_1 - 0.03]^+ + \alpha_2) - (0.1 + S_1)\left(\sum_g \sum_j Z_{gj}[G_{jg}Y_{jg} - 0.2]^+\right) \quad (28)$$

The excitatory and inhibitory shunting terms  $(1 - S_1)$  and  $(0.1 + S_1)$ , respectively, bound cell activity within the interval  $[-0.1, 1]$ . Term  $10[P_1 - 0.03]^+$  describes the driving excitatory input from the PPTN. Term  $\alpha_2$  is a tonic arousal input that ensures baseline dopamine activity is non-zero, providing some dynamic range for dopamine dips. The arousal parameter,  $\alpha_2$ , is set equal to 0.28. The driving, inhibitory term  $\sum_g \sum_j Z_{gj}[G_{jg}Y_{jg} - 0.2]^+$  represents the sum of adaptively timed signals from SD cells across all spectral delays,  $g$  and cue preferences,  $j$ . This term enables dopamine signals to be sensitive to the learned expectations of SD cells regarding the predicted occurrence of rewards and their expected time of delivery.

*Dopamine reinforcement signals.* The effective dopamine signal is determined using transient deviations of dopamine signals from a tonic or baseline dopamine level,  $S_2$ :

$$\frac{1}{5}\dot{S}_2 = S_1 - S_2 \quad (29)$$

Equation (29) has a much slower rate constant than that used in Equation (28). This allows  $S_2$  to compute a time-average of momentary dopamine cell activity,  $S_1$ .

Transient deviations from the baseline signal  $S_2$  constitute phasic dopamine reinforcement signals (Wickens et al., 1996). Equations (30)-(33) describe these phasic

dopamine signals. Equations (30) and (31) define the effective dopamine burst,  $N_1$ , and dopamine dip,  $N_2$ , signals in the striatum:

$$N_1 = [S_1 - S_2 - 0.02]^+ \quad (30)$$

and

$$N_2 = [S_2 - S_1 - 0.02]^+. \quad (31)$$

Dopamine cell activity,  $S_1$ , elicits a dopamine burst or dip ( $N_1$  or  $N_2$ ) when it exceeds baseline dopamine activity,  $S_2$ , by an amount in excess of the threshold parameter, 0.0175. This threshold helps prevent minor fluctuations in DRIVE and GUS inputs from controlling learning.

Equations (32) and (33) define the effective dopamine burst,  $D_1$ , and dopamine dip,  $D_2$ , signals in the AMYG and ORB where dopamine clearance is slower than in the striatum (Garris and Rebec, 2002):

$$\frac{1}{3}\dot{D}_1 = -D_1 + 20N_1 \quad (32)$$

and

$$\frac{1}{3}\dot{D}_2 = -D_2 + 20N_2. \quad (33)$$

Terms  $20N_1$  and  $20N_2$  are dopamine burst and dip inputs.

#### 4.4 Medium Term Memory Equations.

**4.4.1 LH\_in to LH\_gus Drive Habituation.** The habituating transmitter term  $Y_{si}$  gates signals from LH\_in cells to LH\_gus cells:

$$\frac{1}{1}Y_{si} = (1 - Y_{si}) - 1.2Y_{si}[H_{si}^I]^+ \quad (34)$$

The habituating process depends on a recovery process,  $(1 - Y_{si})$ , and an activity-dependent depletion process,  $-1.2Y_{si}[H_{si}^I]^+$ . Activity-dependent depletion is driven by LH\_in cell activity,  $H_{si}^I$ . The rate constant of Equation (34) has a value of 1, hence habituating MTM processes adapt to inputs much more slowly than do STM processes. The habituating transmitter term,  $Y_{si}$ , plays an important role generating rebound responses in the LH and gated dipole circuit (Grossberg, 1972b; see Section 2.3.3 2.1.3). During the first phase of every task, habituating transmitter levels,  $Y_{si}$ , adapt to the tonic arousal inputs that drive the early activity of LH\_in cells. When transient, stimulus-related signals activate different LH\_in cells, the habituated transmitter gates of the LH\_in cells become imbalanced. When an arousal burst occurs or if these transient signals are removed, the imbalanced habituating transmitter gates can persist, during which time rebound responses can result from opponent processing in the LH circuit.

**4.4.2 GUS to LH\_gus Taste Habituation.** GUS inputs to RHIN cells,  $R_i$ , and LH\_gus cells,  $H_{si}^G$  are gated by a habituating transmitter,  $X_m$ :

$$200\dot{X}_m = 0.01(1 - X_m) - (G_m X_m). \quad (35)$$

The habituation involves a slow recovery process,  $0.01(1 - X_m)$  and a faster depletion process,  $-(G_m X_m)$ . Depletion occurs in a consumption dependent fashion and is driven by GUS inputs,  $G_m$ . The taste-specific habituation of GUS inputs,  $X_m$ , is a second mechanism of food specific

satiety (FSS) that complements the model's "drive reduction" mechanism. This second mechanism of FSS can explain the observation that the kind of FSS that is assessed by actual consumption rather than by CS preference remains intact even after ORB and AMYG lesions (Hatfield et al., 1996; Dunn and Everitt, 1988).

#### 4.5 Long Term Memory Equations.

*4.5.1 AMYG to ORB LTM Weights.* Connections from AMYG cells to ORB cells have LTM weights,  $W_{rji}^{AO}$ , that obey a dopamine-modulated instar learning rule (see Equation (9)). LTM weights  $W_{rji}^{AO}$  assign incentive motivation to object-value cells in the ORB:

$$\frac{1}{5} \dot{W}_{rji}^{AO} = \left[ O_{ji}^L \right]^+ \left[ (D_1 + 0.01) \left( [A_{ri}]^+ - W_{rji}^{AO} \right) - (D_2) 2W_{rji}^{AO} \right] \quad (36)$$

Postsynaptic ORB activity,  $[O_{ji}^L]^+$  gates the learning and decay of LTM weights. Term  $(D_1 + 0.01) \left( [A_{ri}]^+ - W_{rji}^{AO} \right)$  describes how dopamine spikes,  $D_1$ , modulate the learning of presynaptic values,  $[A_{ri}]^+$ , by the LTM weights,  $W_{rji}^{AO}$ . In the absence of dopamine, this learning process occurs at a base rate of 0.01. Term  $-(D_2) 2W_{rji}^{AO}$  indicates that LTM weights,  $W_{rji}^{AO}$ , decay when a dopamine dip occurs,  $D_2$ .

*4.5.2 LH\_out to AMYG LTM Weights.* Connections from LH\_out cells to AMYG cells have LTM weights,  $W_{rsi}^{HA}$ , that obey an activity-gated, steepest descent instar learning rule (see Equation (7)). These weights gate the convergence of LH\_out homeostatic signals onto US-specific drive value category cells in the AMYG:

$$\frac{8}{1} \dot{W}_{rsi}^{HA} = [A_{ri} - 0.027]^+ \left( [H_{si}^O]^+ - W_{rsi}^{HA} \right) \quad (37)$$

Term  $[A_{ri} - 0.027]^+$  provides a postsynaptic gate on learning while the term  $\left( [H_{si}^O]^+ - W_{rsi}^{HA} \right)$  describes the process by which LTM weights learn to reflect the activity of LH\_out cells. Term  $[A_{ri} - 0.027]^+$  specifies that AMYG cell activity must exceed a threshold of 0.027 before learning can take place. This threshold was selected to ensure learning takes when the AMYG is activated by cortical inputs, and not simply by feedback from the positive recurrent signal function,  $g(A)$ .

*4.5.3 ITA to AMYG LTM Weights.* Connections from ITA cells to AMYG cells have LTM weights,  $W_{jri}^{IA}$ , that follow a dopamine-modulated outstar learning rule (see Equation (8)). LTM weights  $W_{jri}^{IA}$  enable object categories represented in the ITA to acquire conditioned reinforcer properties:

$$\frac{1}{5} \dot{W}_{jri}^{IA} = [I_j]^+ \left[ (D_1 + 0.01) \left( [A_{ri}]^+ - W_{jri}^{IA} \right) - (D_2) 2W_{jri}^{IA} \right] \quad (38)$$

Presynaptic ITA activity,  $[I_j]^+$ , gates the learning and decay of LTM weights. Term  $(D_1 + 0.01) \left( [A_{ri}]^+ - W_{jri}^{IA} \right)$  describes how dopamine spikes,  $D_1$ , modulate the learning of postsynaptic activation,  $[A_{ri}]^+$ . In the absence of dopamine, the learning process occurs at a

base rate of 0.01. Term  $-(D_2)2W_{jri}^{IA}$  indicates that LTM weights decay when a dopamine dip occurs,  $D_2$ .

**4.5.4 RHIN to AMYG LTM Weights.** Connections from RHIN cells to AMYG cells have LTM weights,  $W_{tri}^{RA}$ , that are governed by dopamine-modulated, outstar learning (Equation (8)). LTM weights  $W_{tri}^{RA}$  enable multimodal food reward categories cells in RHIN to acquire reinforcer properties:

$$\frac{1}{5}\dot{W}_{tri}^{RA} = [R_t]^+ \left[ (D_1 + 0.01) \left( [A_{ri}]^+ - W_{tri}^{RA} \right) - (D_2)2W_{tri}^{RA} \right] \quad (39)$$

Pre-synaptic RHIN cell activity  $[R_t]^+$  gates learning and decay processes for the LTM weights,  $W_{tri}^{RA}$ . Term  $(D_1 + 0.01) \left( [A_{ri}]^+ - W_{tri}^{RA} \right)$  describes the dopamine-modulated process by which weights,  $W_{tri}^{RA}$ , come to reflect postsynaptic AMYG activity,  $[A_{ri}]^+$ . In the absence of dopamine, this process takes place at a rate of 0.01. Term  $-(D_2)2W_{tri}^{RA}$  indicates dopamine dips induce the decay of LTM weights.

**4.5.5 AMYG to LH\_out LTM Equations.** Connections from AMYG cells to LH\_out cells have LTM weights,  $W_{sri}^{AH}$ , that obey an activity-gated, steepest-descent outstar learning law (see Equation (6)). LTM weights  $W_{sri}^{AH}$  enable US-specific drive value categories in the AMYG to excite drive cells in the LH activated by the consumption of specific food rewards:

$$\frac{8}{1}\dot{W}_{sri}^{AH} = [A_{ri} - 0.027]^+ \left( [H_{si}^O]^+ - W_{sri}^{AH} \right) \quad (40)$$

Term  $[A_{ri} - 0.027]^+$  provides a presynaptic gate on learning. Term  $\left( [H_{si}^O]^+ - W_{sri}^{AH} \right)$  describes the process by which LTM weights learn to reflect LH\_out cell activity,  $H_{si}^O$ . Term  $[A_{ri} - 0.027]^+$  specifies that AMYG cell activity must exceed a threshold of 0.027 before learning can take place. This process allows LTM weights  $W_{sri}^{AH}$  to learn to encode a prototype of the metabolic activations in the LH associated with US consumption.

**4.5.6 AMYG to MORB LTM Equations** Connections from AMYG cells to MORB cells have LTM weights,  $W_{rti}^{AM}$ , that obey a dopamine-modulated instar learning rule (Equation (9)). LTM weights  $W_{rti}^{AM}$  assign incentive motivational value to food rewards represented in the MORB:

$$\frac{1}{5}\dot{W}_{rti}^{AM} = [O_{ti}^M]^+ \left[ (D_1 + 0.01) \left( [A_{ri}]^+ - W_{rti}^{AM} \right) - (D_2)2W_{rti}^{AM} \right] \quad (41)$$

Postsynaptic MORB activity,  $[O_{ti}^M]^+$ , gates learning and decay. Term  $(D_1 + 0.01) \left( [A_{ri}]^+ - W_{rti}^{AM} \right)$  indicates dopamine bursts modulate the learning of presynaptic activity,  $[A_{ri}]^+$ , by LTM weights,  $W_{rti}^{AM}$ . The term  $-(D_2)2W_{rti}^{AM}$  indicates that LTM weights decay when a dopamine dip occurs,  $D_2$ .

**4.5.7 ORB to VS LTM Equations.** Connections from ORB cells to VS cells have LTM weights,  $W_{jri}^{OV}$ , that obey a dopamine-gated instar rule (Brown et al., 2004; Equation (9)). LTM weights  $W_{jri}^{OV}$  allow conditioned stimuli to strongly activate dopaminergic responses:

$$\frac{1}{10}W_{jri}^{OV} = V_{ri} \left[ N_1(2.5O_{ji}^L - W_{jri}^{OV}) - 0.2N_2W_{jri}^{OV} \right]. \quad (42)$$

Post-synaptic VS activity,  $V_{ri}$ , gates learning and decay. Term  $N_1(2.5O_{ji}^L - W_{jri}^{OV})$  indicates that dopamine bursts,  $N_1$ , doubly gate steepest descent learning. LTM weights  $W_{jri}^{OV}$  learn to reflect the value of postsynaptic ORB cells,  $2.5O_{ji}^L$ . Term  $-0.2N_2W_{jri}^{OV}$  indicates that dopamine dips,  $N_2$ , gate weight decay.

**4.5.8 Adaptive Timing LTM Equations.** Adaptive LTM weights  $Z_{jg}$  determine the strength of timed inhibitory outputs from SD cells to the SNc/VTA. These LTM weights,  $Z_{jg}$ , adapt according to an equation similar to dopamine-gated steepest descent:

$$\frac{1}{300}\dot{Z}_{jg} = \left[ G_{jg}Y_{jg} - 0.2 \right]^+ (N_1 - Z_{jg}N_2) \quad (44)$$

Term  $\left[ G_{jg}Y_{jg} - 0.2 \right]^+$  gates learning, ensuring learning occurs only during the peak phase of a  $Ca^{2+}$  spike in an SD cell. Term  $(N_1 - Z_{jg}N_2)$  describes the process by which weights,  $Z_{jg}$  grow when dopamine spikes are present and decay when dopamine dips are present. The growth and decay of LTM weights,  $Z_{jg}$ , tracks reward history, inhibiting dopamine responses to predictable rewards. This ensures dopamine signals generated by the SNc/VTA reflect a reward prediction error.

## 4.6 Outputs.

**4.6.1 Blood Pressure Output.** The blood pressure response, BPR, is influenced by appetitive and aversive stimuli (Braesicke, et al. 2005). This influence is relayed from the hypothalamus to cardiovascular regulatory neurons in the medulla (Smith et al., 1990; Nakamura et al., 1992; Zhang et al., 2005, 2006). The component of the blood pressure variable attributable to appetitive or aversive stimuli,  $B$ , is calculated as follows:

$$\dot{B} = -B + 15 \sum_i \left[ H_{si}^O - \frac{1}{4} \sum_{s=1}^4 H_{si}^O \right]^+, \quad (44)$$

$$\text{Where } BPR = 120 + B. \quad (45)$$

The term  $15 \sum_i \left[ H_{si}^O - \frac{1}{4} \sum_{s=1}^4 H_{si}^O \right]^+$  relays excitatory information from LH\_out cells regarding stimuli and rewards. The blood pressure response, BPR, reflects the integral of the LH input added to a baseline blood pressure level of 120.

**4.6.2 Saccadic Output.** Saccadic responses are generated by object category-selective cells in the FEF. In vivo, the FEF receives projections from both the ORB and ITA (Barbas, 1992; Bullier et al., 1996). Correspondingly, in the model, FEF cells integrate inputs from the ORB and ITA along with arousal inputs. When an FEF cell activity  $F_j$  exceeds a threshold of 0.3, a saccade was elicited in response to the corresponding cue. Only one saccade could be elicited per trial:



$$10\dot{F}_j = -F_j + 4[O_{j1}]^+ + [I_j]^+ + \alpha_3 + \varepsilon_j. \quad (46)$$

Activity  $F_j$  passively decays at rate 1 via term  $-F_j$ . Additive excitatory terms  $4[O_{j1}]^+$  and  $[I_j]^+$  describe inputs that arise from the ORB and ITA. Terms  $\alpha_3$  and  $\varepsilon_j$  are constants that take on non-zero values during stimulus presentation.

Term  $\alpha_3$  is an arousal input that activates the FEF depending on whether or not an instrumental response is required to gain reward. For the SVD Task where the CS was presented for 450 ms before a saccade was made,  $\alpha_3 = 6.5$ . For the Pavlovian CS Task for which the CS was presented for 2 seconds,  $\alpha_3 = 0$ . This arousal term helps to realize the different response and timing requirements of the CS and SVD tasks. In the CS task,  $\alpha_3 = 0$  because responses are incidental to the acquisition of reward and may be generated anytime across the 2 second CS presentation time, or not at all. For the SVD task, a response must be made within 400 ms in order to gain reward. The choice of  $\alpha_3 = 6.5$  ensures that a decision is made within this time frame.

The noise terms  $\varepsilon_j$  are constants that break symmetry during decision-making. Values for  $\varepsilon_j$  are randomly selected at the start of each trial from a uniform distribution over the interval  $[0, 1]$ . The noise terms acknowledge that signals from unknown sources can influence behavior and decision-making, generating behavioral variation or breaking the symmetry between closely matched options. In tasks such as the SVD reversal task, where stimulus and response contingencies change, the response variability introduced by the term,  $\varepsilon_j$ , can drastically speed the learning of new associations.

In essence, model FEF cells function as cumulative spike counters, integrating ORB and ITA activity along with some noise and arousal inputs to elicit saccades using a “race to threshold” rule (Schall and Thompson, 1999). For a more detailed representation of the FEF, see Brown et al. (2004). Parameter values were selected to ensure reaction times during the SVD task fall between 300ms and 450ms and that no responses are made in the CS task prior to learning taking place. ITA inputs play a critical role in driving saccadic behavior when ORB inputs are silent or after ORB has been lesioned. Such behaviors are simulated in Grossberg et al. (2008).

#### 4.7 Initialization of Variables.

STM equations for cells that process drive and arousal inputs and are persistently active (LH cell Equations (15-17), PPTN and SNc/VTA Equations (21-22, 28)) were initialized to the equilibrium level expected from the previous trial or from their initial inputs. All other STM equations were set to an initial value of zero at the start of each trial.

MTM equations describing gustatory habituation,  $X_m$  (Equation (35)), were initialized to an initial value of 1, the equilibrium level that would be reached after a long period of disuse. The MTM equations describing habituating gates associated with LH\_in cell hunger and satiety channels,  $Y_{si}$  (Equation (34)), were set to initial values of 0.5 for hunger-related and 1 for satiety-related gates.

With the exception of LTM weights ( $W_{rsi}^{HA}$ ,  $W_{sri}^{AH}$ ; Equations (37, 40)) linking AMYGD and LH output cells, all LTM weights were initialized to zero at the start of training. On

subsequent trials and tasks, all LTM weights were assigned the final value of the LTM weights reached in the previous trial as an initial value. At the start of training, LTM weights linking the AMYG and LH values were set to an initial value of 0.22 with noise added from a uniform random distribution over the interval  $[-0.05 \ 0.05]$ . The noise served to break the initial symmetry across the AMYG and LH cell connections. Nonzero initial LTM weight values ensure that AMYG drive value category cells are sensitive to LH inputs prior to learning (see Equation (12)). The magnitude of the initial values for the LTM weights was also selected to ensure that there is a large degree of mismatch between the patterns of LH activation that are recognized by well trained AMYG drive value category cells and novel patterns of LH input.

## 5. REFERENCES

- Amaral, D.G., Price, J.L., 1984. Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J Comp Neurol.* 230, 465-496.
- Amsel, A., 1968. Secondary reinforcement and frustration. *Psychol Bull.* 69, 278.
- Balliène, B.W., 2005. Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol Behav.* 86, 717-30.
- Balleine, B.W., Garner, C., Gonzalez, F., Dickinson, A., 1995. Motivational control of heterogeneous instrumental chains. *J Exp Psychol: Anim Beh Proc.* 21, 203-217.
- Barbas, H., 1992. Architecture and cortical connections of the prefrontal cortex in the rhesus monkey. *Adv Neurol.* 57, 91-115.
- Barbas, H., 1993. Organization of cortical afferent input to orbitofrontal areas in the rhesus monkey. *Neurosci.* 56, 841-864.
- Barbas, H., 2000. Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Res Bull.* 52, 319-330.
- Barbas, H., Ghashghaei, H., Dombrowski, S.M., Rempel-Clower, N.L., 1999. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *J Comp Neurol.* 410, 343-367.
- Barbas, H., Saha, S., Rempel-Clower, N., Ghashghaei, T., 2003. Serial pathways from primate prefrontal cortex to autonomic areas may influence emotional expression. *BMC Neurosci.* 4, 25.
- Barefoot, H.C., Baker H.F., Ridley R.M., 2000. Synergistic effects of unilateral immunolesions of the cholinergic projections from the basal forebrain and contralateral ablations of the inferotemporal cortex and hippocampus in monkeys. *Neurosci.* 98, 243-251.
- Baxter, M.G., Parker, A., Lindner, C.C., Izquierdo, A.D., Murray, E.A., 2000. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci.* 20, 4311-4319.
- Baylis, L.L., Gaffan, D., 1991. Amygdectomy and ventromedial prefrontal ablation produce similar deficits in food choice and in simple object discrimination learning for an unseen reward. *Exp Brain Res.* 86, 617-622.
- Bergen, H.T., Monkman, N., Mobbs, C.V., 1996. Injection with gold thioglucose impairs sensitivity to glucose: Evidence that glucose-responsive neurons are important for long-term regulation of body weight. *Brain Res.* 734, 332-336.
- Bernardis, L.L., Bellinger, L.L., 1996. The lateral hypothalamic area revisited: Ingestive behavior. *Neurosci Biobehav Rev.* 20, 189-287.
- Berridge, K.C., 2001. Reward learning: reinforcement, incentives, and expectations. In *The Psychology of Learning and Motivation. Advances in Research and Theory*, Vol 40, Medin, ed. Academic Press, San Diego, pp. 223-278.
- Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Res Rev.* 28, 309-369.
- Bertrand, D., Changeux, J.P., 1995. Nicotinic receptor: An allosteric protein specialized for intercellular communication. *Sem Neurosci.* 7, 75-90.
- Bichot, N.P., Thompson, K.G., Chenchal Rao, S., Schall, J.D., 2001. Reliability of macaque frontal eye field neurons signaling saccade targets during visual search. *J Neurosci.* 21, 713-725.

- Braesicke, K., Parkinson, J.A., Reekie, Y., Man, M.S., Hopewell, L., Pears, A., Crofts, H., Schnell, C.R., Roberts, A.C., 2005. Autonomic arousal in an appetitive context in primates: A behavioural and neural analysis. *Eur J Neurosci.* 21, 1733-1740.
- Brown, J.W., Bullock, D., Grossberg, S., 1999. How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *J Neurosci.* 19, 10502-10511.
- Brown, J.W., Bullock, D., Grossberg, S., 2004. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Netw.* 17, 471-510.
- Brown, T.H., Kairiss, E.W., Keenan, C.L., 1990. Hebbian synapses: Biophysical mechanisms and algorithms. *Annu Rev Neurosci.* 13, 475-511.
- Buffalo, E.A., Ramus, S.J., Clark, R.E., Teng, E., Squire, L.R., Zola, S.M., 1999. Dissociation between the effects of damage to perirhinal cortex and area TE. *Learn Mem.* 6, 572-599.
- Bullier, J., Schall, J.D. and Morel, A. 1996. Functional streams in occipito-frontal connections in the monkey. *Behavioral Brain Research*, 76, 89-97.
- Burwell, R.D., Amaral, D.G., 1998. Cortical afferents of the perirhinal, postrhinal, and entorhinal cortices of the rat. *J Comp Neurol.* 398, 179-205.
- Busemeyer, J.R., Townsend, J.T., Stout, J.C., 2002. Motivational Underpinnings of Utility in Decision Making: Decision Field Theory Analysis of Deprivation and Satiation. John Benjamins, Amsterdam.
- Cabanac, M., Lafrance, L., 1992. Duodenal preabsorptive origin of gustatory alliesthesia in rats. *Am J Physiol.* 263, R1013-1017.
- Cardinal, R.N., Parkinson, J.A., Hall, J., Everitt, B.J., 2002. Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev.* 26, 321-352.
- Carmichael, S.T., Price, J.L., 1995. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol.* 363, 615-641.
- Carpenter, G., Grossberg, S., 1987. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Comput Vis, Graphics, and Image Proc.* 37, 54-115.
- Carpenter, G.A., Grossberg, S., 1991. *Pattern Recognition by Self-Organizing Neural Networks.* MIT Press, Cambridge MA.
- Chance, F.S., Abbott, L.F., Reyes, A.D., 2002. Gain modulation from background synaptic input. *Neuron.* 35, 773-782.
- Cohen, J.D., McClure, S.M., Yu, A.J., 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos Trans R Soc Lond B Biol Sci.* 362, 933-42.
- Conover, K.L., Shizgal, P., 1994. Differential effects of postingestive feedback on the reward value of sucrose and lateral hypothalamic stimulation in rats. *Behav Neurosci.* 108, 559-572.
- Conover, K.L., Woodside, B., Shizgal, P., 1994. Effects of sodium depletion on competition and summation between rewarding effects of salt and lateral hypothalamic stimulation in the rat. *Behav Neurosci.* 108, 549-558.
- Corbit, L.H., Balleine, B.W., 2003. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogenous instrumental chain. *J Exp Psych Anim Behav Process.* 29, 99-106.

- Corbit, L.H., Balleine, B.W., 2005. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci.* 25, 962-70.
- Corbit, L.H., Muir, J.L., Balleine, B.W., 2001. The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J Neurosci.* 21, 3251-3260.
- Critchley, H.D., Rolls, E.T., 1996. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *J Neurophysiol.* 75, 1673-1686.
- Davidson, T.L., Altizer, A.M., Benoit, S.C., Walls, E.K., Powley, T.L., 1997. Encoding and selective activation of "metabolic memories" in the rat. *Behav Neurosci.* 111, 1014-1030.
- Dayan, P., 2001. Motivated reinforcement learning. *NIPS.* 14, 11-18.
- Deco, G., Rolls, E.T., 2005. Synaptic and spiking dynamics underlying reward reversal in the orbitofrontal cortex. *Cereb Cortex.* 15, 15-30.
- DeFalco, J., Tomishima, M., Liu, H., Zhao, C., Cai, X., Marth, J.D., Enquist, L., Friedman, J.M., 2001. Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. *Science.* 291, 2608-2613.
- Denny, M.R., 1970. Relaxation theory and experiments. In *Aversive Conditioning and Learning*, F.R. Brush, ed. Academic Press, New York.
- Dickinson, A., Dearing, M.F., 1979. Appetitive-Aversive Interactions and Inhibitory Processes. Erlbaum, Hillsdale NJ.
- Dickinson, A., Balleine, B.W., 2001. The role of learning in the operation of motivational systems. In *Steven's Handbook Of Experimental Psychology*, 3rd Edition, H.E. Pashler, R. Gallistel, eds. John Wiley & Sons, New York, pp 497-533.
- Dranias, M.R., Bullock, D., Grossberg, S., 2006. A neural network model of conditioning and outcome-specific revaluation: interactions of the orbitofrontal cortex, amygdala, hypothalamus, and basal ganglia. *Soc for Neurosci.* 2006. 216.13. Atlanta, GA.
- Dranias, M.R., Bullock, D., Grossberg, S., 2007a. Emotional and behavioral responses to the devaluation of stimuli by satiation: A computational model. *Computat and Systems Neurosci.* III-68. Salt Lake City, UT.
- Dranias, M.R., Bullock, D., Grossberg, S., 2007b. Neural dynamics of conditioning and outcome-specific revaluation: Cortical, amygdala, hypothalamic, and basal ganglia. interactions. *Eleventh Internatl Conf on Cognit and Neural Syst.* Boston, MA .
- Dunn, L.T., Everitt, B.J., 1988. Double dissociations of the effects of amygdala and insular cortex lesions on conditioned taste aversion, passive avoidance, and neophobia in the rat using the excitotoxin ibotenic acid. *Behav Neurosci.* 102, 3-23.
- Easton, A., Gaffan, D., 2000. Comparison of perirhinal cortex ablation and crossed unilateral lesions of the medial forebrain bundle from the inferior temporal cortex in the rhesus monkey: Effects on learning and retrieval. *Behav Neurosci.* 114, 1041-1057.
- Ferry, A.T., Ongur, D., An, X., Price, J.L., 2000. Prefrontal cortical projections to the striatum in macaque monkeys: Evidence for an organization related to prefrontal networks. *J Comp Neurol.* 425, 447-470.
- Fiala, J.C., Grossberg, S., Bullock, D., 1996. Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. *J Neurosci.* 16, 3760-3774.

- Fiorillo, C.D., Tobler, P.N., and Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*. 299, 1899-1902.
- Fiorillo, C.D., Tobler, P.N., and Schultz, W., 2005. Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behav Brain Func.* 1, 7.
- Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev.* 113, 300-326.
- Friedman, D.P., Aggleton, J.P., Saunders, R.C., 2002. Comparison of hippocampal, amygdala, and perirhinal projections to the nucleus accumbens: Combined anterograde and retrograde tracing study in the Macaque brain. *J Comp Neurol.* 450, 345-365.
- Fukuda, M., Ono, T., Nakamura, K., 1987. Functional relations among inferotemporal cortex, amygdala, and lateral hypothalamus in monkey operant feeding behavior. *J Neurophysiol.* 57, 1060-1077.
- Gaffan, D., 1994. Dissociated effects of perirhinal cortex ablation, fornix transection and amygdectomy: Evidence for multiple memory systems in the primate temporal lobe. *Exp Brain Res.* 99, 411-422.
- Gallagher, M., McMahan, R. W., Schoenbaum, G., 1999. Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci.* 19, 6610-6614.
- Garris, P.A., Rebec, G.V., 2002. Modeling fast dopamine neurotransmission in the nucleus accumbens during behavior. *Behav Brain Res.* 137, 47-63.
- Ghashghaei, H.T., Barbas, H., 2002. Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neurosci.* 115, 1261-1279.
- Gove, A., Mingolla, E., Grossberg, S., 1995. Brightness perception, illusory contours, and corticogeniculate feedback. *Vis Neurosci.* 12, 1027-1052.
- Grossberg, S., 1968. Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity. *Proc. Natl. Acad. Sci.* 59, 368-372.
- Grossberg, S., 1971. On the dynamics of operant conditioning. *Journal of Theoretical Biology.* 33, 225-255.
- Grossberg, S., 1972a. A neural theory of punishment and avoidance, I: Qualitative theory. *Math Biosci.* 15, 39-67.
- Grossberg, S., 1972b. A neural theory of punishment and avoidance, II: Quantitative theory. *Math Biosci.* 15, 253-285.
- Grossberg, S., 1973. Contour enhancement, short-term memory, and constancies in reverberating neural networks. *S App Math.* 52, 213-257
- Grossberg, S., 1975. A neural model of attention, reinforcement and discrimination learning. *Int Rev Neurobiol.* 18, 263-327.
- Grossberg, S., 1976. Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biol Cyb.* 23, 121-134.
- Grossberg, S., 1980. How does a brain build a cognitive code? *Psychol Rev.* 87, 1-51.
- Grossberg, S. (1982). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. *Psychol Rev.* 89, 529-572.
- Grossberg, S., 1984. Some psychophysiological and pharmacological correlates of a developmental, cognitive, and motivational theory. In R. Karrer, J. Cohen, and P. Tueting eds. *Brain and information: Event related potentials*. New York : New York Academy of Sciences, pp 58-142.

- Grossberg, S., 1988. Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Nets*. 1, 17-61.
- Grossberg, S., 1999. How does the cerebral cortex work? Learning, attention and grouping by the laminar circuits of visual cortex. *Spat Vis*. 12, 163-186.
- Grossberg, S., 2000a. The imbalanced brain: From normal behavior to schizophrenia. *Biol Psychiatry*. 48, 81-98.
- Grossberg, S., 2000b. The complementary brain: Unifying brain dynamics and modularity. *Trends Cogn Sci*. 4, 233-246.
- Grossberg, S., 2003. How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex. *Behav and Cogn Neurosci Revs*. 2, 47-76.
- Grossberg, S., Bullock, B., Dranias, M. R., 2008. Neural dynamics underlying impaired autonomic and conditioned responses following amygdala and orbitofrontal lesions. *Behav Neuro*. In press.
- Grossberg, S., Gutowski, W.E., 1987. Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychol Rev*. 94, 300-318.
- Grossberg, S. and Levine, D.S. (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. *Applied Optics* 26, 5015-5030.
- Grossberg, S., Levine, D., Schmajuk, N., 1987. Predictive regulation of associative learning in a neural network by reinforcement and attentive feedback. *Int J Neurol*. 21-22, 83-104.
- Grossberg, S., Merrill, J.W., 1992. A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cogn Brain Res*. 1, 3-38.
- Grossberg, S. and Merrill, J.W.L., 1996. The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *J Cog Neurosci*. 8, 257-277.
- Grossberg, S., Schmajuk, N.A., 1987. Neural dynamics of attentionally modulated pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. *Psychobiol*. 15, 195-240.
- Grossberg, S. and Schmajuk, N.A. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. *Neural Networks*, 2, 79-102.
- Haber, S.N., Lynd-Balta, E., Mitchell, S.J., 1993. The organization of the descending ventral pallidal projections in the monkey. *J Comp Neurol*. 329, 111-128.
- Haber, S.N., Fudge, J.L., McFarland, N.R., 2000. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci*. 20, 2369-2382.
- Hall, G., 2001. Conditioning. In *Steven's Handbook of Experimental Psychology*, H.E. Pashler, ed. John Wiley & Sons, New York, pp 4.
- Hatfield, T., Han, J.S., Conley, M., Gallagher, M., Holland, P., 1996. Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci*. 16, 5256-5265.
- Hikosaka, K., Watanabe, M., 2000. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb Cortex*. 10, 263-271.
- Hodgkin, A.L., Huxley, A.F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*. 117, 500-544.
- Hosokawa, T., Kato, K., Inoue, M., Mikami, A., 2004. Neurons in the orbitofrontal cortex code both visual shapes and reward types. *Neuroreport*. 15, 1493-1496.



- Insausti, R., Amaral, D.G., Cowan, W.M., 1987. The entorhinal cortex of the monkey: II. Cortical afferents. *J Comp Neurol.* 264, 356-395.
- Jagadeesh, B., Chelazzi, L., Mishkin, M., Desimone, R., 2001. Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. *J Neurophysiol.* 86, 290-303.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica.* 47, 263-291.
- Kantak, K.M., Green-Jordan, K., Valencia, E., Kremin, T., Eichenbaum, H.B., 2001. Cognitive task performance after lidocaine-induced inactivation of different sites within the basolateral amygdala and dorsal striatum. *Behav Neurosci.* 115, 589-601.
- Karadi, Z., Oomura, Y., Nishino, H., Scott, T.R., Lenard, L., Aou, S., 1992. Responses of lateral hypothalamic glucose-sensitive and glucose-insensitive neurons to chemical stimuli in behaving rhesus monkeys. *J Neurophysiol.* 67, 389-400.
- Kobayashi, Y., Inoue, Y., Yamamoto, M., Isa, T., Aizawa, H., 2002. Contribution of pedunculopontine tegmental nucleus neurons to performance of visually guided saccade tasks in monkeys. *J Neurophysiol.* 88, 715-731.
- Kondoh, T., Mori, M., Ono, T., Torii, K., 2000. Mechanisms of umami taste preference and aversion in rats. *J Nutr.* 130, 966S-970S.
- Lauwereyns, J., Watanabe, K., Coe, B., Hikosaka O., 2002. A neural correlate of response bias in monkey caudate nucleus. *Nature.* 418, 413-417.
- Lavoie, B., Parent, A., 1994. Pedunculopontine nucleus in the squirrel monkey: Cholinergic and glutamatergic projections to the substantia nigra. *J Comp Neurol.* 344, 232-241.
- Liu, Z., Richmond, B.J., 2000. Response differences in monkey TE and perirhinal cortex: Stimulus association related to reward schedules. *J Neurophysiol.* 83, 1677-1692.
- Ljungberg, T., Apicella, P., Schultz, W., 1992. Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol.* 67, 145-163.
- Luck, S.J., Chelazzi, L., Hillyard, S.A., Desimone, R., 1997. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol.* 77, 24-42.
- Malkova, L., Gaffan, D., Murray, E.A., 1997. Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. *J Neurosci.* 17, 6011-6020.
- McDonald, A.J., 1998. Cortical pathways to the mammalian amygdala. *Prog Neurobiol.* 55, 257-332.
- Mogami, T., Tanaka, K., 2006. Reward association affects neuronal responses to visual stimuli in macaque TE and perirhinal cortices. *J Neurosci.* 26, 6761-70.
- Muramoto, K., Ono, T., Nishijo, H., Fukuda, M., 1993. Rat amygdaloid neuron responses during auditory discrimination. *Neurosci.* 52, 621-636.
- Murray, E.A., Mishkin, M., 1998. Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *J Neurosci.* 18, 6568-82.
- Murray, E.A., Gaffan, E.A., Flint, R.W., Jr., 1996. Anterior rhinal cortex and amygdala: Dissociation of their contributions to memory and food preference in rhesus monkeys. *Behav Neurosci.* 110, 30-42.
- Nakamura, K., Ono, T., 1986. Lateral hypothalamus neuron involvement in integration of natural and artificial rewards and cue signals. *J Neurophysiol.* 55, 163-181.

- Nakamura, K., Ono, T., Tamura, R., 1987. Central sites involved in lateral hypothalamus conditioned neural responses to acoustic cues in the rat. *J Neurophysiol.* 58, 1123-1148.
- Nakamura, K., Ono, T., Fukuda, M., Uwano, T., 1992. Paraventricular neuron chemosensitivity and activity related to blood pressure control in emotional behavior. *J Neurophysiol.* 67, 255-264.
- Nakano, Y., Oomura, Y., Lenard, L., Nishino, H., Aou, S., Yamamoto, T., Aoyagi, K., 1986. Feeding-related activity of glucose- and morphine-sensitive neurons in the monkey amygdala. *Brain Res.* 399, 167-172.
- Nishijo, H., Ono, T., Nishino, H., 1988a. Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *J Neurosci.* 8, 3570-3583.
- Nishijo, H., Ono, T., Nishino, H., 1988b. Topographic distribution of modality-specific amygdalar neurons in alert monkey. *J Neurosci.* 8, 3556-3569.
- Niv, Y., Duff, M.O., Dayan, P., 2005. Dopamine, uncertainty, and TD learning. *Behav Brain Func.* 1, 6.
- Oakman, S.A., Faris, P.L., Kerr, P.E., Cozzari, C., Hartman, B.K., 1995. Distribution of pontomesencephalic cholinergic neurons projecting to substantia nigra differs significantly from those projecting to ventral tegmental area. *J Neurosci.* 15, 5859-5869.
- Olson, S. and Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Netw.* 11, 189-208.
- Ongur, D., Price, J.L., 2000. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex.* 10, 206-219.
- Ono, T., Nakamura, K., Nishijo, H., Fukuda, M., 1986a. Hypothalamic neuron involvement in integration of reward, aversion, and cue signals. *J Neurophysiol.* 56, 63-79.
- Ono, T., Sasaki, K., Nishino, H., Fukuda, M., Shibata, R., 1986b. Feeding and diurnal related activity of lateral hypothalamic neurons in freely behaving rats. *Brain Res.* 373, 92-102.
- Owens, D.F., Kriegstein, A.R., 2002. Is there more to GABA than synaptic inhibition? *Nat Rev Neurosci.* 3, 715-727.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature.* 441, 223-226.
- Papini, M.R., 2003. Comparative psychology of surprising nonreward. *Brain Behav Evol.* 62, 83-95.
- Parker, A., Gaffan, D., 1998. Lesions of the primate rhinal cortex cause deficits in flavour-visual associative memory. *Behav Brain Res.* 93, 99-105.
- Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature.* 439, 865-870.
- Pecina, S., Berridge, K.C., 1996. Brainstem mediates diazepam enhancement of palatability and feeding: microinjections into fourth ventricle versus lateral ventricle. *Brain Res.* 727, 22-30.
- Petrovich, G.D., Canteras, N.S., Swanson, L.W., 2001. Combinatorial amygdalar inputs to hippocampal domains and hypothalamic behavior systems. *Brain Res Rev.* 38, 247-289.
- Peyron, C., Tighe, D.K., van den Pol, A.N., de Lecea, L., Heller, H.C., Sutcliffe, J.G., Kilduff, T.S., 1998. Neurons containing hypocretin (orexin) project to multiple neuronal systems. *J Neurosci.* 18, 9996-10015.

- Price, J.L., 2003. Comparative aspects of amygdala connectivity. *Ann N Y Acad Sci.* 985, 50-58.
- Reep, R.L., Corwin, J.V., King, V., 1996. Neuronal connections of orbital cortex in rats: Topography of cortical and thalamic afferents. *Exp Brain Res.* 111, 215-232.
- Rempel-Clower, N.L., Barbas, H., 2000. The laminar pattern of connections between prefrontal and anterior temporal cortices in the Rhesus monkey is related to cortical structure and function. *Cereb Cortex.* 10, 851-865.
- Rescorla, R.A., 2001. Retraining of extinguished pavlovian stimuli. *J Exp Psych: Anim Behav Proc.* 27, 115-124.
- Reynolds, S.M., Berridge, K.C., 2002. Positive and negative emotion in nucleus accumbens shell: Bivalent rostrocaudal gradients for GABA-elicited eating, taste "liking"/"disliking" reactions, place preference/avoidance, and fear. *J Neurosci.* 22, 7308-7320.
- Richmond, B.J., Sato, T., 1987. Enhancement of inferior temporal neurons during visual discrimination. *J Neurophysiol.* 58, 1292-1306.
- Risold, P.Y., Thompson, R.H., Swanson, L.W., 1997. The structural organization of connections between the hypothalamus and cerebral cortex. *Brain Res Rev.* 24, 197-254.
- Roesch, M.R., Olson, C.R., 2003. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J Neurophysiol.* 90, 1766-1789.
- Roesch, M.R., Olson, C.R., 2004. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science.* 304, 307-310.
- Rolls, E.T., 2000. The orbitofrontal cortex and reward. *Cereb Cortex.* 10, 284-294.
- Rolls, E.T., Baylis, L.L., 1994. Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *J Neurosci.* 14, 5437-5452.
- Rolls, E.T., Critchley, H.D., Mason, R., Wakeman, E.A., 1996. Orbitofrontal cortex neurons: Role in olfactory and visual association learning. *J Neurophysiol.* 75, 1970-1981.
- Rolls, E.T., Murzi, E., Yaxley, S., Thorpe, S.J., Simpson, S.J., 1986. Sensory-specific satiety: Food-specific reduction in responsiveness of ventral forebrain neurons after feeding in the monkey. *Brain Res.* 368, 79-86.
- Rolls, E.T., Critchley, H.D., Browning, A.S., Hernadi, I., Lenard, L., 1999. Responses to the sensory properties of fat of neurons in the primate orbitofrontal cortex. *J Neurosci.* 19, 1532-1540.
- Rolls, E.T., Sanghera, M.K., Roper-Hall, A., 1979. The latency of activation of neurons in the lateral hypothalamus and substantia innominata during feeding in the monkey. *Brain Res.* 164, 121-135.
- Russchen, F.T., 1986. Cortical and subcortical afferents of the amygdaloid complex. *Adv Exp Med Biol.* 203, 35-52.
- Sah, P., Faber, E.S., Lopez De Armentia, M., Power J., 2003. The amygdaloid complex: Anatomy and physiology. *Physiol Rev.* 83, 803-834.
- Saleem, K.S., Tanaka, K., 1996. Divergent projections from the anterior inferotemporal area TE to the perirhinal and entorhinal cortices in the macaque monkey. *J Neurosci.* 16, 4757-4775.
- Salinas, E., Their, P., 2000. Gain modulation: A major computational principle of the central nervous system. *Neuron.* 27, 15-21.
- Sato, T., Murthy, A., Thompson, K.G., Schall, J.D., 2001. Search efficiency but not response interference affects visual selection in frontal eye field. *Neuron.* 30, 583-591.

- Schall, J.D., Thompson, K.G., 1999. Neural selection and control of visually guided eye movements. *Annu Rev Neurosci.* 22, 241-259.
- Schoenbaum, G., Setlow, B., Saddoris, M.P., Gallagher M., 2003. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron.* 39, 855-867.
- Schultz, W., 2007. Behavioral dopamine signals. *Trends Neurosci.* 30, 203-210.
- Schultz, W., Tremblay, L., Hollerman, J.R., 2000. Reward processing in primate orbitofrontal cortex and basal ganglia. *Cereb Cortex.* 10, 272-284.
- Sclafani, A., 2004. Oral and postoral determinants of food reward. *Physiol Behav.* 81, 773-779.
- Scott, T.R., Yan, J., Rolls, E.T., 1995. Brain mechanisms of satiety and taste in macaques. *Neurobio (Bp).* 3, 281-292.
- Semba, K., Fibiger, H.C., 1992. Afferent connections of the laterodorsal and the pedunculopontine tegmental nuclei in the rat: A retro- and antero-grade transport and immunohistochemical study. *J Comp Neurol.* 323, 387-410.
- Sewards, T.V., Sewards, M.A., 2001. Cortical association areas in the gustatory system. *Neurosci Biobehav Rev.* 25, 395-407.
- Shi, C.J., Cassell, M.D., 1997. Cortical, thalamic, and amygdaloid projections of rat temporal cortex. *J Comp Neurol.* 382, 153-175.
- Shimizu, N., Oomura, Y., Sakata, T., 1984. Modulation of feeding by endogenous sugar acids acting as hunger or satiety factors. *Am J Physiol Regul Integr Comp Physiol.* 246, 542-550.
- Small, D.M., Bender, G., Veldhuizen, M.G., Rudenga, K., Nachtigal, D., Felsted, J., 2007. The Role of the Human Orbitofrontal Cortex in Taste and Flavor Processing. *Ann NY Acad Sci. Annals PrePrint*, published online ahead of print. doi: 10.1196/annals.1401.002 from the Annals volume Linking Affect to Action: Critical Contributions of the Orbitofrontal Cortex <http://www.annalsnyas.org/cgi/rapidpdf/annals.1401.002v1> retrieved 9-29-2007.
- Smith, K.S., Berridge, K.C., 2005. The ventral pallidum and hedonic reward: Neurochemical maps of sucrose "liking" and food intake. *J Neurosci.* 25, 8637-8649.
- Solomon, R.L., 1980.. The opponent process theory of acquired motivation. *Am Psychol.* 35, 691-712.
- Spooren, W.P., Lynd-Balta, E., Mitchell, S., Haber, S.N., 1996. Ventral pallidostriatal pathway in the monkey: Evidence for modulation of basal ganglia circuits. *J Comp Neurol.* 370, 295-312.
- Suzuki, W., Saleem, K.S., Tanaka, K., 2000. Divergent backward projections from the anterior part of the inferotemporal cortex (area TE) in the macaque. *J Comp Neurol.* 422, 206-228.
- Suzuki, W.A., Miller, E.K., Desimone, R., 1997. Object and place memory in the macaque entorhinal cortex. *J Neurophysiol.* 78, 1062-1081.
- Swanson, L.W., 2000. Cerebral hemisphere regulation of motivated behavior. *Brain Res.* 886, 113-164.
- Tan, C.O., Bullock, D. 2008a. Neuropeptide co-release with GABA may explain functional non-monotonic uncertainty responses in dopamine neurons. *Neurosci Lett.* 430: 218-23.
- Tan, C. O. and Bullock, D., 2008b. A local circuit model of learned striatal and dopamine cell responses under probabilistic schedules of reward. *J Neurosci.* Submitted for publication.

- Thorpe, S.J., Rolls, E.T., Maddison, S., 1983. The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Exp Brain Res.* 49, 93-115.
- Torii, K., Kondoh, T., Mori, M., Ono, T., 1998. Hypothalamic control of amino acid appetite. *Ann N Y Acad Sci.* 855, 417-425.
- Touzani, K., Sclafani, A., 2002. Lateral hypothalamic lesions impair flavour-nutrient and flavour-toxin trace learning in rats. *Eur J Neurosci.* 16, 2425-2433.
- Tremblay, L., Schultz, W., 2000a. Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *J Neurophysiol.* 83, 1877-1885.
- Tremblay, L., Schultz, W., 2000b. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. *J Neurophysiol.* 83, 1864-1876.
- Usuda, I., Tanaka, K., Chiba, T., 1998. Efferent projections of the nucleus accumbens in the rat with special reference to subdivision of the nucleus: Biotinylated dextran amine study. *Brain Res.* 797, 73-93.
- van den Pol, A.N., Trombley, P.Q., 1993. Glutamate neurons in hypothalamus regulate excitatory transmission. *J Neurosci.* 13, 2829-2836.
- Veazey, R.B., Amaral, D.G., Cowan, W.M., 1982. The morphology and connections of the posterior hypothalamus in the cynomolgus monkey (*Macaca fascicularis*). II. Efferent connections. *J Comp Neurol.* 207, 135-156.
- Voytko, M.L., 1986. Visual learning and retention examined with reversible cold lesions of the anterior temporal lobe. *Behav Brain Res.* 22, 25-39.
- Wachsmuth, E., Oram, M.W., Perrett, D.I., 1994. Recognition of objects and their component parts: Responses of single units in the temporal cortex of the macaque. *Cereb Cortex.* 4, 509-522.
- Wagner, A.R., Mazur, J.E., Donegan, N.H., Pfautz, P.L., 1980. Evaluation of blocking and conditioned inhibition to a CS signaling a decrease in US intensity. *J Exp Psychol Anim Behav Process.* 6, 376-385.
- Weiss, S.J., Thomas, D.A., Weissman, R.D., 1996. Combining operant-baseline-derived conditioned exciters and inhibitors from the same and different incentive classes: An investigation of appetitive-aversive interactions. *Q J Exp Psychol B.* 49, 357-381.
- Wickens, J.R., Begg, A.J., Arbuthnott, G.W., 1996. Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neurosci.* 70, 1-5.
- Wilson, F.A., Rolls, E.T., 1990a. Neuronal responses related to reinforcement in the primate basal forebrain. *Brain Res.* 509, 213-231.
- Wilson, F.A., Rolls, E.T., 1990b. Neuronal responses related to the novelty and familiarity of visual stimuli in the substantia innominata, diagonal band of Broca and periventricular region of the primate basal forebrain. *Exp Brain Res.* 80, 104-120.
- Wilson, F.A., Rolls, E.T., 1993. The effects of stimulus novelty and familiarity on neuronal activity in the amygdala of monkeys performing recognition memory tasks. *Exp Brain Res.* 93, 367-382.
- Xiang, J.Z., Brown, M.W., 1999. Differential neuronal responsiveness in primate perirhinal cortex and hippocampal formation during performance of a conditional visual discrimination task. *Eur J Neurosci.* 11, 3715-3724.
- Yan, J., Scott, T.R., 1996. The effect of satiety on responses of gustatory neurons in the amygdala of alert cynomolgus macaques. *Brain Res.* 740, 193-200.

- Yonemori, M., Nishijo, H., Uwano, T., Tamura, R., Furuta, I., Kawasaki, M., Takashima, Y., Ono, T., 2000. Orbital cortex neuronal responses during an odor-based conditioned associative task in rats. *Neuroscience*. 95, 691-703.
- Zahm, D.S., 1999. Functional-anatomical implications of the nucleus accumbens core and shell subterritories. *Ann N Y Acad Sci*. 877, 113-128.
- Zahm, D.S., 2000. An integrative neuroanatomical perspective on some subcortical substrates of adaptive responding with emphasis on the nucleus accumbens. *Neurosci Biobehav Rev*. 24, 85-105.
- Zhang, S., Blache, D., Vercoe, P.E., Adam, C.L., Blackberry, M.A., Findlay, P.A., Eidne, K.A., Martin, G.B., 2005. Expression of orexin receptors in the brain and peripheral tissues of the male sheep. *Regul Pept*. 124, 81-87.
- Zhang, W., Sakurai, T., Fukuda, Y., Kuwaki, T., 2006. Orexin neuron-mediated skeletal muscle vasodilation and shift of baroreflex during defense response in mice. *Am J Physiol Regul Integr Comp Physiol*. 290, R1654-1663.

**Acknowledgements**

M.D. was supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409), the National Institutes of Health (NIH R29-DC02952), the National Science Foundation (NSF IIS-97-20333 and NSF SBE-0354378), and the Office of Naval Research (ONR N00014-01-1-0624). D.B. was supported in part by the National Institutes of Health (NIH R01-DC007683) and the National Science Foundation (NSF SBE-0354378). S.G. was supported in part by the National Science Foundation (NSF SBE-0354378) and the Office of Naval Research (ONR N00014-01-1-0624).