

**SOVEREIGN:  
An Autonomous Neural System for Incrementally Learning  
Planned Action Sequences to Navigate Towards a Rewarded Goal**

William Gnatd<sup>1</sup> and Stephen Grossberg<sup>2</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  
Phone: 617-353-7858  
Fax: 617-353-7755

Technical Report CAS/CNS TR-07-015

Submitted: February, 2007  
Revised: September, 2007

Correspondence should be addressed to:  
Stephen Grossberg  
Department of Cognitive and Neural Systems  
Boston University, 677 Beacon Street  
Boston, MA 02215  
Phone: 617-353-7858/7  
Fax: 617-353-7755  
Email: [steve@bu.edu](mailto:steve@bu.edu)

<sup>1</sup>Supported in part by Riverside Research Institute, the Defense Advanced Research Projects Agency (N00014-92-J-4015), the Air Force Office of Scientific Research (F49620-92-J-0225), the National Science Foundation (IRI 90-24877 and SBE-0354378), the Office of Naval Research (N00014-91-J-4100, N00014-92-J-1309, and N00014-95-1-0657), and Pacific Sierra Research (PSR 91-6075-2).

<sup>2</sup>Supported in part by the National Science Foundation (SBE-0354378), and the Office of Naval Research (N00014-01-1-0624).

## ABSTRACT

How do reactive and planned behaviors interact in real time? How are sequences of such behaviors released at appropriate times during autonomous navigation to realize valued goals? Controllers for both animals and mobile robots, or animats, need reactive mechanisms for exploration, and learned plans to reach goal objects once an environment becomes familiar. The SOVEREIGN (Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation) animat model embodies these capabilities, and is tested in a 3D virtual reality environment. SOVEREIGN includes several interacting subsystems which model complementary properties of cortical *What* and *Where* processing streams and which clarify similarities between mechanisms for navigation and arm movement control. As the animat explores an environment, visual inputs are processed by networks that are sensitive to visual form and motion in the *What* and *Where* streams, respectively. Position-invariant and size-invariant recognition categories are learned by real-time incremental learning in the *What* stream. Estimates of target position relative to the animat are computed in the *Where* stream, and can activate approach movements toward the target. Motion cues from animat locomotion can elicit head-orienting movements to bring a new target into view. Approach and orienting movements are alternately performed during animat navigation. Cumulative estimates of each movement are derived from interacting proprioceptive and visual cues. Movement sequences are stored within a motor working memory. Sequences of visual categories are stored in a sensory working memory. These working memories trigger learning of sensory and motor sequence categories, or plans, which together control planned movements. Predictively effective chunk combinations are selectively enhanced via reinforcement learning when the animat is rewarded. Selected planning chunks effect a gradual transition from variable reactive exploratory movements to efficient goal-oriented planned movement sequences. Volitional signals gate interactions between model subsystems and the release of overt behaviors. The model can control different motor sequences under different motivational states and learns more efficient sequences to rewarded goals as exploration proceeds.

## 1. Introduction

**1.1. Three Basic Design Themes.** This article describes the SOVEREIGN (Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation) neural model to clarify how an animal, or animat, can learn to reach valued goal objects through planned sequences of navigational movements. The SOVEREIGN model embodies a self-organizing control system that attempts to learn and perform such behaviors autonomously. As the name SOVEREIGN indicates, this control system unifies visual, recognition, cognitive, cognitive-emotional, and motor competences. We believe that this is the first neural model that embodies and coordinates such a wide range of behavioral competences in an autonomous self-organizing control system that can operate in real time. These results have been briefly reported in Gnatd and Grossberg (2005a, 2005b, 2006).

SOVEREIGN contributes to three large themes about how the brain works. The first theme concerns how brains learn to balance between reactive and planned behaviors. During initial exploration of a novel environment, many reactive movements occur in response to unexpected and unfamiliar environmental cues (Leonard and McNaughton, 1990). These movements may initially appear to be locally random, as an animal orients toward and approaches many local stimuli. As such an animal becomes familiar with its surroundings, it learns to discriminate between objects likely to yield a reward and those that lead to punishment. Such approach-avoidance behavior is often learned via a perception-cognition-emotion-action cycle in which an action and its consequences elicit sensory cues that are associated with them. Rewards and punishments affect the likelihood that the same actions will be repeated in the future. When objects are out of direct sensory range, multiple reactive exploratory movements may be needed to reach them. Eventually, reactive exploratory behaviors are replaced by more efficient planned sequential trajectories within a familiar environment. One of the main goals of SOVEREIGN is to explain how erratic reactive exploratory behaviors can give rise to organized planned behaviors, and how both reactive and planned behaviors may remain balanced so that planned behaviors can be carried out where appropriate, without losing the ability to respond quickly to novel reactive challenges.

The second design theme illustrates the hypothesis that advanced brains are organized into parallel processing streams with complementary properties (Grossberg, 2000a). Each stream's properties are related to those of a complementary stream much as a lock fits its key, or two pieces of a puzzle fit together. The mechanisms that enable each stream to compute one set of properties prevent it from computing a complementary set of properties. As a result, each of these streams exhibits complementary strengths and weaknesses. How, then, do these complementary properties get synthesized into a consistent behavioral experience? It is proposed how interactions between these processing streams overcome their complementary deficiencies and generate behavioral properties that realize the unity of conscious experiences. In this sense, pairs of complementary streams are the functional units because only through their interactions can key behavioral properties be competently computed. SOVEREIGN clarifies how these complementary properties interact together to control goal-orienting sequences of navigational behaviors. For example, it is well-known that there are What and Where (or Where/How) cortical processing streams (Goodale and Milner, 1992; Mishkin, Ungerleider and Macko, 1983; Ungerleider and Mishkin, 1982). In particular, key properties of the What and Where cortical processing streams seem to be complementary. Neural models have clarified that properties of invariant object learning and recognition in the What cortical stream are complementary to properties of goal-oriented movement control in the Where cortical processing stream.

A third design theme underlying the SOVEREIGN model is that brains use homologous circuits to compute navigational and hand/arm movements. In other words, movements of the body and of the hand/arms are controlled by circuits that share many properties. This proposed homology clarifies how navigational and arm movements can be coordinated when a body moves with respect to a goal object with the intention of grasping or otherwise manipulating it using the hand/arm system.

A considerable body of neural modeling of arm movement trajectory control (e.g., the VITE model: Bullock and Grossberg, 1988; Cisek, Bullock, and Grossberg, 1998) suggests that cortical arm movement control circuits compute a representation of where the arm wants to move (i.e., a target position) and compare this with an outflow representation of where the arm is now (i.e., the present position) by computing a difference vector between target position and present position representations. The difference vector represents the direction and distance that the arm needs to move to realize its goal position. Basal ganglia volitional signals of various kinds, such as a GO signal, translate the difference vector into a motor trajectory of variable speed. Additional cortical, spinal, and cerebellar circuitry is needed to ensure that the brain generates the forces that are needed to actually carry out such a commanded trajectory (e.g., the FLETE model: Bullock and Grossberg, 1991; Contreras-Vidal, Grossberg, and Bullock, 1997).

A key difference between navigation and hand/arm movement control concerns how present position is calculated. Because the arm is attached to the body, present position of the arm can be directly computed using outflow, or corollary discharge, movement commands that explicitly code the commanded arm position. In contrast, when a body moves with respect to the world, no such immediately available present position command is available. This difference requires more elaborate brain machinery to compute present position of the body in the world during navigational movements. The brain needs to use a variety of sensory cues, both proprioceptive and visual, to create a representation of present position that can be compared with representations of target position, so that a difference vector and volitional commands can move the body towards desired goal objects. In summary, both navigational movement in the world and movement of limbs with respect to the body use a difference vector computational strategy.

**1.2. What SOVEREIGN Does.** SOVEREIGN's perceptual competences include on-line, albeit simplified, visual representations of a 3D virtual reality environment in which the model controls navigation. SOVEREIGN computes, in parallel, both visual form and motion information about the world. As in the brain, the visual form of objects is computed within the What cortical processing stream, whereas visual motion is computed within the Where cortical processing stream. In this way, the brain can process both what objects are and where and how to track and act upon them.

SOVEREIGN uses the visual form information to incrementally learn spatially-invariant and size-invariant object recognition categories whereby to recognize visually perceived objects in the world. These recognition categories, in turn, learn to read out top-down attentive expectations of the visual objects that they code. Object categories in the What stream are spatially-invariant and size-invariant to prevent a combinatorial explosion from occurring in which each position and size of an object would need its own representation. The Where stream represents the spatial locations of these objects. In particular, visual motion information is used to guide reactive orienting movements and attention shifts to locations at which changes occur in SOVEREIGN's visual world. What-Where inter-stream interactions are needed to enable both

recognition and acquisition of desired goal objects. These parallel streams help SOVEREIGN to balance between reactive and planned behaviors, in a manner that is further discussed below.

SOVEREIGN also includes cognitive processes, notably mechanisms to temporarily store sequences of events in working memory, and to learn sequential plans, or chunks, of these sequences with which to predict and control future planned behaviors. Parallel object and spatial working memories and sequential chunking networks are modeled. The object working memory and chunking network are in the model's What stream, and the spatial working memory and chunking network are in its Where stream. SOVEREIGN clarifies how these parallel cognitive processes cooperate to acquire desired goal objects that can only be reached through a sequence of actions, and to disambiguate sequential navigational decisions in contexts where only one of them would be insufficient.

Cognitive-emotional mechanisms include the role of rewards and punishments in shaping goal-oriented behaviors. In particular, reinforcement learning can influence which learned cognitive chunks will be attended and selected to elicit behaviors that acquire desired goals within a familiar environment. Learned interactions between cognitive and emotional representations, notably motivationally-mediated attention, play an important role in this context-sensitive selection process.

The SOVEREIGN model thus contributes solutions to three key problems: (1) How an animal, or animat that embodies biologically-inspired designs, learns to balance between reactive and planned behaviors in a task-appropriate way. (2) How plans are learned during erratic reactive behaviors in such a way that, after learning, they can be read out fluently at the correct times and in the correct spatial contexts. (3) How, in particular, an animat coordinates its reactive and planned behaviors so that its perception-cognition-emotion-action cycles of exploration, real-time vision, learned recognition, sequential working memory storage, learning of sequential plans, reinforcement learning, and planned action sequences are activated when needed as the animat navigates novel and familiar environments.

A number of problems need to be solved in realize these competences:

1. How are visual inputs used for both target identification and target position estimation?
2. What coordinate system is used for representing target positions? How is it calibrated through learning?
3. What coordinate system is used for representing body position in space? How is it calibrated through learning?
4. What is the role of proprioceptive feedback in calibrating body position?
5. How are target position and body position information combined to control the next planned movement? Are movement sequences whose units are sensory and motor items (cue-command pairs) sufficient for some types of spatial navigation?
6. How are items within movement sequences stored in short-term working memory?
7. How can a memory system be designed to store and recall sequential plans, or chunks? How can such learning solve the classical goal paradox (Grossberg, 1978), which notes that, on all sequence learning trials, the goal always occurs last, whereas during planned sequential performance, knowledge that the goal is being sought is activated first? That is, how can learning be order-preserving, yet also break the learned order to permit the desired goal to activate a plan capable of realizing it?
8. How does reinforcement enable an animal to selectively attend to rewarded goal objects?

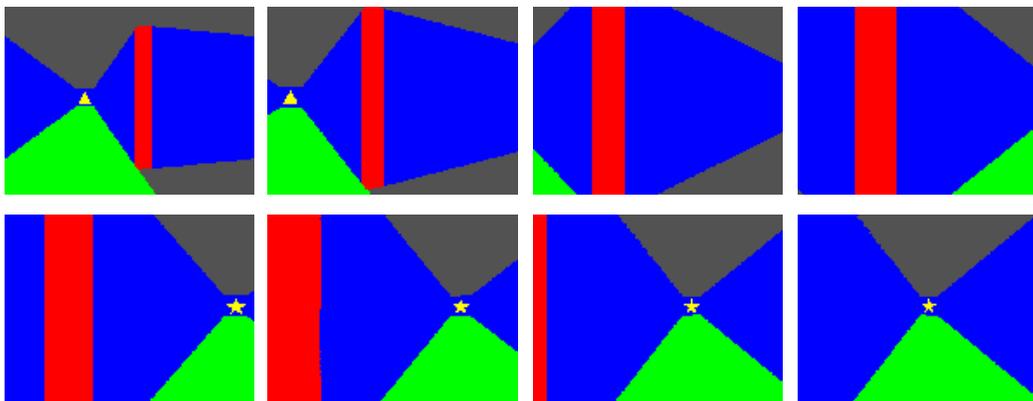
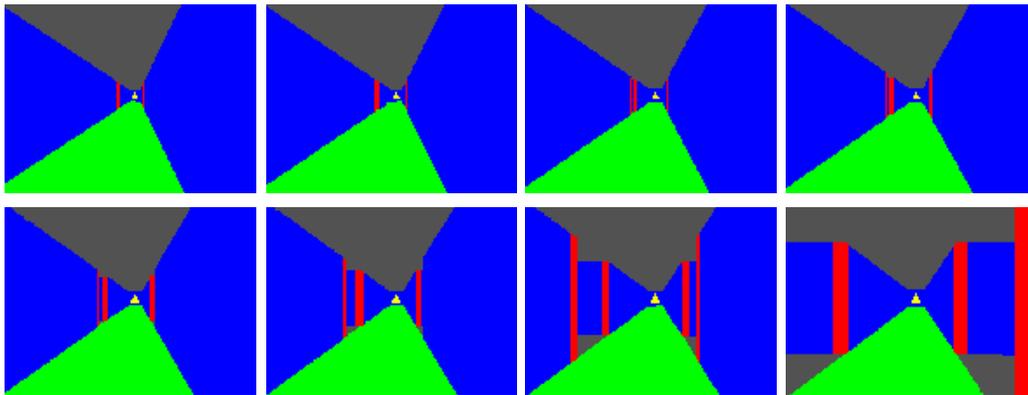
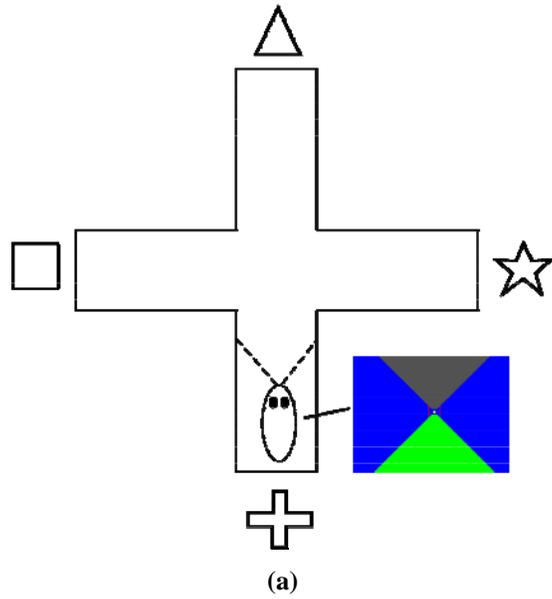
9. How does reinforcement interact with the memory representation of experienced sequences in order to preferentially choose and perform them?
10. How can latent learning, or learning without reward, be explained?
11. How does competitive decision-making resolve conflicts between visually reactive and planned movement commands?

## **2. SOVEREIGN Model**

### **2.1. Approach-Orienting Navigation and Learning in a 3D Virtual Reality Environment.**

The SOVEREIGN model simulates these processes for an animat that experiences a 3D visual world in a virtual reality environment. This world is the relatively simple spatial format of a plus-maze (Munn, 1950). Although simple, this environment tests in a clear way many of the critical competences that the animat needs to achieve. Much of the problem's difficulty arises because an animat may navigate the maze in different ways, including different speeds and directions of movement, on successive learning trials. Despite this variability of each experience of the maze, the animal can learn to navigate the maze to achieve valued goals in an efficient way. For our purposes, it is sufficient to assume that a learning trial starts after placing the animat in the maze, at the end of one arm. The goal location, in one of the other three arms, is baited with a cue that the animal finds rewarding. By shrouding the top of the maze, only route-based visual and motor cues can be used for navigation (O'Keefe and Nadel, 1978). Thus the model does not attempt to explain how spatial navigation, as supported by hippocampal place cells, is achieved. Only one visual cue is assumed to be visible at a time, at the end of each maze arm, from any location within the maze. A schematic diagram of the experimental setup appears in Figure 1a.

A sequence of images from the 3D virtual reality simulation during reactive approach toward a visual cue appears in Figure 1b. As the animat approaches the choice point, a competitive struggle occurs between the salience of form and motion signals. Suppose that the form signals have led to previous object category learning and have led to positive reinforcement that increases their motivational salience. Such motivational salience enhances the strength of the form representation through attentional feedback. Then the form signals may more effectively compete with the motion signals to determine the animat's momentary behavior. If the form cues win the competition, then the animat can continue to carry out an approach movement that is consistent with its recognition. If the motion signals win the competition, then they may trigger a reactive head-orienting movement to the right or left at a choice point, revealing another source of form signals at the end of an adjacent corridor. The outcome of this form-motion competition is sensitive to navigational variations that change from trial to trial. The sequence of visual scenes that are processed during a typical head-orienting behavior is illustrated in Figure 1c. An alternation of approach and orienting movements is characteristic of the animat's exploration of a novel environment.



**Figure 1.** (a) The 3D graphical simulation generates perspective-views from any location within the plus-maze. (b) Snapshots from the 3D virtual reality simulation depict changes in the scene during reactive homing toward the triangle cue. (c)

During reactive approach to the triangle cue, visual motion signals trigger a reactive head orienting movement to bring the star cue into view.

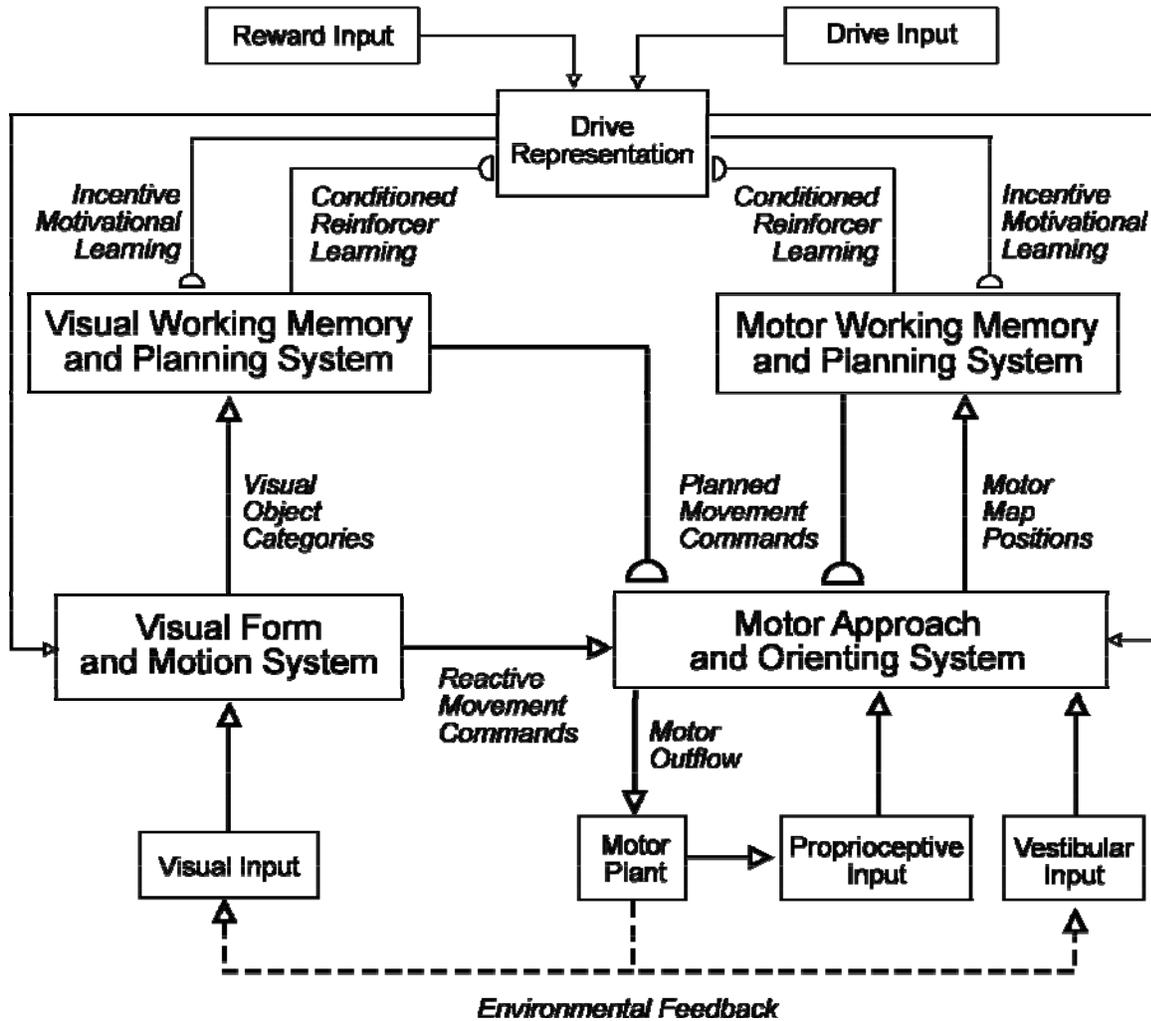
**2.2. Parallel Visual and Motor Working Memory and Chunking Networks.** The animat's control system is split into a number of subsystems shown in the macrocircuit of Figure 2. The primary input is via the visual system. The Visual Form and Motion System processes visual cues within the What and Where cortical streams, respectively. The What cortical stream learns position-invariant and size-invariant object categories via on-line incremental learning. The Where stream computes measures of the relative location of visual cues from the animat. In particular, the distance and direction of the animat from a prescribed visual cue are used to cause approach movements towards that visual cue, or from memory. Motion cues result from the animat's self-motion, and determine whether the animat will make a left or right turn, and how big a turn, instead of continuing to approach a target cue.

These visual form and motion signals compete for control of the animat's approach-orienting behaviors within the Visual Form and Motion System. Learned visual categories can be amplified in strength, and thereby more probably attended, by feedback from motivational centers, called Drive Representations, through learned reinforcement-motivational feedback loops that embody their value as events that predict desired rewards. For this to happen, the invariant object categories are amplified by motivational signals that draw attention to them, and amplify, in turn, the approach commands corresponding to that object's position relative to the animat. Such a motivational amplification requires What-Where inter-stream interactions between position-invariant and position-variant information.

When a motivationally-modulated form cue wins, approach persists. When a motion cue wins, an orienting movement often begins. When motion cues are balanced in strength relative to the present gaze direction, the net left vs. right orienting signal is zero, after opponent competition between the opposite directions takes place. A form cue can then win with high probability. However, a suitably strong left/right motion cue difference can win the form-motion competition and direct the Motor Approach and Orienting System to initiate a head-orienting movement in the favored direction. Target position information for approach behaviors, and motion information for head-orienting behaviors, is relayed from the Visual Form and Motion System to the Motor Approach and Orienting System (Figure 2), where they direct body-approach movements or head-and-body orienting movements.

How does the animat know where a target cue is with respect to its current position? As noted in Figure 2, proprioceptive and vestibular signals provide the ground truth upon which the animat's location is estimated relative to its starting point, and with respect to targets in its environment. Proprioceptive and vestibular cues are capable of guiding animat navigation in a familiar environment even in the dark, and can modify movements quickly to cope with uneven or slippery terrains. Visual cues are also used during navigation to estimate body and head position and displacement relative to the animat. These visual signals are associated with, and adaptively calibrated with respect to, the representations that are activated by proprioceptive and vestibular motor signals. These multiple sources of information work together to more accurately guide movements under varying conditions than any one source of positional signals could.

Estimates of spatial displacement compute the NET body displacement and head rotation from a starting point. Sequences of such approach-orienting displacements represent a path that can command an animat to move from a starting location to a goal location in a maze.



**Figure 2.** The key interactions between components of the SOVEREIGN model are shown in this flow diagram. See text for details.

The Visual Working Memory and Planning System temporarily stores sequences of visual object categories in short-term working memory. It also categorizes, or chunks, sequences of stored object categories. Chunks are learned that are sensitive to object sequences of variable lengths. These visual list chunks can learn to activate motor commands in the Motor Approach and Orienting System via top-down learning. The motor commands encode approach-orient movements. The Visual Working Memory and Planning System operates in parallel with a Motor Working Memory and Planning System that temporarily stores sequences of motor commands in working memory. It also categorizes, or chunks, sequences of stored motor commands. These motor list chunks can also learn to activate approach-orient commands within the Motor Approach and Orienting System.

Why are visual and motor list chunks both needed? Together these parallel visual and motor working memories can disambiguate decisions that only one of them, acting alone, may find ambiguous. For example, the sequences of approach distances and head turns in two

different environments may be the same, but their sequence of visual cues may be different. In a different environment, the sequence of visual cues may be the same, but their sequences of motor actions may differ. The visual and motor working memories induce the learning of list chunks, or sequential planning cells, that are sensitive to their respective object and action sequences, and can read out a prediction of the next motor command. The sequence that can disambiguate two different environments will typically win over one that cannot.

Rewards and punishments can modulate animat learn and determine what visual representations are attended and what motor plans are executed. Upon receiving reward, the active chunks are associated with active drives and actions. Drive inputs represent the animat's internal motivational state, and reward inputs represent valued inputs from the external environment. Both types of inputs combine in Drive Representations, which are most highly activated when a drive input representing a strong internal need combines with either a primary reward or a conditioned reinforcer input from the Visual Form and Motion System (Figure 2). After such a combination of cognitive and emotional learning occurs, when the animat sees a familiar sensory cue under a prescribed motivational state, it can recall a motivationally-compatible plan to reach the site of previous reward. Repeated, random exploration of the environment hereby effects a gradual transition from reactive to more efficient, planned control that leads the animat to its various motivated goals. Due to the selective role of motivational feedback, the animat is capable of learning to carry out different plans to satisfy different motivational goals even in response to the same sensory cues.

Visual and motor list chunks may learn to activate different approach-orient commands under different motivational states. How can a single chunk give rise to multiple responses? How this occurs can be seen by noting that emotional centers are often organized into *opponent* affective processes, such as fear and relief, and that oppositely valenced rewards can be conditioned to these opponent channels (Grossberg, 1984b, 2000b). These opponent-processing emotional circuits are called *gated dipoles*. In such a circuit, habituating transmitters "gate," or multiply, signal processing in each of the channels of the opponent "dipole." The response amplitude and sensitivity to external reinforcing inputs and internal drive inputs of these opponent-processing emotional circuits are calibrated by an arousal level and chemical transmitters that slowly inactivate, or habituate, in an activity-dependent way.

Sensory and cognitive representations, no less than emotional representations, can be organized into opponent channels with habituating ON and OFF cells. Unexpected events can trigger a burst, or sudden increment, of nonspecific arousal. When such an arousal burst is received on top of the baseline tonic arousal input of a normal dipole, it can cause an antagonistic rebound of activity in the OFF channel. In other words, the sensory, cognitive, or emotional hypothesis that is represented in a dipole's activity can be disconfirmed by an unexpected event. An unexpected event can hereby reset ongoing processing and lead to a shift of attention. SOVEREIGN expands the gated dipole mechanism into a gated multipole (see Section 2.6.5), which can select between multiple opponent drive channels. Each channel, whether representing an exploratory or consummatory drive state, can be associated with a particular learned response.

SOVEREIGN embodies a system synthesis and further development of biologically-derived neural networks that have been mathematically and computationally characterized elsewhere. These include LAMINART and FORMOTION models for form and motion processing (Berzhanskaya, Grossberg, and Mingolla, 2007; Cao and Grossberg, 2005; Grossberg, Mingolla, and Viswanathan, 2001; Grossberg and Yazdanbakhsh, 2005; Raizada and Grossberg, 2003), ART fast incremental learning classifiers (Carpenter, et. al., 1992), STORE

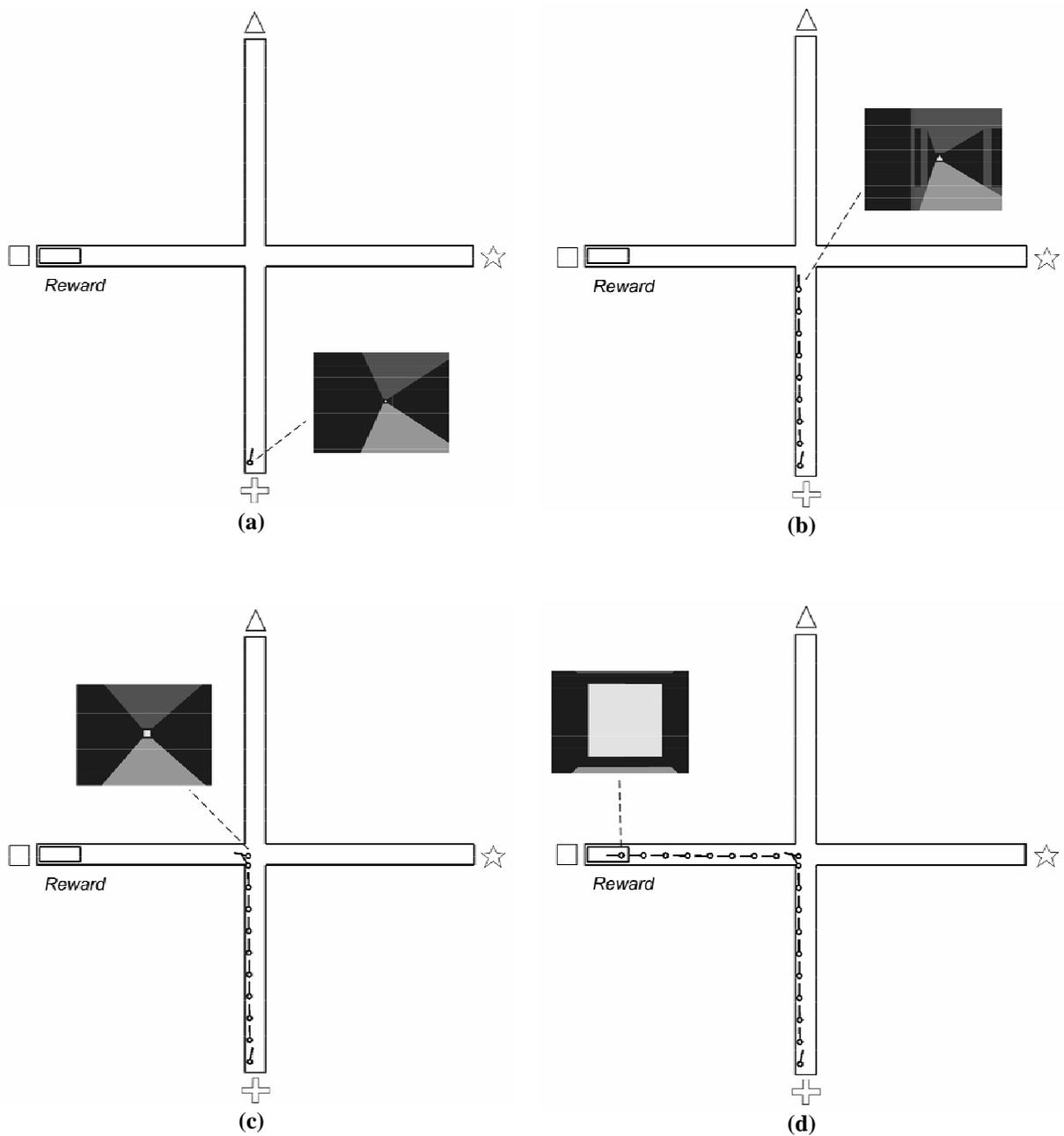
working memories (Bradski, Carpenter, and Grossberg, 1994), Masking Field sequence chunking networks (Cohen and Grossberg, 1986, 1987; Grossberg and Myers, 2000; Grossberg and Pearson, 2007), Gated Dipole opponent processes (Grossberg, 1980, 1984a; Grossberg and Seidman, 2006), CogEM cognitive-emotional circuits for reinforcement learning (Grossberg and Merrill, 1992, 1996; Grossberg, 2000), Spectral Timing circuits for adaptively timed learning (Grossberg and Merrill, 1992; Fiala, Grossberg, and Bullock, 1996), and volitional (GO) and endogenous (ERG, Endogenous Random Generator) gates to release consummatory and exploratory behaviors, respectively (Bullock and Grossberg, 1988; Gaudio and Grossberg, 1991; Pribe, Grossberg, and Cohen, 1997). We are not aware of any other autonomous agent that has yet integrated this range of self-organizing biological competences.

The remainder of this article is organized as follows. This section continues a summary of the model during reactive exploration and then details each of the model systems outlined in Figure 2. Section 3 describes three SOVEREIGN virtual reality simulations and how they illustrate answers to the questions above, while shedding light on relevant experimental data. Section 4 discusses and draws conclusions from these simulations. Finally, a complete mathematical description of SOVEREIGN appears in the Appendix.

**2.3. Reactive Exploration.** The following sequence illustrates the functional flow of the visual input system during reactive exploration in the plus-maze of Figure 1. North designates the vertical direction, with South, East and West following accordingly. For definiteness, assume that the animat is placed into the maze and that all extra-maze cues are suppressed. Furthermore, the animat is motivated under both an exploratory and a hunger drive. The drive and reward inputs to the Drive Representation and then into the Visual and Motor Working Memory and Planning Systems are shown in Figure 2. The exploratory drive is assumed to be excited by an Endogenous Random Generator, or ERG, which is an internal arousal source. Such a source is active when the animat is placed into a new environment. The exploratory drive is inhibited by consummatory drive activity that can support realization of a valued goal. The animat receives a reward (e.g., food) upon reaching the goal location, which is located at the end of the West arm. We show how reactive visual signals during exploration eventually lead the animat toward the goal location, and reinforcement signals strengthen the association between stored plan items and the current motivational state. A step-by-step description of the model under reactive visual guidance follows.

Suppose that, by chance, the animat starts the maze shifted to the left of the corridor, with its head facing slightly to the right of the visual cue (Figure 3a). The left shift reduces the distance to motion cues on the left side of the maze. Because of this positional bias, motion signals within the Visual Form and Motion System (Figure 2) will receive a strong leftward bias. These assumptions are used to demonstrate an exploratory trial which ensures that the animat makes its first head-orienting movement toward the goal location. During the experimental trial, the animat moves forward (Figure 3b), turns left (Figure 3c) and approaches the goal location (Figure 3d) under reactive control.

Movement is organized into orienting and approach movements. In particular, a visually-activated motor command from the Visual Form and Motion System triggers a Motor Outflow command (Figure 2) that specifies a head-orienting angle to align the head with the triangle target. The resulting signals activate the Motor Plant (Figure 2), which converts the movement command into a physical displacement. A head-orienting movement towards the triangle target is thereby initiated. The head turn continues until the NET head-orienting displacement equals the commanded displacement angle.



**Figure 3.** (a) Animat position and head direction facing the triangle cue at the start of the trial. Perspective-views of the 3D virtual reality scene at key locations within the maze are shown by a dashed line. (b) Animat position and head direction while approaching the triangle cue and nearing the choice point. (c) Animat position and head direction after a head orienting turn toward the square cue. (d) Animat position and head direction after reaching the goal location at the square cue.

When the animat faces the triangle cue, a Motor Outflow command from the Visual Form and Motion System activates the Motor Plant (Figure 2) to initiate an approach movement toward the triangle cue. When the Motor Plant converts the commanded approach movement into a physical displacement, the animat's body is passively aligned with the head during an approach movement to maintain a stable posture. Such dynamic stability control is assumed to be present, but is beyond the scope of this work.

During the approach movement, the Motor Approach and Orienting System continues to compute the NET head and body displacement toward the visual target cue. In the absence of competing cues, the body-approach movement could continue until the animat reaches the cue. However, the Visual Form and Motion System processes both form and motion signals while the animat continues to move. A sufficiently strong motion signal in the model's visual periphery can win a competition between Parvo form target locations and Magno motion cues. If a motion cue wins, then it can terminate the approach movement and trigger a reactive head-orienting movement away from the visual target cue.

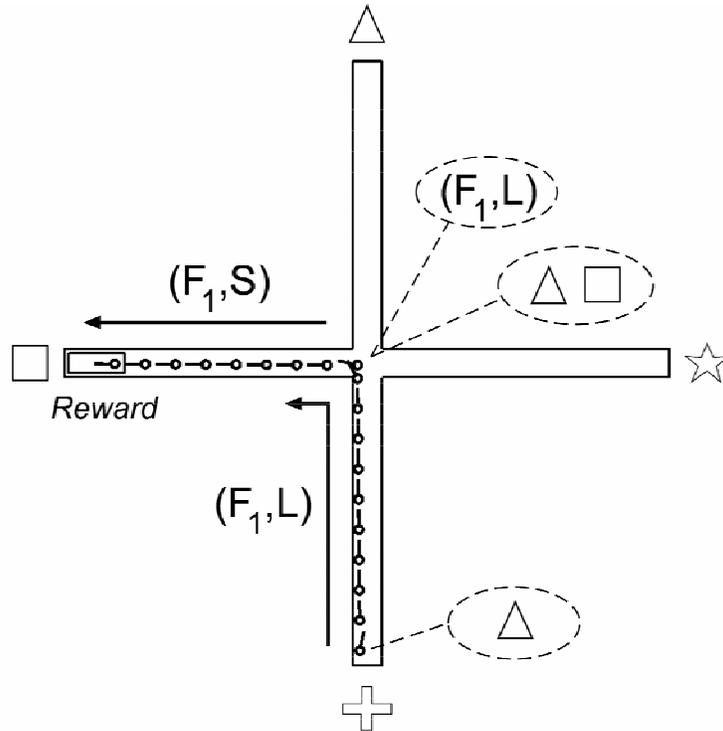
As noted above, when the animat starts in a position that is shifted to the left side of the corridor, as in Figure 3a, motion signals in the left visual hemifield are stronger than those in the right hemifield. Left vs. right motion signals accumulate in the Visual Form and Motion System. When the left motion signal is sufficiently strong relative to the right motion cue and the form signal, a reactive head-orienting command is sent to the Motor Approach and Orienting System.

As the animat carries out these movements, it learns an invariant object category, or chunk, for the triangle visual cue. Top-down signals from the Visual Working Memory and Planning System (Figure 2) corresponding to the Triangle chunk learn the NET body approach and orienting movements computed by the Motor Approach and Orienting System. The triangle cue hereby learns to predict the Forward-Left body movement. The Forward-Left body movements are also stored in the Motor Working Memory and Planning System.

After the animat turns left, invariant preprocessing and learned ART categorization within the Visual Form and Motion System encode a 3D representation of the square cue. This 3D representation is stored in the Visual Working Memory and Planning System (Figure 2), while the NET body displacement in the Motor Approach and Orienting System is reset to prepare for the next movement. Then the cycle of computing the NET head and body displacements begin again, as the animat navigates toward the square cue.

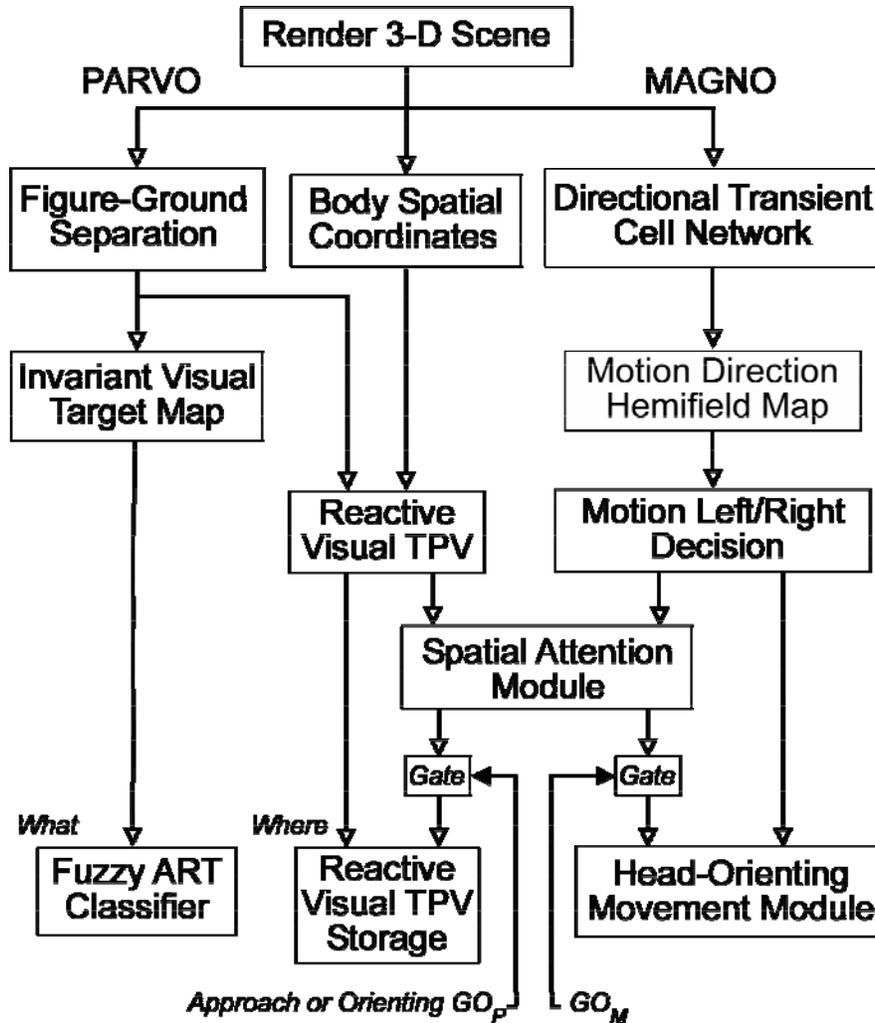
The square cue is at the rewarded location. When the animat reaches this location, it receives a reward, such as food. The active hunger drive representation is then associated with the currently active plan chunks stored in both the Visual Working Memory and Planning System and the Motor Working Memory and Planning System (Figure 2). In particular, the visual Triangle-Square list chunk is learned and associated with the hunger drive representation. In addition, signals from the Triangle-Square chunk learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System, and thereby learns to predict the Forward body movement that brings the animat to the square cue after it turns left in the West arm of the maze. Figure 4 summarizes this sequence of events.

One additional point should be made: All animat behaviors are motivated by some Drive Representation (Figure 2). During initial exploratory activities, an exploratory drive is active. As learning occurs, the exploratory drive is supplanted by the consummatory motivational sources that correspond to the reward; e.g., the hunger drive when the animat is rewarded by food. These processes are now described in greater detail.



**Figure 4.** An initial maze trial in which the animat is under reactive visual guidance is shown in this diagram. An approach movement toward the triangle cue is interrupted by motion signals to the left. After a reactive head orienting movement, the square cue comes into view. After approaching the square cue, the rewarded location is reached and adaptive weights are adjusted to strengthen the association between the forward-left-forward sequence and the current motivational state. The arrows and symbols  $(F_1, L)$  and  $(F_1, S)$ , along with the triangle and triangle-square symbols in the dotted ellipses, summarize that a forward-left movement sequence with a forward distance of  $F_1$  is associated with the Triangle list category, and a forward-straight movement also with a forward distance of  $F_1$  is associated with the Triangle-Square list category.

**2.4. Visual Form and Motion System.** The Visual Form and Motion System processes signals from the What Parvo cortical processing stream and the Where Magno cortical processing stream (Figure 5). This separation of functionality endows the animat with three major capabilities. First, the animat can utilize target object recognition and cognitive-emotional conditioning circuitry to learn, choose, and execute motivationally-compatible movements within an overall plan. Second, the animat can use form information to localize visual references, or beacons, to measure its progress over varying terrain. Finally, the animat can process motion boundaries generated during movement toward a choice point within a maze. As the animat nears a choice point, its field-of-view and the intensity of boundary-derived motion signals increase, which can trigger a reactive head-orienting movement. The visual system also drives several important control signals within the model, as described below.



**Figure 5.** The Visual Form and Motion System flow diagram depicts the stages of visual processing in the model. See text for details.

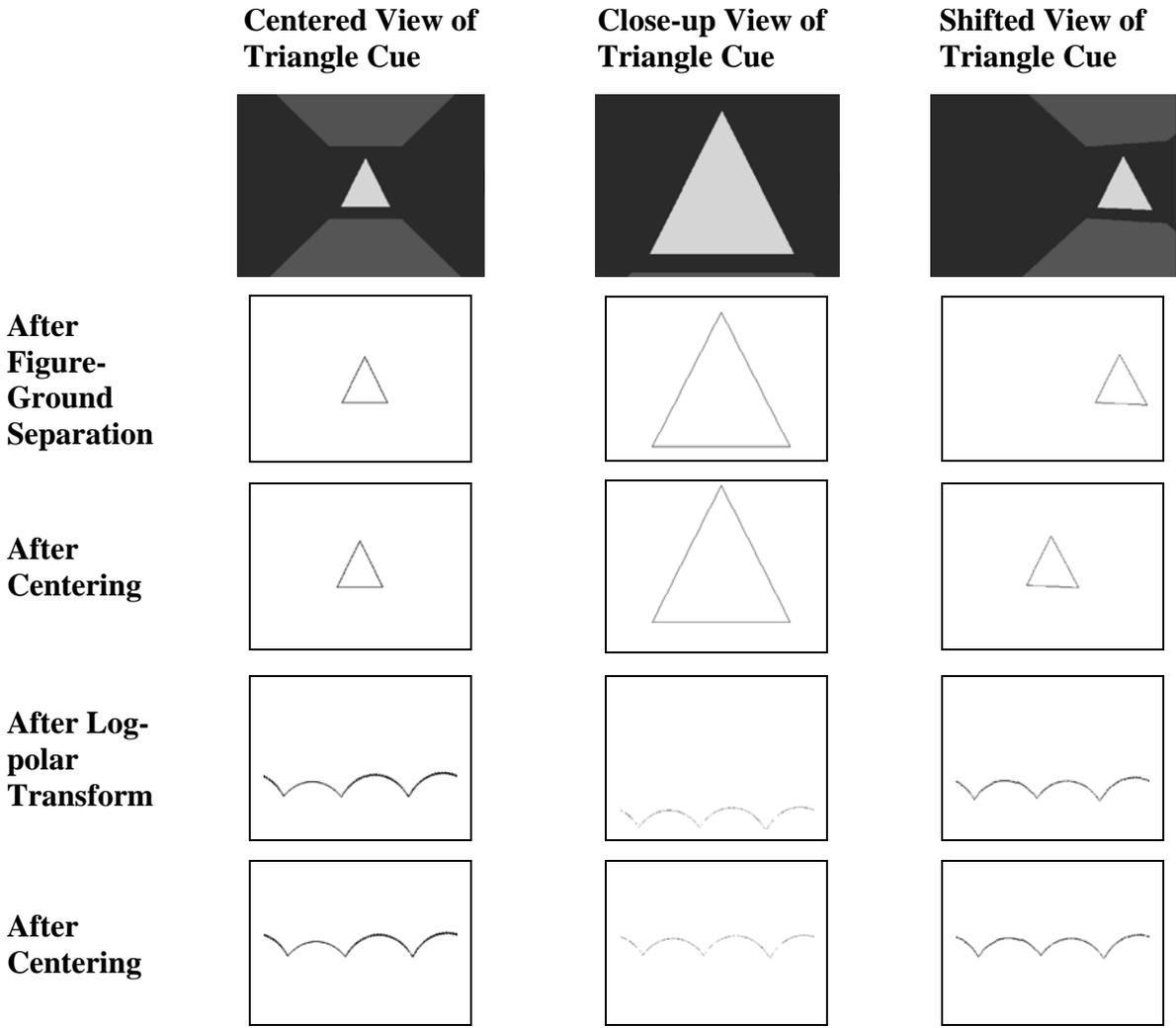
The visual environment is simulated in a virtual reality environment by rendering 3D chromatic scenes as 2D “snapshots” at regular intervals during head-orienting and body-approach movements. For ease of processing, these 2D color snapshots are converted to grayscale using a linear combination of color RGB values (Section A1, equation (5)). As indicated in Figure 1, the visual environment is simplified in SOVEREIGN, which focuses on the various learning and navigational aspects of sequential goal-oriented navigation. A visual target object is separated from the background by a two-stage Figure-Ground Separation module that is within the Parvo stream (Figure 5, left stream). At present, the first processing stage is accomplished in a simple way by using visual targets that are yellow (Figure 1), or have the grayscale corresponding to yellow, and are thereby selected from the background. The second processing stage selects object boundaries via convolution with a 2D Laplacian-of-Gaussian filter (Section A1, equations (6), (7)). Future model developments will include more sophisticated neural models for 3D vision and figure-ground separation (Cao and Grossberg, 2005; Fang and Grossberg, 2007; Grossberg and Yazdanbakhsh, 2005; Kelly and Grossberg, 2000).

When an object falls within the visual field and it is separated from its background, the coordinates of its centroid, in the 2D image plane, are computed (cf., Russell, 1932) and passed to the Reactive Visual TPV module (Figure 5). The Reactive Visual TPV module converts the centroid from image plane coordinates to head-centered spatial coordinates by using the perspective transformation (Section A1, equations (8), (9) and (10)). The Body Spatial Coordinates module computes the angle between the head and body, before combining this information with the target coordinates in the Reactive Visual TPV module to compute the body-centered distance and angle coordinates of the visual target. The Reactive Visual TPV module updates the Reactive Visual TPV Storage module until the volitional Approach and Orienting GO signal ( $GO_p$ ) releases a head-orienting or body-approach movement (Figure 5). The head-orienting movement brings the visual target to the center of gaze. Such a transformation into body-centered coordinates can be learned by using a more elaborate network (Greve et. al., 1993; Grossberg et. al., 1993; Guenther et al., 1993).

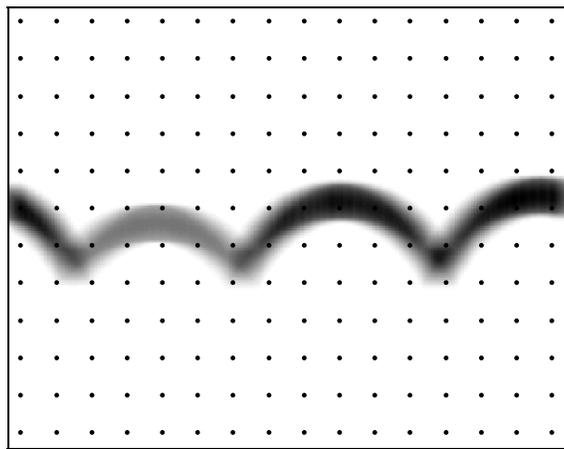
The left path of the Parvo stream in Figure 5 is devoted to learning a size-invariant and position-invariant object category representation of a visual target within the Invariant Visual Target Map. In order to accomplish this, the figure-ground-separated visual target undergoes a log-polar transformation followed by Gaussian coarse-coding (Baloch and Waxman, 1991; Bradski and Grossberg, 1993). The log-polar transform computes a representation of the visual target object that is size-invariant and position-invariant. Examples of invariance processing of the triangle cue from various perspectives are shown in Figure 6a. Further processing by Gaussian coarse-coding reduces storage requirements and 3D foreshortening effects (Figure 6b). This invariant map representation of the target is then transformed into an object category, leading to further compression and invariance under modest changes in object shape, by using unsupervised incremental learning by a Fuzzy ART classifier (Carpenter et al., 1991). The grid of points in Figure 6b show the location of optimal sample points for the vector input to the Fuzzy ART classifier (Section A1; Bradski and Grossberg (1993)). The Fuzzy ART classifier and Reactive Visual TPV Storage modules comprise What and Where cortical representations of visual target objects.

The Fuzzy ART classifier can be generalized in a future version of SOVEREIGN to enable learning of 3D target objects from one of multiple views. This requires additional processing stages to learn individual view categories which can be associatively linked to a view-invariant object category (Baloch and Waxman, 1991; Bradski and Grossberg, 1993; Fazl, Grossberg, and Mingolla, 2007).

The Magno stream generates visual motion signals in response to changes in the visual field during approach movements (Figure 5, right stream). An expanded diagram shows the five stages of motion processing in Figure 7. The Directional Transient Cell Network (Grossberg, Mingolla and Viswanathan, 2001) comprises the first three stages of motion processing. It is designed to sense directional motion signals at a wide range of speeds. Unidirectional Transient Cells compute responses using two consecutive 2D “snapshots”, or doublets, with spatial separation which varies with the approach speed of the animat (Section A2, equations (20), (21)). The Unidirectional Transient Cell responses persist for longer durations at higher approach speeds. Directional Interneurons perform a direction-specific (left/right) time-average of signals from Unidirectional Transient Cells (Section A2, equation (22)). Finally, Directional Transient Cells compute a direction-specific (left/right) response to signals from the Unidirectional Transient Cells and Directional Interneurons (Section A2, equation (23)).



(a)



(b)

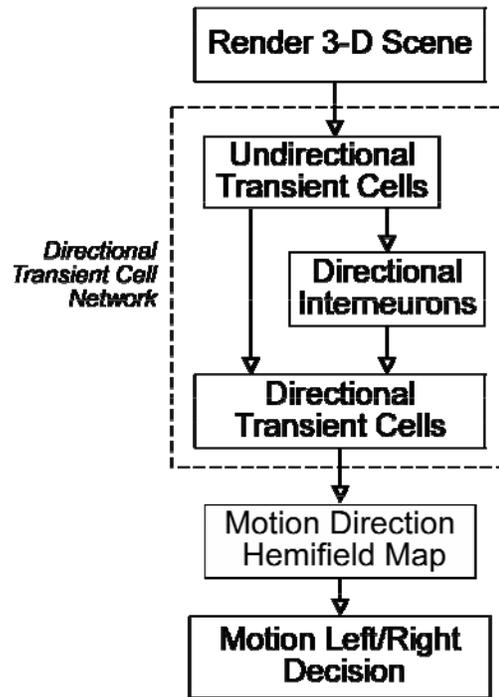
**Figure 6. (a)** Stages of invariant form, or parvo, processing by the Visual Form and Motion System. In the centered (left), close-up (center) or shifted (right) columns, the effect of four stages of invariance processing is a very similar representation of the cue from each perspective. **(b)** Coarse-coding of the centered log-polar representation reduces sensitivity to 3D foreshortening effects and improves compression by reducing Fuzzy ART category requirements. The coarse-coded and centered log-polar map is the *Invariant Visual Target Map*.

Motion Direction Hemifield Cells in the left hemifield are sensitive to the difference between Directional Transient Cell activity in the left and right visual hemifields. Cell responses are proportional to the visual speed of boundaries, such as those formed by vertical maze walls (Figure 1b), in head-centered coordinates. Vertical boundaries help to segment the image and to compute the relative azimuth of regions of interest. Since these boundaries are continuous vertical lines in this simplified maze environment (Figure 1b), only nine horizontal rows of motion cells are utilized, spaced thirty-two pixels apart (Section A2). Using this set of motion cells, the Motion Left/Right Decision module computes directional signals that are sensitive to the total difference between left and right Motion Direction Hemifield Cell activation (Section A2, equations (24), (25)). Such signals are capable of eliciting a reactive left or right head-orienting signal.

The Magno motion signals compete with Parvo object position coordinates and random ERG signals in the Spatial Attention module (Figure 5). Strong motion signals from the Motion Left/Right Decision module can activate the Head-Orienting Movement module and trigger a head-orienting attentional shift. The Head-Orienting Movement module sends a left/right directional burst to downstream motor circuits which turn the head. The head-orienting burst activates an Orienting GO signal ( $GO_M$ ) which initiates the movement and persists until a new cue comes into view. Finally, activation of the  $GO_M$  temporarily halts new head-orienting signals from reaching the Head-Orienting Movement module by closing a gate which inhibits output from the Spatial Attention module. If neither Parvo nor Magno signals are sufficient to win the competition, then an endogenously-generated random signal, or ERG, can elicit head-orienting or body-approach movements to continue exploration.

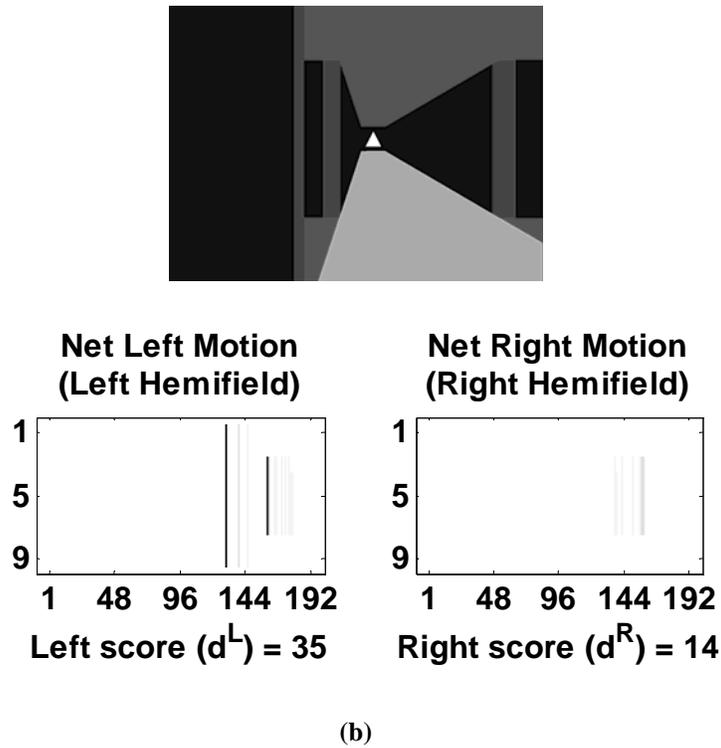
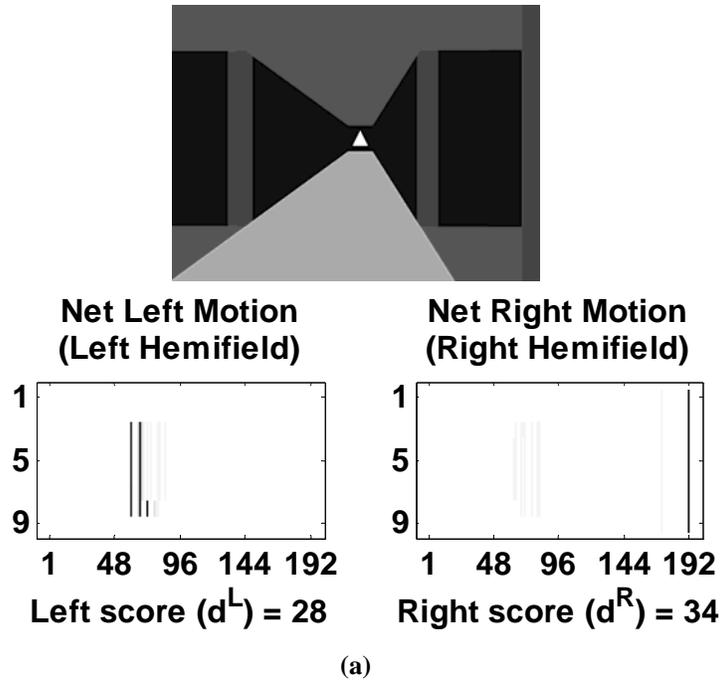
An example of the response of the Motion Left/Right Decision module when motion inputs in the right hemifield are larger than in the left appears in Figure 8a. The top panel of Figure 8a shows the perspective-view of the 3D virtual reality scene as the animat approaches the choice point. The animat is shifted to the right of the corridor, thereby biasing motion signals in the right hemifield. The bottom left panel of Figure 8a shows the left hemifield response of the Motion Direction Hemifield map. This map is sensitive to the net difference between left and right directional transient cells in the left hemifield. Similarly, the bottom right panel of Figure 8a shows the right hemifield response of the Motion Direction Hemifield map. The vertical lines in each map correspond to vertical edges in the 3D perspective-view. Each map has only nine rows since the motion cells are spaced evenly along the vertical direction every thirty-two rows (Section A2). The left and right scores ( $d^L$  and  $d^R$ ) are computed by summing the activity over each map (Section A2, equations (24), (25)). The imbalance between left and right motion scores are processed in the Spatial Attention module (Figure 5). Similarly, an example of Motion Left/Right Decision module response when motion input in the left hemifield is larger than in the right appears in Figure 8b. This situation typically occurs when the animat is moving down the

left side of the corridor. The Motion Left/Right Decision module comprises one input to the Spatial Attention module.



**Figure 7.** Detailed stages of motion, or magno, processing within the Visual Form and Motion System are shown in this diagram. The Directional Transient Cell Network module comprises multiple stages of processing. The Motion Left/Right Decision generates signals which are capable of eliciting a reactive left or right head-orienting signal.

The Spatial Attention module controls either approach or head-orienting movements. If a visual target in the Parvo stream wins, then the Reactive Visual TPV Storage module initiates a reactive approach movement toward the target. If the Magno stream wins, then the Head-Orienting Movement module initiates a head-orienting movement towards the motion cue and away from the visual target. The salience of the target is influenced by rewards that have been correlated with its prior occurrences. In particular, previous rewards lead to greater incentive motivational amplification of the target’s activity. This amplification comes from a Drive Representation module (Figure 2) which enhances the amount of motivated attention that a target receives based on its reinforcement history. In other words, if a visual target has acquired the properties of a conditioned reinforcer, or is an unconditioned stimulus (US), such as a food reward, it can activate a significant level of incentive motivational amplification if the drive that was rewarded is currently not satisfied.

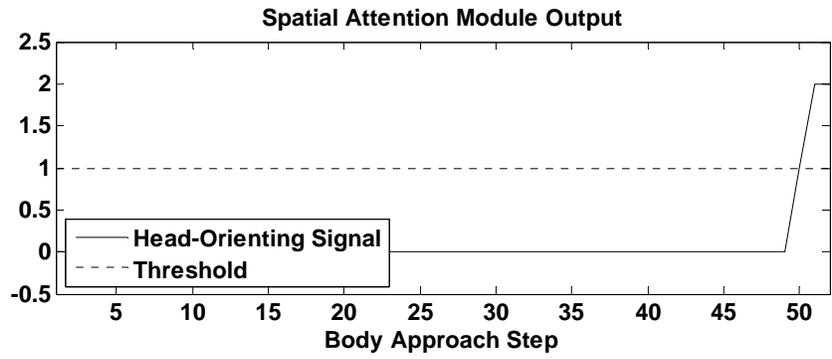
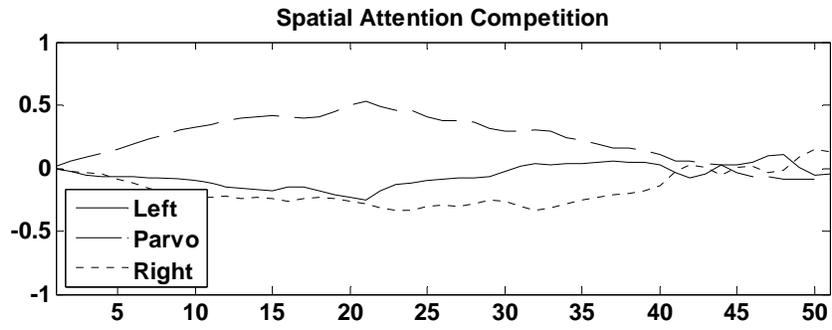


**Figure 8.** (a) An example of the *Motion Left/Right Decision* module response, during an approach movement toward the triangle cue. The animat is shifted to the right of the corridor, causing motion input in the right hemifield to be larger than in the left. At the top of the panel, a perspective-view is shown of the 3D virtual reality scene as the animat approaches the choice point. The lower left panel contains a map of the

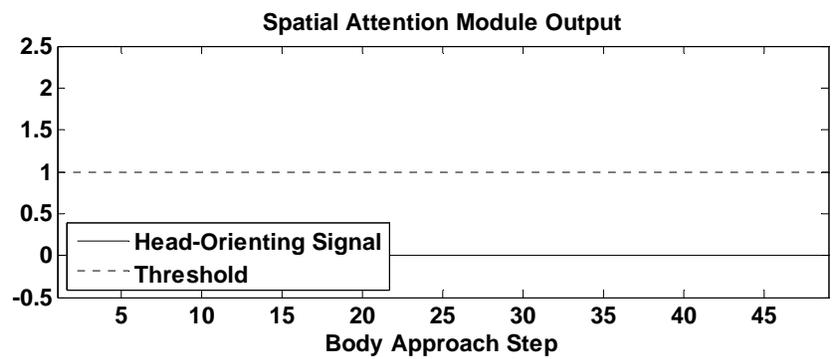
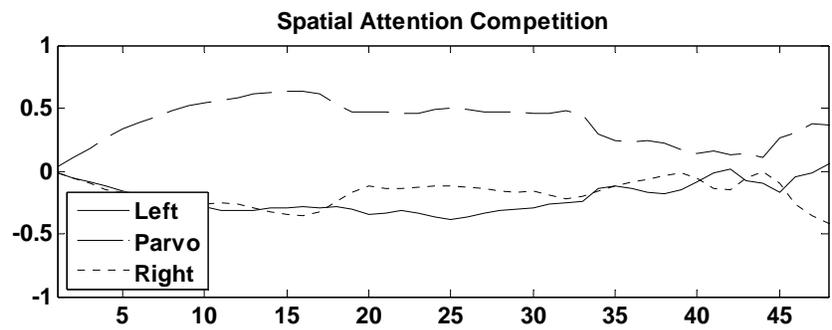
difference between left and right directional transient cells in the left hemifield. The lower right panel contains a map of the difference between right and left directional transient cells in the right hemifield. The left and right scores ( $d^L$  and  $d^R$ ) represent the net left- and right-ward motion signals in the left- and right-hemifields, respectively. The imbalance between left and right motion scores are processed in the Spatial Attention module and may result in head-orienting movement away from the triangle cue. **(b)** An example of the *Motion Left/Right Decision* module response when motion input in the left hemifield is larger than in the right. These panels depict the opposite of the situation outlined above.

The Spatial Attention module (Figure 5) comprises a two-layer feedforward competitive network. By employing winner-take-all dynamics, the Spatial Attention module forces a choice from among the left/right Magno, or Parvo input signals. Parvo signals, which represent the salience of the visual target, are time-averaged in the first stage of processing (Section A3, equation (27)). These signals are multiplicatively scaled by a gain factor which is large if incentive motivational signals are strong enough, and smaller otherwise (Section A3, equation (29)). The amplitude of Parvo input is inversely proportional to the distance to the target object, as computed by the Reactive Visual TPV module (Section A3, equation (30)). Targets that are closer naturally have greater salience and are more likely to yield immediate reward. On the other hand, signals from the Motion Left/Right Decision module reflect the overall level of visual motion in the left or right hemifields. These signals are also time-averaged in the first stage of processing (Section A3, equations (26), (28)). In the second stage of processing, a winner is selected between the Parvo and Magno signals (Section A3, equations (31)-(33)). Decision signals are obtained by thresholding each of the three outputs (Section A3, equations (34)-(36)). The final output is then determined by how these decision variables activate an heuristic motor plant, as detailed in Section A3. If no output is consistently strong enough, then random signals from the ERG can elicit head-orienting or body-approach movements to continue exploration.

The output of the Spatial Attention module (Figure 5) is compared under two different conditions in Figure 9. In the first example, the animat approaches the choice point under reactive control from a starting point that is shifted right in the corridor (Figure 1c). The upper panel of Figure 9a displays the Left, Parvo and Right decision signals during an approach movement toward the choice point. The difference between the Left and Right signals results from the animat starting the trial shifted to the right in the corridor. In the lower panel of Figure 9a, the orienting decision reaches the speed-sensitive threshold (equal to 2 in this case) toward the end of the approach. The baseline Parvo gain (Section A3, equation (29)) allows the Right signal to win the competition, thus triggering the Head-Orienting Movement module (Figure 5) and a head-orienting movement to the right. In Figure 9b, the animat approaches the choice point along the centerline of the corridor. Because of prior experience with the visual target, incentive motivational feedback from the Drive Representation module (Figure 2) increases the Parvo attentional gain (Section A3, equation (29)), and the animat's approach speed. This motivational bias to the Parvo input enables the Parvo signal to inhibit Reactive Head-Orienting Movement signals (Section A3, equations (31)-(33)). At this increased approach speed, the Right motion signals are stronger than those to the Left, but they never exceed the amplitude of the Parvo, or approach, decision signal, so the animat continues toward the visual cue.



(a)



(b)

**Figure 9. (a)** An example of the Spatial Attention module response while the animat approaches the choice point. The top panel displays the time-averaged left, parvo and right decision signals as a function of distance. The difference between left and right response is a result of the animat starting the trial shifted to the right of the corridor. In the bottom panel, the orienting decision reaches the speed-sensitive threshold ( $=2$  in this case) toward the end of the approach. Nominal parvo gain allows the right signal to ultimately win the competition, thus triggering the Head-Orienting Movement module to command a head orienting movement to the right. **(b)** The animat is approaching the choice point while shifted to the left of the corridor. Motivational feedback has increased both the running speed and parvo attentional gain and simultaneously reduced the Head-Orienting Movement module threshold. While the left motion signals are stronger than the right, yet never exceed the amplitude of the parvo, or approach, decision signal. Thus, the animat continues toward the visual target.

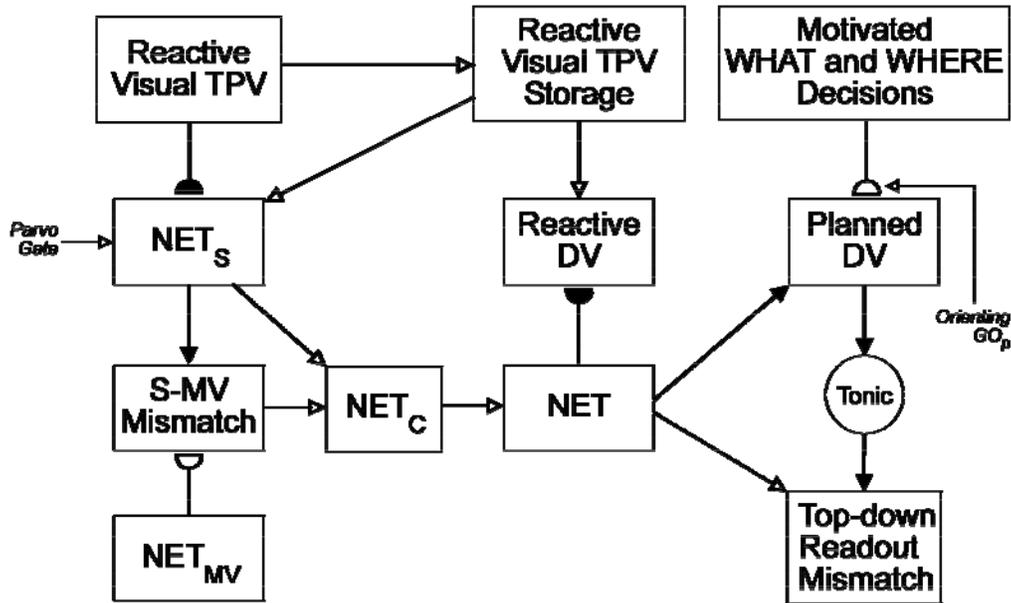
In summary, the sequence of processing in the Visual Form and Motion System when navigation begins (Figure 3a) is as follows: Form signals in the visual field are processed by the Figure-Ground Separation module (Figure 5). The Reactive Visual TPV module computes the head-centered position of the target and converts its head-centered position to body-centered coordinates. While attending to the target, the animat learns to categorize the target using the Fuzzy ART Classifier module (Figure 5). The animat updates the activation of the Spatial Attention module, which may receive reinforcement signals via the Sensory-Drive Hierarchy module (see Figure 17 below). Competition between Magno and Parvo signals in the Spatial Attention module (Figure 5) chooses a winning attentional focus. The winning visual signal opens a movement gate that enables either the Reactive Visual TPV Storage module to command a head-orienting or body-approach movement toward the target, or the Head-Orienting Movement module to command a head-orienting movement to orient away from it.

Head-orienting movements can more easily be elicited as the animat approaches the choice point in the maze because the vertically-oriented corners of the corridor walls, shown as a sequence of images in Figure 1b, move outward on the retina at an increasing rate as the animat nears the choice point. As these signals grow, they may become large enough to win the competition within the Spatial Attention module. Such a head turn is assumed to continue as long as the motion system is active. For example, as the animat turns its head at the choice point, motion signals will arise from new visual information appearing in the periphery of the visual field (Figure 1c). The head movement tracks this motion cue until the head is facing down the corridor and the motion system becomes quiet. At that time, the Parvo visual system can again win the Spatial Attention competition and select a visual target to approach. Model body-approach and head-turning movements tend to alternate, thereby accumulating sequences of approach-orient movement commands.

The Fuzzy ART Classifier (Figure 5) undergoes a match-mismatch-reset cycle as the animat repeatedly categorizes and turns away from potential targets (Figure 1c). A mismatch results when fewer features (i.e., a different view) of an object fall within the visual field. This is followed by a reset, as the Fuzzy ART module selects another category. A match may then be computed on the remaining features of the visual target object. To compute its match-mismatch-reset cycle, Fuzzy ART uses a combination of ON-cell and oOFF-cell activities that is called *complement coding* (Section A1.2, equation (11)). Ultimately, no features of the visual object are

visible. Soon thereafter, a match is computed as features of a new target object come into view. The match-mismatch cycle continues until the head-orienting command is satisfied and the turn comes to an end.

**2.5 Motor Approach and Orienting System.** As noted above, the Motor Approach and Orienting System directs body-approach and head-orienting movements (Figure 2). Cumulative estimates of each approach-orienting movement that is processed within the Motor Approach and Orienting System are stored in the Motor Working Memory and Planning System (Figure 2), which is discussed in Section 2.6. This section summarizes how these estimates are computed.



**Figure 10.** The Motor Approach and Orienting System flow diagram depicts the control hierarchy which generates motor outflow commands. See text for details.

The Motor Approach and Orienting System flow diagram is shown in Figure 10. Target position information originates from one of two sources. First, it can be received from the body-centered distance and angle coordinates of the visual target object computed in the Reactive Visual TPV module (Figure 5). Second, it can be received from learned top-down signals from the processing stage that computes Motivated WHAT and WHERE Decisions (Figure 10). These decisions comprise responses from the animat's learned experience which are compatible with the current motivational state. An approximate measure of head-orienting and body displacement is computed by the NET module (Figure 10 and Section A10, equations (120), (121)). Target position information flows from the Reactive Visual TPV to the Reactive Visual TPV Storage module (Section A10, equation (116)). The NET activity is subtracted from the Stored TPV via learned weights to compute the Reactive Difference Vector, or DV (Section A9, equation (89)), which represents the angle and distance to move. The learned weights from the NET activity are necessary to calibrate the DV activity. Similarly, learned top-down commands from the Motivated WHAT and WHERE Decisions activate the Planned DV (Section A8, equations (78), (79)), where NET movement signals are subtracted, yielding a planned angle and distance to

move. Calibration of planned commands is accomplished entirely by the top-down adaptive weights. The Reactive DV and Planned DV are the first motor control stages which can elicit head-orienting and body-approach movements.

The NET estimates of head-orienting and body-approach displacement requires multiple stages of processing to be computed (Figure 10). NET estimates during navigation replace the outflow present position estimates that are computed during hand/arm movements. The NET<sub>S</sub> module (Figure 10) calculates this displacement using target positions computed by the Visual Form and Motion System (Figure 5). Body-centered spatial coordinates are denoted by an “s” subscript. The NET<sub>S</sub> field activity encodes the net body movement toward a target in spatial coordinates by subtracting the Reactive Visual TPV activity from the Reactive Visual TPV Storage module activity (Section A10, equation (117)).

Initially, target position information flows from the Reactive Visual TPV to the Reactive Visual TPV Storage module and a short burst of learning zeros the difference at the NET<sub>S</sub> module (Section A10, equation (118)). As the animat moves toward a target, updates to the Reactive Visual TPV Storage module cease and the Reactive Visual TPV decreases, thereby allowing the NET<sub>S</sub> module activity to grow. Vestibular and proprioceptive feedback signals are integrated into distances by the NET<sub>MV</sub> module (Figure 10 and Section A10, equation (109)).

Learning at the output of the NET<sub>MV</sub> module calibrates the vestibulo-motor signals relative to the visual signals at the S-MV Mismatch module (Section A10, equation (115)). This adaptive process uses a slow learning rate while visual signals are available from the Visual Form and Motion System (Figure 5). The resulting activity at the S-MV Mismatch module (Section A10, equation (110)) serves as a correction factor which can account for the animat’s progress either without visual feedback (e.g., in the dark) or over uneven (e.g., slippery) terrain. When the NET<sub>S</sub> and NET<sub>MV</sub> module activity are identical, then the correction factor is zero and the S-MV Mismatch module activity is also zero.

The NET module (Section A10, equations (120), (121)) combines signals from the NET<sub>S</sub> and S-MV Mismatch modules into a robust sensory-motor representation of body displacement (Section 2.3.1). The NET<sub>S</sub> module is only active when Parvo signals are present in the Invariant Visual Target Map module (Figure 5). Learned weights from the NET module inhibit activity of the Reactive DV (Figure 10). When the animat has reached the target under visual guidance, the Reactive Visual TPV reaches zero and these adaptive weights are updated, thereby inhibiting the Reactive DV, and stopping the movement. On future trials, the Reactive DV module can be driven to zero by a calibrated level of activity in the NET module, regardless of whether visual input is available.

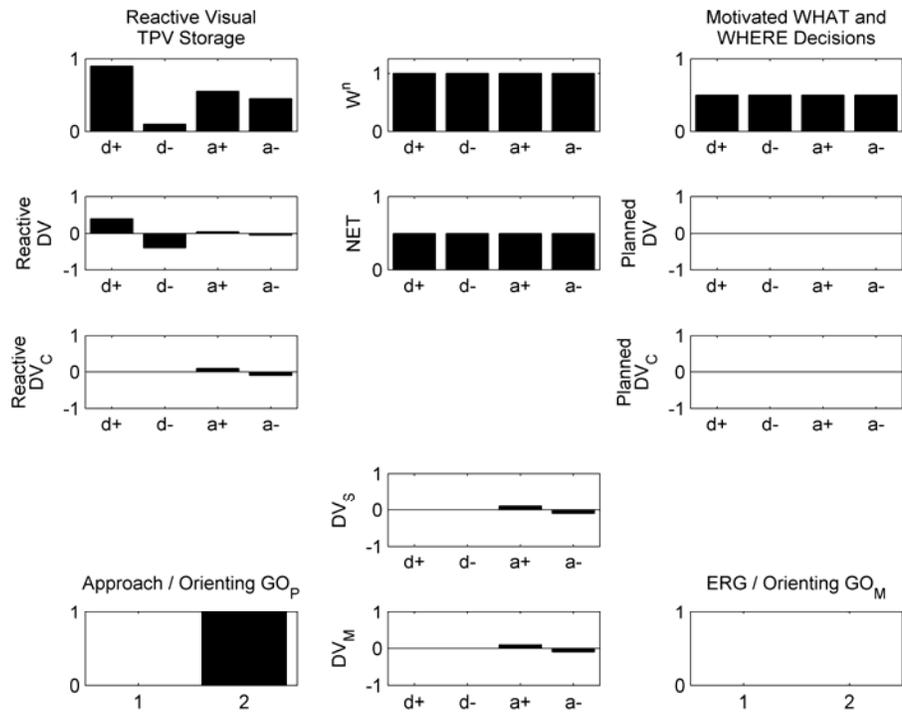
While under reactive control, visual target coordinates flow into the Reactive Visual TPV Storage module (Section A10, equation (116)) and activate the Reactive DV module (Section A9, equation (89)), which initiates head and body movements. The Approach or Orienting GO<sub>P</sub> control signals (Section A9.3, equations (101), (103)), are activated when the Reactive or Planned DV command is released under volitional control. The activation of the Approach or Orienting GO<sub>P</sub> allows the DV signals to initiate a head-orienting or body-approach movement. Updates to the Reactive Visual TPV Storage module (Figure 10) continue until the Approach or Orienting GO<sub>P</sub> is activated. However, under planned control, Motivated WHAT and WHERE Decisions (Figure 10) learn to read out planned head-orienting and body-approach movements. Top-down commands are computed in the Planned DV module, which can initiate head and body movements in response to motivationally-compatible plan items. As then plan unfolds, NET increases until the Planned DV approaches zero, thus terminating a planned movement. The

Top-down Readout Mismatch module compares the activity of the learned top-down command and the NET module. A sufficiently large discrepancy between these fields can elicit a control signal to select a different top-down signal from the Motivated WHAT and WHERE Decisions. For instance, the control signal is released when a planned response is interrupted by a strong motion signal which activates the Head-Orienting Movement module (Figure 5) and the animat turns away from the planned response direction.

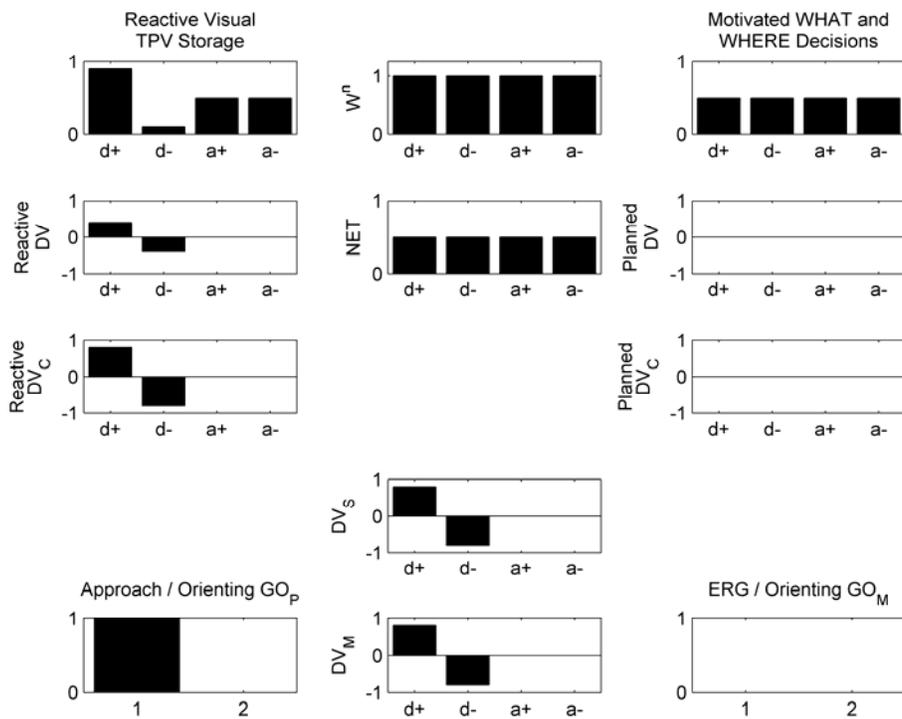
To summarize, there are four sites which utilize adaptive weights in the Motor Approach and Orienting System (Figure 10). First, there is fast learning from the Reactive Visual TPV module to the  $NET_S$  module when a new visual target cue is categorized and loaded. This learning effectively calibrates differences between the Reactive Visual TPV and Reactive Visual TPV Storage to ensure accurate visual distance and angle measurements. Second there is slow, continuous learning from the  $NET_{MV}$  module to the S-MV Mismatch module to calibrate proprioceptive and vestibular signals to be consistent with the visual signals. Third, there is fast learning from the NET module to the Reactive DV module while visual signals are on to calibrate their signals against the angle or distance fields of the Reactive Visual TPV module, thereby enabling reactive movements to stop accurately. After new plan items are stored in the Visual or Motor Working Memory and Planning System, there is top-down learning of the NET movement from the Motivated WHAT and WHERE Decisions (Figure 10) to the Planned DV module. An example follows that describes the flow of spatial and motor signals through the model.

The following sequence illustrates the flow of both spatial and motor information during a maze trial after exploratory learning. This trial follows the same sequence as the initial reactive exploratory trial summarized in Section 2.1, but with an emphasis on motor processing. As in the previous case, the animat by chance starts the maze shifted left in the corridor, with its head turned slightly to the right of the visual cue (Figure 3a). The model initiates processing when a target in the visual field attracts the animat's attention, as follows.

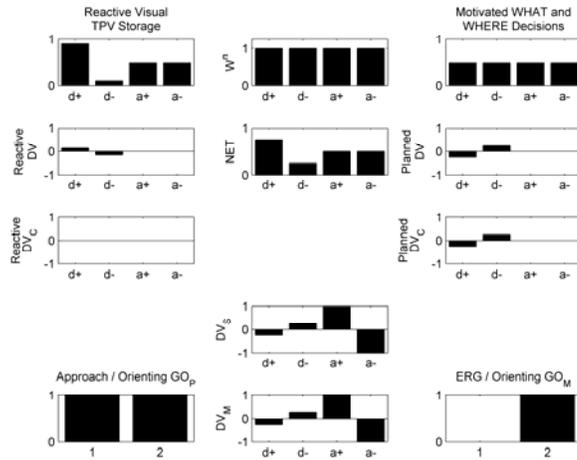
Invariant preprocessing and learned ART categorization within the Visual Form and Motion System (Figure 5) encode a compressed representation of the triangle cue. The NET body displacement, from a prior trial, is reset by zeroing the proprioceptive and vestibular and motor integrator activity (Section A10.1). The head-centered azimuth and distance to the triangle cue are computed and converted to body-centered coordinates by the Reactive Visual TPV module (Figure 5). The animat loads the body-centered Reactive Visual TPV module into the Reactive Visual TPV Storage module. The reactive movement command is then relayed to the Reactive DV module (Figure 10) to initiate reactive approach. A short burst of Vector Associative Map, or VAM, learning (Gaudiano and Grossberg, 1991) is used at the  $NET_S$  field to zero any differences between the coordinates of the Reactive Visual TPV and Reactive Visual TPV Storage fields. The Motor Approach and Orienting System response under these conditions is shown in Figure 11a. In the panels of Figure 11, the cell activity of each field of the Motor Approach and Orienting System, the GO and ERG control signals, and the resulting motor outflow commands,  $DV_S$  and  $DV_M$ , are shown. In Figure 11a, the Reactive and Planned commands have been calibrated with respect to the NET displacement via VAM learning.



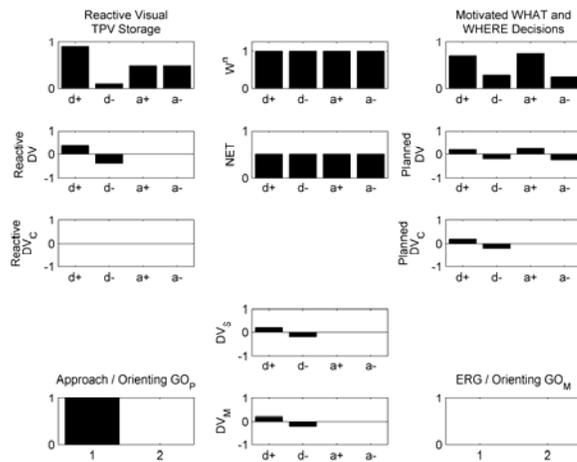
(a)



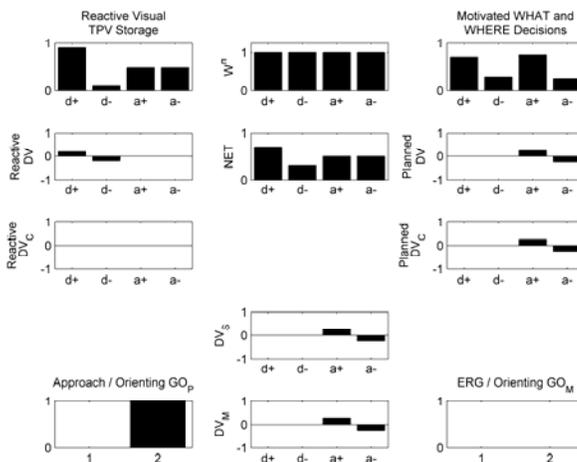
(b)



(c)



(d)



(e)

**Figure 11.** A series of simulations depicting the Motor Approach and Orienting System (Figure 10) response under various conditions ( $d$  = distance,  $a$  = angle). The on- and off-channel activities of each module, denoted by “+” and “-“ symbols, are shown graphically within the following six panels. The activity of ERG and GO control signals appear for reference. **(a)** The animat has identified a visual target object, computed its body-centered coordinates, and loaded the Reactive Visual TPV Storage module. The activity of the Reactive Visual TPV Storage module is shown graphically in the upper left corner. The vertical bars represent the activation levels, which show the target some distance away, and offset to the left. There is no planned movement command from the Motivated WHAT and WHERE Decision module, hence the on- and off-channel responses are equal. The control hierarchy implemented by the Motor Approach and Orienting System causes the reactive orienting command to inhibit the approach command. The resulting movement command ( $DV_S$ ) gates the Orienting  $GO_P$  control signal and enables the outflow command ( $DV_M$ ) to initiate a head turn to face the visual cue. **(b)** The animat has turned to face the visual target object, at which time the Reactive Visual TPV Storage module is updated. The activity of the Reactive Visual TPV Storage module, shown in the upper left corner, indicates that the visual target object is some distance directly ahead of the animat. The resulting movement command ( $DV_S$ ) gates the Approach  $GO_P$  control signal and enables the outflow command ( $DV_M$ ) to initiate a body movement to approach the visual cue. **(c)** During a reactive body approach movement, the Visual Form and Motion System (Figure 5) responds to peripheral motion signals. When such signals are strong enough, the Head-Orienting Movement module commands a head-orienting movement to turn away from the visual cue. A large head turn command ( $DV_S$ ) is gated by the Orienting  $GO_M$  and causes the motor outflow ( $DV_M$ ) to command a head turn away from the visual target object, even while still approaching it. **(d)** The Motor Approach and Orienting System response is shown after top-down readout from the Motivated WHAT and WHERE Decisions overrides the reactive approach command. Then, activation of the Approach  $GO_P$  gates the top-down body approach command ( $DV_S$ ), causing the motor outflow ( $DV_M$ ) to initiate a forward movement under planned control. **(e)** The Motor Approach and Orienting System response is shown after the planned approach movement is complete and a planned orienting command inhibits the reactive approach command. The planned head turn ( $DV_S$ ) gates the Orienting  $GO_P$  and the motor outflow command ( $DV_M$ ) initiates an orienting head turn under planned control.

The Reactive DV module (Figure 10) computes a motor command to initiate a head-orienting movement toward the triangle cue. The motor command drives alternating body orienting and approach movements. The Motor Approach and Orienting System also combines visual, proprioceptive and vestibular inputs to compute a NET head and body displacement toward the visual target cue. The head turn continues until the NET turn signals zero the angle component of the motor command and the animat is facing the triangle cue. The Motor Approach and Orienting System response under these conditions is shown in Figure 11b.

Once the animat is facing the triangle cue, the Reactive DV module (Figure 10) computes a motor command to initiate an approach movement toward the triangle cue. The animat’s body

is passively aligned with the head during an approach movement. The Motor Approach and Orienting System response under these conditions is shown in Figure 11c. When the volitional signal to approach the target is activated, the Visual Working Memory and Planning System, or Visual Memory System (Figure 2), is updated by storing the learned category in the Fuzzy ART Classifier module (Figure 5). After additional processing, described below, a motivationally-compatible plan chunk, called the Motivated WHAT Decision, is chosen. Furthermore, the Drive Representation module can energize nonspecific reinforcement signals via its incentive motivation output. This incentive arousal is added to the Parvo input to the Spatial Attention module to enhance attention to the visual cue (Figure 2). Finally, the selected sensory chunk reads-out a spatial pattern representing a distance and turn angle to the Planned DV module in the Motor Approach and Orienting System (Figure 10). This target movement vector can override the visual representation of body-centered target position in a later stage of processing. The Motor Approach and Orienting System response under these conditions is shown in Figure 11d.

During the approach movement, the Motor Approach and Orienting System (Figure 10) updates the multimodal NET head and body displacement via visual, vestibular and proprioceptive signals. NET inhibits the Reactive DV via calibrated weights. When the activity of the Reactive DV is zero, then the body-approach movement stops because the animat has reached the triangle cue. Simultaneously, the Visual Form and Motion System (Figure 5) processes both form and motion signals which compete via learned interactions. A suitably strong motion signal in the periphery can win the competition and trigger a head-orienting movement away from the visual target cue. If motion signals were able to win the competition in the Spatial Attention module, then the Head-Orienting Movement module (Figure 5) would be activated, as shown in Figure 11e.

Because the animat is shifted to the left side of the corridor, motion signals in the left visual hemifield are stronger than those in the right hemifield (Figure 3b). A left/right orienting signal accumulates in the Visual Form and Motion System (Figure 5). Once the left orienting signal is sufficiently strong relative to the form signal, an attentional shift to the left occurs. However, in this case, incentive motivation signals from the Drive Representation module (Figure 2) increase the gain on Parvo inputs to the Spatial Attention module. Motion signals in the periphery are unable to overcome the attention focus on the currently visible triangle cue. Now that the distance field of the Planned DV module (Figure 10) has zero activity, the angle field of the Planned DV module can initiate the previously learned head turn to the left. The Visual Form and Motion System sends a command to the Motor Approach and Orienting System (Figure 10), which then computes a Motor Outflow command to initiate an orienting head turn to the left. The Motor Approach and Orienting System updates the NET body displacement during the head-orienting turn. The Motor Approach and Orienting System response under these conditions is shown in Figure 11f.

The head-orienting movement continues until the angle field of the NET module inhibits the angle field of the Planned DV module (Figure 10) and the square cue is centered in the visual field. Top-down signals from the Visual Working Memory and Planning System (Figure 17) corresponding to the Triangle chunk learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System (Figure 10). Top-down learning is gated when the Approach or Orienting  $GO_p$  signal is shut off. The representation of the triangle cue is thereby associated with the exploratory drive and learns to predict the Forward-Left body

movement. Finally, the Forward-Left body orienting and approach movement is stored in the Motor Working Memory and Planning System (Figure 2).

Invariant preprocessing and learned ART categorization within the Visual Form and Motion System (Figure 5) encode a compressed representation of the square cue. Then, the NET body displacement in the Sensory-Motor NET Movement module (Figure 10) is reset by zeroing the proprioceptive and vestibular and motor integrator activity. The head-centered azimuth and distance to the square cue are computed and converted to body-centered coordinates. This reactive movement command is relayed to the Motor Approach and Orienting System to begin reactive approach. The animat's body is passively aligned with the head during an approach movement. If a volitional signal to approach the target is activated, then the Visual Memory System (Figure 2) is updated by storing the learned category from the Fuzzy ART Classifier module (Figure 5). After multiple stages of processing, described below, a motivationally-compatible plan chunk is chosen and the Motivated WHAT and WHERE Decisions (Figure 10) are computed. The learned Drive Representation module (Figure 2) response can energize nonspecific reinforcement signals via its incentive motivation output. This incentive arousal is added to the Parvo input to the Spatial Attention module (Figure 5) to enhance attention to the previously-reinforced square cue. Finally, the sensory chunk reads-out a spatial pattern representing a distance and turn angle to the Planned DV module (Figure 10). This target movement vector may then override the visual representation of body-centered target position via a high-gain pathway. During the approach movement, the Motor Approach and Orienting System updates the NET head and body displacement toward the visual target cue. The body-approach movement can continue until the NET approach signals zero the approach component of the Motor Outflow command and the animat reaches the square cue.

When the animat reaches the rewarded location, it receives a food reward. Plans stored in both the Visual and Motor Working Memory and Planning System are associated with the currently active hunger drive. Top-down signals from the Visual Working Memory and Planning System corresponding to the Triangle-Square chunk learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System. The short-term memory representation of both the Triangle and Triangle-Square chunks are associated with the hunger drive. However, only the Triangle-Square chunk learns to predict the Forward-Straight body movement in this context. Top-down signals from the Motor Working Memory and Planning System corresponding to the Forward-Left chunk learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System. The short-term memory representations of both the Forward and Forward-Left chunks are associated with the hunger drive. However, only the Forward chunk learns to predict the Forward-Straight body movement.

To summarize this learned maze trial, the animat is initially under reactive visual guidance. Then, previously learned planned responses override the reactive signals. An overview of the movement trajectory and related learning appears in Figure 4. An approach movement toward the triangle is followed by a head-orienting movement to the left. After the head-orienting movement, the square cue comes into view. Based on this new visual information, and the NET movement, a new plan item to approach the square cue is recalled. After approaching the square cue, the rewarded location is reached and adaptive weights are adjusted to strengthen the association between the currently active plan and the current motivational state. These learned expectations can be read out at a later time under the same motivational state.

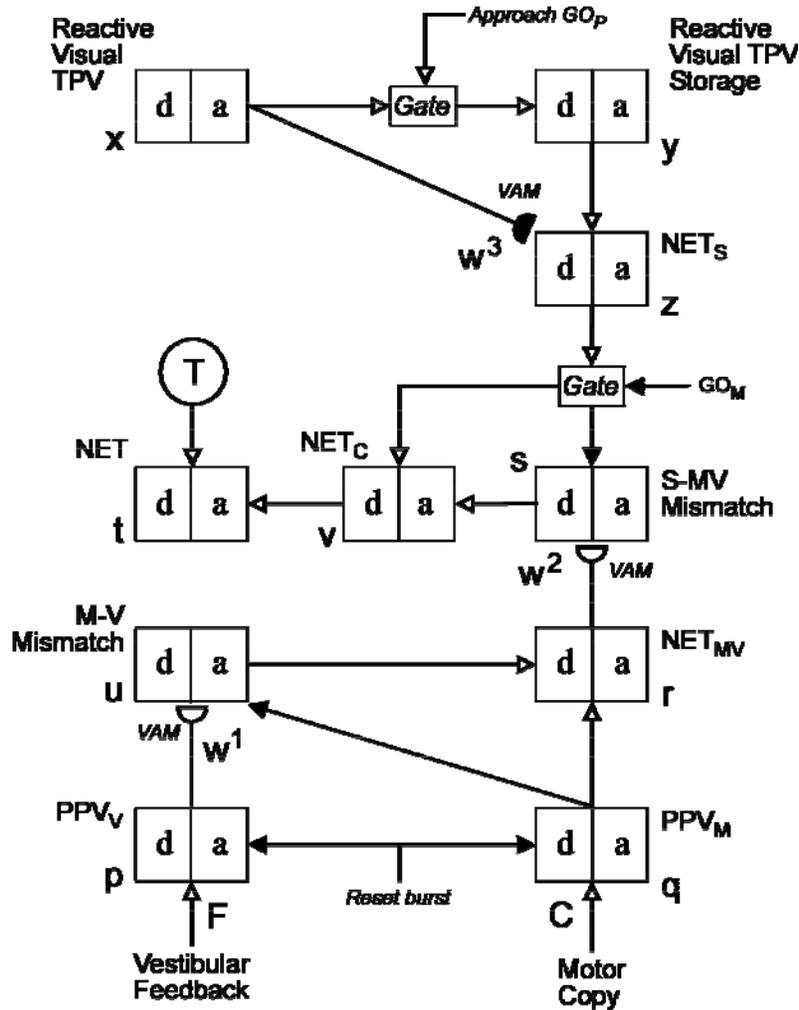
**2.5.1 Sensory-Motor NET Movement.** The Sensory-Motor NET Movement module (Figure 12) implements the body-centered spatial representation and computes net body displacement during

reactive and planned orienting and approach movements. The motor control system is a variant of the DIRECT model of arm reaching (Bullock, Grossberg, and Guenther, 1993) which has also been adapted as the DIVA model of speech articulator control (Guenther, 1995). Each of the model vectors is modeled by fast differential equations, which are approximated using algebraic equations. In Figure 12, “d” denotes the distance representation, whereas “a” denotes the angle representation of each field.

The Reactive Visual TPV module, at the upper left, receives input from the Visual Form and Motion System (Figure 5). Similarly, the Fuzzy ART Classifier module receives input from the Invariant Visual Target Map, then categorizes the input, which comprises the WHAT stream representation of the visual target object. A Reactive Visual TPV, the WHERE stream representation of the visual target object, arrives from the visual input system simultaneously. The visual target coordinates then flow into the Reactive Visual TPV Storage module (Section A10, equation (116)) and activates the Reactive DV module (Figure 10 and Section A9, equation (89)), which initiates head and body movements via the Motor Approach and Orienting System. As the Reactive Visual TPV Storage module is loaded, any differences between it and the Reactive Visual TPV module are registered at the  $NET_S$  field (Section A10, equation (117)). These signals are in body-centered spatial coordinates and denoted by an “s” subscript. The  $NET_S$  module, which depends on visual input, is only active when Parvo signals are present in the Invariant Visual Target Map module (Figure 5).

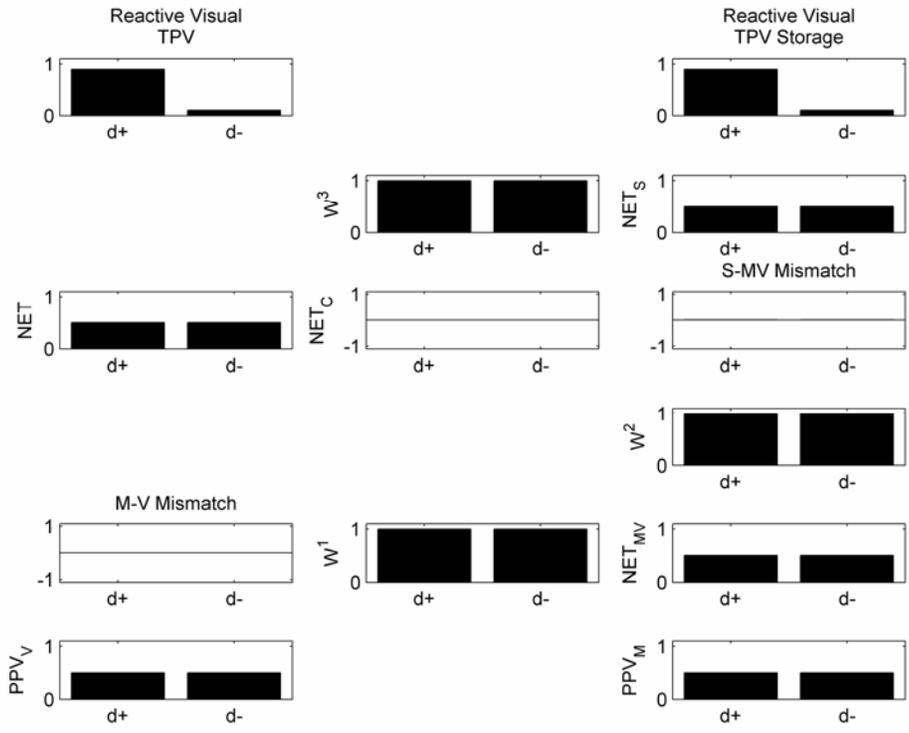
At the onset of the Approach  $GO_P$ , which is activated by the volitional release of the Reactive DV, adaptive weights learn to zero the activity of the  $NET_S$  module, which computes the visual representation of body displacement, via a burst of VAM learning at the adaptive weights,  $w^3$  (Section A10, equation (118)). A simulation of the Sensory-Motor NET Movement module in which a Reactive Visual TPV is stored and the adaptive weights are updated is shown in Figure 13a. In each of the panels of Figure 13, the distance fields of the Sensory-Motor NET Movement module are depicted. In Figure 13a, note that the initial target distance has been loaded and the adaptive weights have zeroed differences at the  $NET_S$  module. Furthermore, with zero response at the  $PPV_V$  and  $PPV_M$ , the NET indicates zero activity.

Updates to the Reactive Visual TPV Storage module (Figure 12) continue until the Approach or Orienting  $GO_P$  is activated (Section A9.3, equations (101), (103)). As the Reactive Visual TPV module activity changes during a movement, the  $NET_S$  module computes the difference between the current and stored initial target position. The NET module (Section A10, equations (120), (121)) combines both sensory and motor information by adding the  $NET_S$  and a vestibulo-motor correction computed by the S-MV Mismatch module. As such, the NET module comprises the primary inhibitory signal reaching the Reactive DV and Planned DV modules. The Planned DV module (Section A8, equations (78), (79)) computes the difference between the completed approach movement and the top-down planned target position. The NET module learns to inhibit the Reactive DV module via learned adaptive weights in the Motor Approach and Orienting System (Figure 10). Alternatively, the NET module can directly inhibit the Planned DV module. A simulation of the Sensory-Motor NET Movement module in which the NET displacement is computed using both vision and proprioception is shown in Figure 13b.

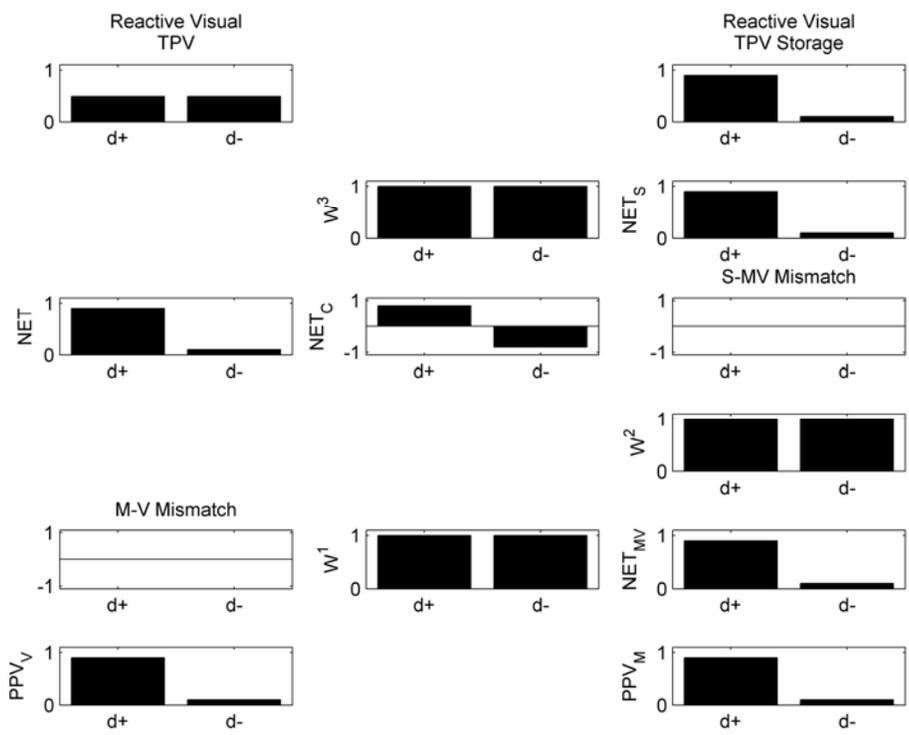


**Figure 12.** The Sensory-Motor NET Movement module learns to compute a robust representation of body displacement using a combination of visual, vestibular and motor copy signals (d = distance, a = angle).

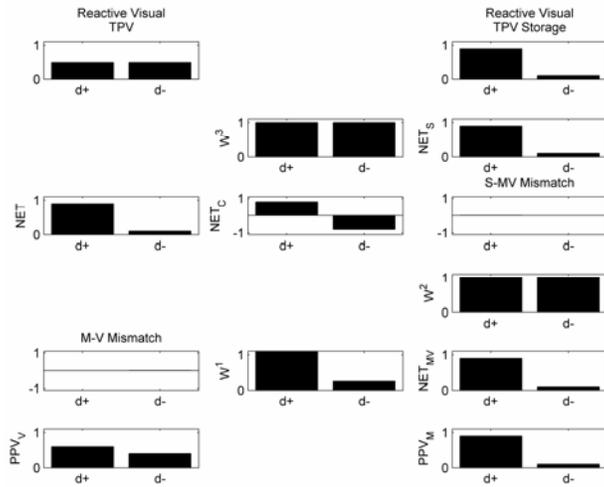
The NET<sub>MV</sub> module (Figure 12 and Section A10, equation (109)) computes a cumulative measure of body displacement from motor copy and vestibular signals. The Motor Copy input signal represents the directional efference copy feedback available internally to the animat. Motor efference signals are represented by a present position vector after integration by the PPV<sub>M</sub> module (Section A10.1). In addition, a measure of passive movement is computed from acceleration sensors in the vestibular system and integrated by the PPV<sub>V</sub> field (Section A10.1). The angular integrator is reset at an attentional shift, and both the angle and distance integrators are reset at the onset of a volitional movement signal.



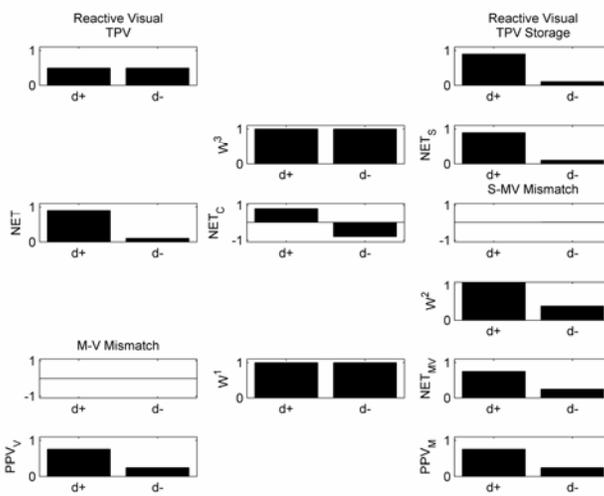
(a)



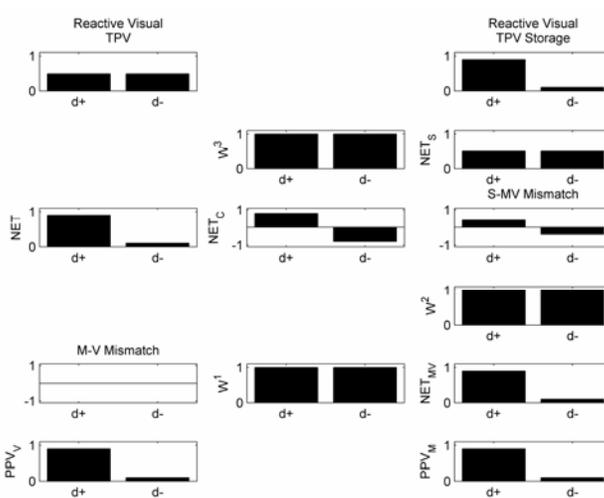
(b)



(c)



(d)



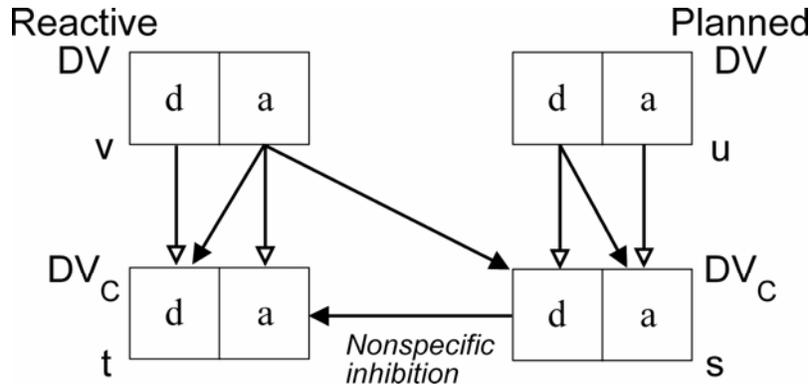
(e)

**Figure 13.** A series of simulations which depict the response of the Sensory-Motor NET Movement module under various conditions ( $d$  = distance). Only the distance response of each field within the network is shown for simplicity. **(a)** The Sensory-Motor NET Movement module response after a Reactive Visual TPV is stored and the adaptive weights ( $w^1$  and  $w^3$ ) are updated. The NET displacement is depicted by the differential height of the dark bars in the NET field on the left side of this panel. The NET field response at the outset of a movement is zero. **(b)** The Sensory-Motor NET Movement module response computed using both vision and proprioception. The large NET displacement is the result of computing the difference between the current distance to the target (Reactive Visual TPV) and the initial distance to the target (Reactive Visual TPV Storage). This difference, or NET displacement, is shown at the  $NET_S$  field. Similarly, the response of vestibular and proprioceptive integrators, are combined at the  $NET_{MV}$  field. Since the  $NET_S$  and  $NET_{MV}$  have identical activity, the S-MV Mismatch field is zero. **(c)** The relative gain of the vestibular and proprioceptive integrators is calibrated by adaptive weights ( $w^1$ ). The Sensory-Motor NET Movement module response is shown after any differences between vestibular and proprioceptive feedback are calibrated by VAM learning. The calibration process ensures that vestibular signals can compensate for potential inaccuracies in proprioceptive signals. **(d)** The Sensory-Motor NET Movement module response after the NET displacement is computed using visual signals. Simultaneously, the  $NET_{MV}$  module feedback is calibrated by adjusting adaptive weights ( $w^2$ ) using VAM learning. This learning occurs when the  $NET_S$  and  $NET_{MV}$  activity should be equal, as is the case when the animat has reached the visual target object. **(e)** The Sensory-Motor NET Movement module response after the NET displacement is computed without visual signals. Calibrated vestibular and proprioceptive feedback signals are able to provide an accurate estimate of body displacement.

The difference between the  $PPV_V$  and  $PPV_M$  vectors are computed by the M-V Mismatch module (Section A10, equation (106)). There, adaptive weights are adjusted by VAM learning such that the  $PPV_V$  signals are calibrated with respect to the  $PPV_M$  signals (Section A10, equation (107)). The VAM learning at the M-V Mismatch module is on continuously, to account for ongoing developmental differences between the motor and vestibular signals. Since the difference is usually zero, the  $NET_{MV}$  module is driven primarily by the  $PPV_M$  signals. However, in the event that the vestibular and motor signals differ, the VAM pathway projects a compensatory difference to the  $NET_{MV}$  module. A simulation of the Sensory-Motor NET Movement module in which a difference between vestibular and proprioceptive feedback are calibrated by adjusting adaptive weights ( $w^1$ ) appears in Figure 13c.

The  $NET_{MV}$  field (Figure 12) learns to zero the  $NET_S$  module activity via VAM learning at the  $w^2$  weights (Section A10, equation (115)) in the S-MV Mismatch module. A simulation of the Sensory-Motor NET Movement module in which the NET displacement is computed with vision and  $NET_{MV}$  module feedback and calibrated by adjusting adaptive weights ( $w^2$ ) is shown in Figure 13d. In the event that visual input is unavailable,  $NET_{MV}$  field can take over via this calibrated pathway, enabling feedback during blind movements. A final simulation shows the Sensory-Motor NET Movement module when the NET displacement is computed without vision appears in Figure 13e.

**2.5.2 Motor DV System.** Reactive movements by the animat are characterized by alternating head-orienting and body-approach movements. One main function of the Motor Approach and Orienting System (Figure 2), described above, is to control the interactions between reactive and planned motor commands. Another control structure is needed to regulate read-out of sequentially occurring alternations between approach and orienting movements. Furthermore, if both reactive and planned commands are present, as in the case of top-down readout at the outset of a learned trial, then the control hierarchy must reconcile the priority of each signal. The Motor DV System (Figure 14) carries out this task. In Figure 14, the “d” refers to the distance field and the “a” refers to the angle field of each module. Each of these fields, in turn, is described by complementary pairs of neural activation.

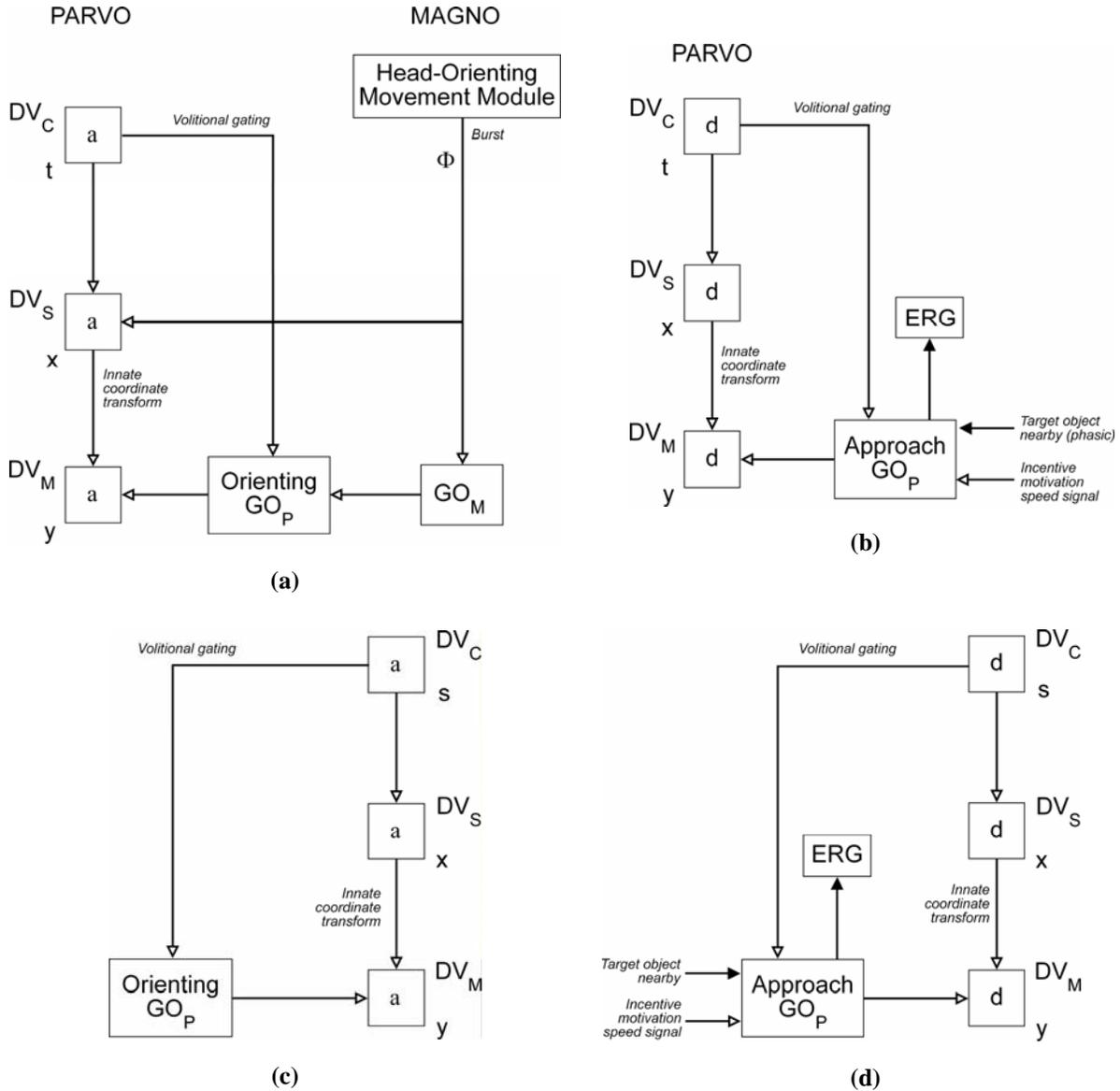


**Figure 14.** The Motor DV System computes the Reactive  $DV_C$  ( $d$  = distance,  $a$  = angle) and Planned  $DV_C$  module response by prioritizing reactive and planned body approach and head orienting movements.

The Motor DV System comprises four modules, shown in Figure 14: Reactive DV, Reactive  $DV_C$ , Planned DV and Planned  $DV_C$ . The Reactive DV module, codes the motor command in the form of an angle and distance to turn toward and approach a visual target cue. The Planned DV module, codes the planned motor command to approach and orient away from a visual target cue. By orienting away from a cue, a new visual target cue may come into the visual field. However, the Reactive DV and Planned DV cannot control the animat at the same time. The Reactive  $DV_C$  and Planned  $DV_C$  form a control hierarchy to resolves which head-orienting and body-approach command is to be performed next. The Reactive  $DV_C$  (Section A9.1, equation (92), (94)) and Planned  $DV_C$  (Section A9.1, equations (95), (96)) modules result from both direct and competitive signals from the corresponding DV modules.

There are four direct and four competitive connections from the distance and angle fields of the DV modules to the  $DV_C$  modules. In the first competition, the angle field of the Reactive DV inhibits the distance field of the Reactive  $DV_C$  module. This ensures that the animat will first orient toward the target visual cue before beginning an approach movement under reactive control. In the second competition, the angle field of the Reactive DV inhibits the distance field of the Planned  $DV_C$  module. This inhibition prevents the animat from releasing a planned approach movement before orienting directly toward the visual target cue. In the third competition, the distance field of the Planned DV inhibits the angle field of the Planned  $DV_C$  module. This prevents the animat from orienting away from a target visual cue before completing the corresponding planned approach movement. Finally, both fields of the Planned

DV<sub>C</sub> nonspecifically inhibit both fields of the Reactive DV<sub>C</sub> module. This strong inhibitory interaction enables a planned movement command to override a reactive movement command. The remaining components of the Motor DV System can now be introduced. These stages appear in the four panels of Figure 15.



**Figure 15.** (a) The Reactive Orienting System computes the head orienting outflow command, DV<sub>M</sub> (a = angle) in response to reactive parvo and magno visual signals. (b) The Reactive Approach System computes the body approach outflow command, DV<sub>M</sub> (d = distance) in response to parvo visual signals. (c) The Planned Orienting System computes the head orienting motor outflow command, DV<sub>M</sub> (a = angle) in response to planned head orienting signals. (d) The Planned Approach System computes the body approach outflow command DV<sub>M</sub> (d = distance) in response to planned body approach signals.

**2.5.3 Reactive Orienting System.** The Reactive Orienting System, or ROS, computes head-orienting outflow commands in response to visual target position signals (Section A9.2). A block diagram of the ROS appears in Figure 15a. The ROS operates in response to two different types of input. In the first, a reactive head-orienting movement toward a visual target cue is computed by the Motor DV System (Figure 14). Such a head-orienting command, in turn, activates the Reactive  $DV_C$  module. The angle field of the Reactive  $DV_C$  module appears in the upper left of Figure 15a. Any activation of the  $DV_C$  angle field projects directly to the  $DV_S$ , which computes a representation of how far to turn in sensory coordinates.

The mapping from  $DV_S$  to  $DV_M$  assumes that the normalized opponent pair of cells representing the body-centered direction of the target can be directly used to compute a motor command. The  $DV_S$  vector is transformed into a  $DV_M$ , or motor coordinate, vector via an innate spatial-to-motor transform. The rule for this transform is that the direction of body movement is in the same direction as the target position. The DIRECT model suggests how this transform may be learned (Bullock et al., 1993).

In the second type of input, the Head-Orienting Movement module (Figure 5) releases a burst of activity in response to peripheral Magno signals. This burst is denoted by  $\Phi$  in the upper right of Figure 15a. The orienting burst can directly activate the  $DV_S$  and gates itself by activating  $GO_M$ , the Magno GO gate. When the  $GO_M$  is active, the animat is said to be performing an attentional shift. Sufficient activation of the  $GO_M$  then excites the Orienting  $GO_P$ , which initiates a reactive head-orienting movement away from the visual target cue.

The motor outflow command for orienting movements is computed by scaling the  $DV_M$  by volitional GO signals from the  $GO_M$  and Orienting  $GO_P$ , which multiply DV commands and thereby control the overall speed of movement (cf. Bullock and Grossberg, 1988). In particular, the Orienting  $GO_P$  signal directly modulates the orienting head movement speed. Furthermore, the GO signal is designed to remain activated until the  $DV_C$  field is zero and the movement is complete. For example, the animat is designed to perform a head turn at a velocity that precludes a body roll over. As mentioned previously, the computation of the volitional signals and stability limits is hypothesized to occur via circuits (e.g., the basal ganglia) that are beyond the scope of this project.

**2.5.4 Reactive Approach System.** The Reactive Approach System, or RAS, computes body-approach outflow commands in response to visual target position signals (Section A9.2). A block diagram of the RAS appears in Figure 15b. The RAS operates in response to reactive head approach movements, as computed by the Motor DV System, discussed above (Figure 14). The distance vector of the Reactive  $DV_C$  module appears in the upper left of Figure 15b. Any activation of the  $DV_C$  distance vector projects directly to the  $DV_S$ , which computes a representation of how far to approach in sensory coordinates. The  $DV_S$  vector is transformed into a  $DV_M$ , or motor coordinate vector, via an innate spatial-to-motor transform (Section 2.3.3).

The Approach  $GO_P$  modulates the activity of the distance field of the  $DV_M$ , in order to control the approach movement. The activity of the Approach  $GO_P$  derives primarily from a volitional signal, released upon activation of the  $DV_C$ . The Approach  $GO_P$  is also modulated by a phasic signal indicating that a target object is nearby. This suffices to stop an approach movement and prevent a collision if the animat is too close. Furthermore, the Approach  $GO_P$  is amplified (Section A9.3, equation (104)) by incentive motivational signals released by the Sensory-Drive Hierarchy (Section 2.4.3). The resulting motivational speedup emulates the well-known increase in running speed as a navigating rat reaches a goal location (Hull, 1932, 1934).

An additional function of the Approach  $GO_P$  is to inhibit random head-orienting signals from the ERG, discussed in Section 2.3.7. Finally, the  $DV_M$  generates an outflow motor command which defines the desired speed and direction of movement.

**2.5.5 Planned Orienting System.** The Planned Orienting System, or POS, computes head-orienting outflow commands in response to planned target position signals (Section A9.2). A block diagram of the POS appears in Figure 15c. The POS operates in response to planned head-orienting movement commands computed by the Motor DV System (Figure 14). Such a head-orienting command, in turn, activates the Planned  $DV_C$  module. The angle vector of the Planned  $DV_C$  module appears in the upper right of Figure 15c. Any activation of the  $DV_C$  angle vector projects directly to the  $DV_S$ , which computes a representation of how far to turn in sensory coordinates. The innate mapping from  $DV_S$  to  $DV_M$  assumes that the normalized opponent pair of cells representing the body-centered direction of the target can be directly used to compute a motor command (Section 2.3.3). Finally, the  $DV_M$  generates an outflow motor command that defines the desired speed and direction of head orienting.

**2.5.6 Planned Approach System.** The Planned Approach System, or PAS, computes body-approach outflow commands in response to planned target position signals (Section A9.2). A block diagram of the PAS appears in Figure 15d. The PAS operates in response to planned head approach movements, as computed by the Motor DV System, discussed above (Figure 14). The distance vector of the Planned  $DV_C$  module appears in the upper right of Figure 15d. Any activation of the  $DV_C$  distance vector projects directly to the  $DV_S$ , which computes a representation of how far to approach in sensory coordinates. The  $DV_S$  vector is transformed into a  $DV_M$ , or motor coordinate vector, via an innate spatial-to-motor transform (Section 2.3.3).

The Approach  $GO_P$  modulates the activity of the distance vector of the  $DV_M$ , in order to control the approach movement. The activity of the Approach  $GO_P$  derives primarily from a volitional signal, released upon activation of the  $DV_C$ . The Approach  $GO_P$  is also modulated by a phasic signal indicating that a target object is nearby. This suffices to stop an approach movement and prevent a collision if the animat is too close. Furthermore, the Approach  $GO_P$  is influenced by incentive motivational signals released by the Sensory-Drive Hierarchy (Section 2.6.3). An additional function of the Approach  $GO_P$  is to inhibit random orienting signals from the ERG (Section 2.3.7). Finally, the  $DV_M$  generates an outflow motor command which defines the desired speed and direction of movement.

**2.5.7 Endogenous Random Generator.** During exploratory movements in the environment, the animat acquires feedback about cues that may result in reward. As shown in the plus-maze example of Figure 1a, the animat may initially see the triangle cue and approach it. Say, for instance, that as it approaches the choice point, visual motion cues are strong enough to win the competition in the Spatial Attention module (Figure 5). Furthermore, say the Magno system generates a head movement to the left. The animat then rapidly performs the head turn, while tracking the motion cue generated by new visual information in the visual field.

When the head turn is complete, the Magno system is quiescent, and a stable Fuzzy ART Classifier module (Figure 5) match signal is computed. The body-centered coordinates of the target are loaded into Reactive Visual TPV module and an approach movement is computed by the Motor Approach and Orienting System (Figure 10), then by the  $DV_S$  to  $DV_M$  transform of the RAS (Section 2.3.4). This step initiates both an approach movement and counter-rotation of the body. During this time, the animat homes toward the visual target. Finally, as the animat reaches

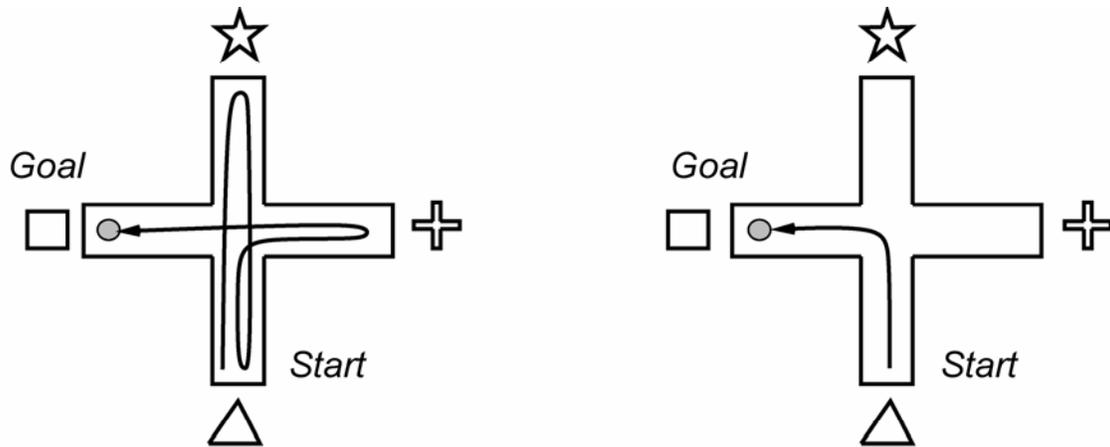
the target, the distance field of the  $DV_s$  signal reaches zero. No visual cues remain to guide approach or orienting movements. If the animat is not rewarded within some period of time, then it will likely perish. Part of the control system designed to address this situation enables the animat to turn around and continue to explore the environment.

A proposed solution for enabling a turn-around is via an endogenous random generator, or ERG, that is adapted from the Vector Associative Map, or VAM, model of Gaudio and Grossberg (1991). The proposed ERG (Section A9.3, equation (105)) creates a random burst of activation in body-centered motor coordinates, as part of the ROS (Section 2.3.3). When the ERG is firing, the animat makes a rapid head turn in one direction. The action of the ERG turns the animat away from the current cue and activates the Magno orienting system. Once the head turn has started, the Magno system takes over and continues the turn until the visual field reaches a stable state, as specified above. In the model, the ERG is disinhibited at a dead-end, after a short delay. This allows the animat to continue exploration if not rewarded at the dead-end.

**2.6 Visual and Motor Working Memory and Planning System.** For each traversal through the plus-maze, the animal can build a representation of the location of the goal in a short-term serial order buffer, or working memory. This representation helps to learn a plan capable of solving the maze. The activation of a stored plan can be strengthened by associating the plan with reward, or satisfaction of a drive. The more often a successful sequence is rewarded, the more likely it will be recalled in the same motivational state. On subsequent trials, that information is used to direct the return to the goal location.

Experiments with rats have shown that learned spatial movement plans, or motor habits (O'Keefe and Nadel, 1978), can be very strong. This conclusion is the result of testing rats in a complex Hampton-court maze (Carr and Watson, 1908). Once the rats had become familiar with the apparatus, several experimental manipulations were attempted. The major findings were twofold. First, when one arm of the maze was shortened, the rats would run into the end of the corridor, even if not visually impaired. Second, when cul-de-sacs were constructed where turns to open corridors once had been, the rats again ran into the wall. It was subsequently shown that when previously correct arms end in open space, the rats run off the end of the arm (Dennis, 1932). Finally, it has been observed that when a rat has learned to run to a small food reward at the end of an alley, and a large pile of food is placed in the middle, the rat runs over the food for more than twenty trials (Stoltz and Lott, 1964). In the competition between visual stimuli and the learned plan, the planned response often takes precedence.

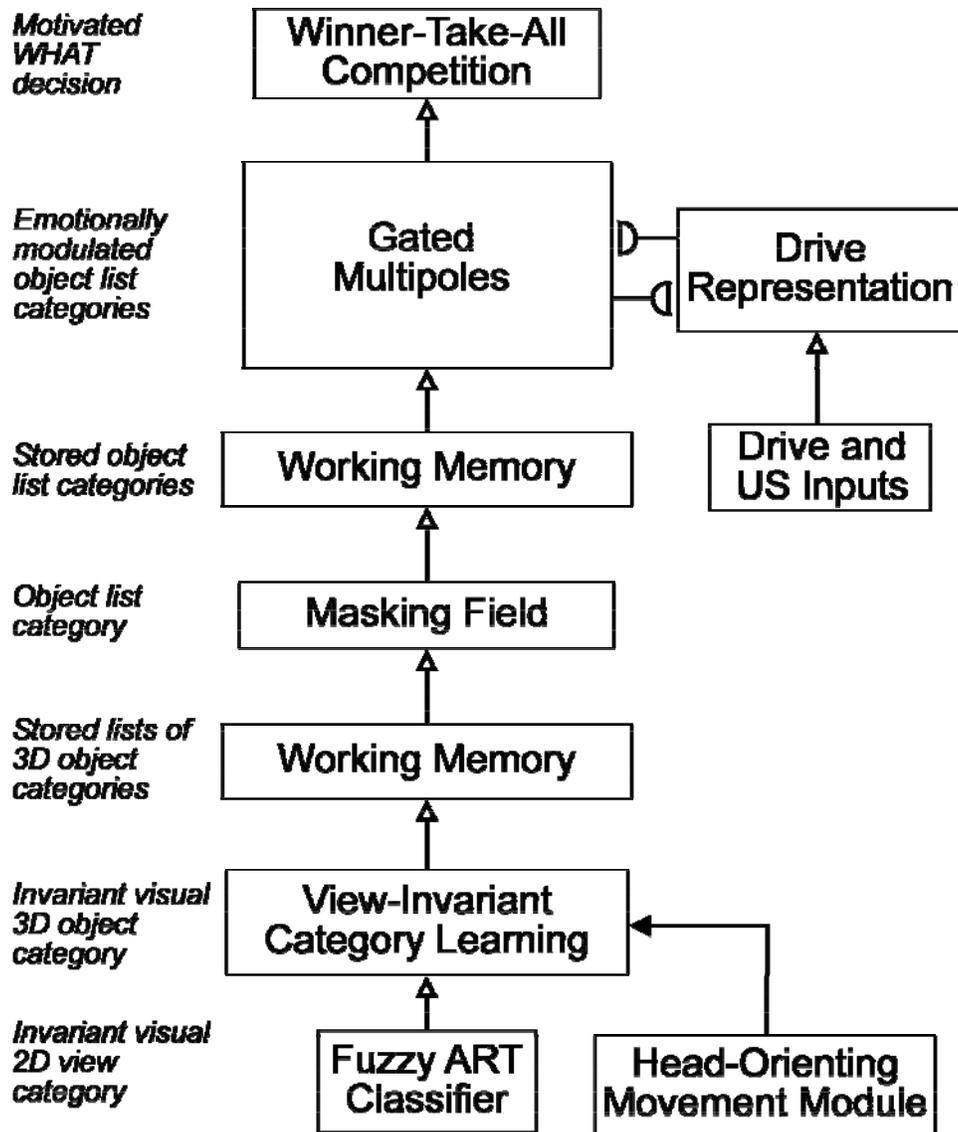
The storage process during learning of plans is important as it can affect the overall efficiency of planned behavior. For example, during an exploratory run of the plus-maze, the animat may execute the path shown on the left panel of Figure 16. This sequence contains extra loops before reaching the goal. The corresponding turn sequence would be: Forward, Turnaround, Forward, Turnaround, Forward, Right, Forward, Turnaround and Forward. For the optimal path, shown in the right panel of Figure 16, the sequence is simply: Forward, Left, and Forward. However, the two paths lead to the same goal and therefore both satisfy the animal's internal need. How can the practiced path be simplified in memory?



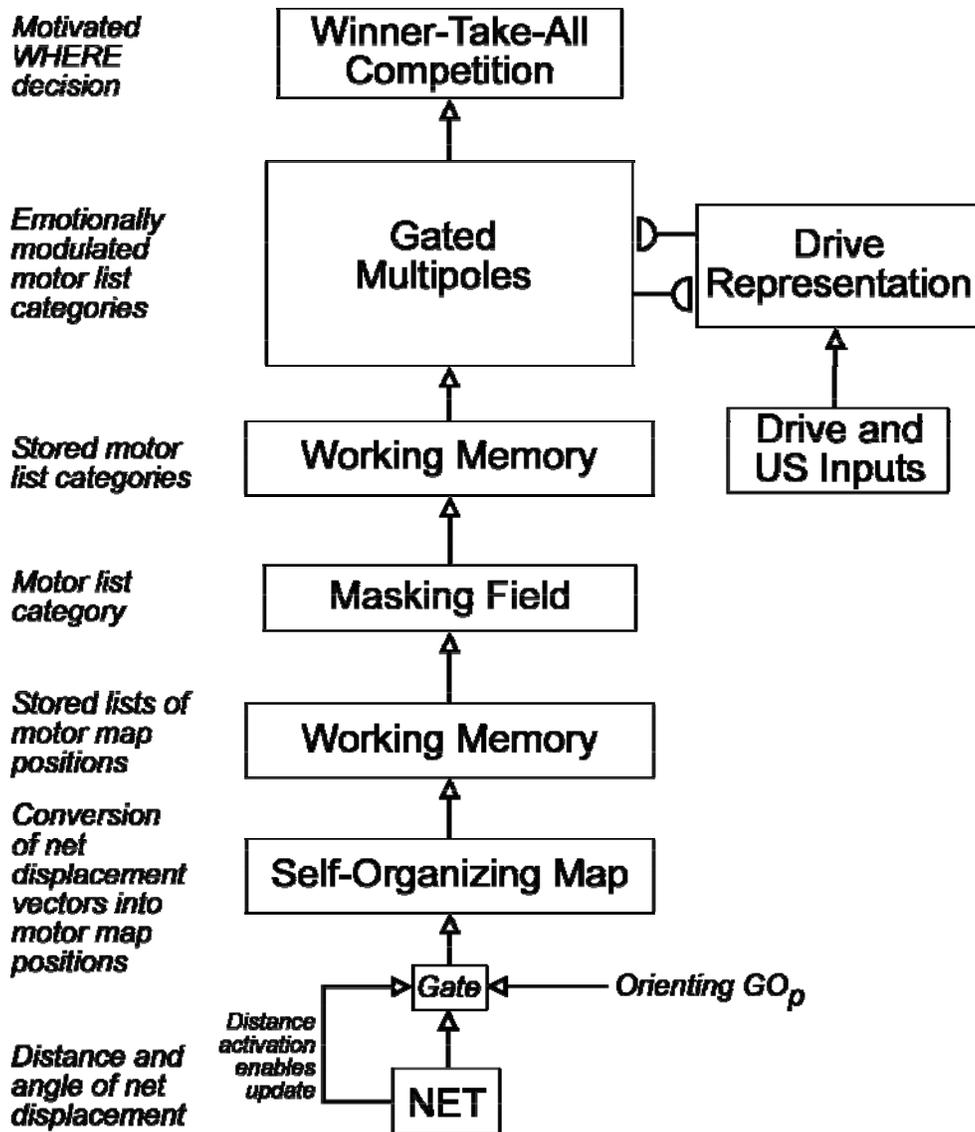
**Figure 16.** In this diagram, the overhead view of two plus-mazes show the exploratory path (left panel) and stored path to the goal (right panel). This type of simplification involves a translation between spatially-equivalent movement sequences.

Such a simplification would require a translation that is sensitive to the equivalence of the two movements. This is similar to the notion of motor equivalence (Saltzman, 1979), in which spatial movements with different trajectories have the same spatial consequence. It is possible that rats, like humans, are capable of a geometric sense (Leonard and McNaughton, 1990), which might also be applied in the generation of novel trajectories between familiar places. Another possible explanation is that repeated exploration of the environment might allow stored representations of successful, rewarded plans to be selected by conditioning. Thus, more frequently rewarded plans would be selectively amplified relative to others. This is what happens in SOVEREIGN.

The Visual Working Memory and Planning System temporarily stores and categorizes, or chunks, sequences of categorized objects and learns the corresponding approach-orient motor commands, via top-down learning from the Motivated WHAT and WHERE Decisions. A detailed diagram of the Visual Working Memory and Planning System, which includes the stages that compute the Motivated WHAT and WHERE Decisions, is shown in Figure 17. Sequence, or list chunks, are learned that selectively respond to object sequences of variable lengths. This happens in the Masking Field module (Figure 17). The Visual Working Memory and Planning System operates in parallel with a Motor Working Memory and Planning System, which temporarily stores and categorizes, or chunks, sequences of motor commands. A detailed diagram of the Motor Working Memory and Planning System is shown in Figure 18. Together these parallel Visual and Motor working memories can disambiguate decisions that only one of them would find ambiguous. When the animat receives reward, the active list chunks, which play the role of plans, are associated with active drives and actions. The animat is hereby capable of learning different plans to satisfy different motivational goals. After such learning occurs, when the animat sees a familiar sensory cue under a prescribed motivational state, it can recall a motivationally-compatible plan to reach the site of previous reward. Repeated, random exploration of the environment effects a gradual transition from reactive to more efficient, planned control that leads the animat to its various motivated goals. Each component of the multi-stage Visual and Motor Memory System is discussed in the following sections.



**Figure 17.** The Visual Working Memory and Planning System computes motivationally-reinforced representations of sequences of 3D object categories.



**Figure 18.** The Motor Working Memory and Planning System computes motivationally-reinforced representations of sequences of motor map positions.

**2.6.1 Working Memory.** This model working memory circuit uses the STORE 2 working memory model (Bradski et al., 1992) for both sensory and motor items. Two working memory modules appear in the Visual Working Memory and Planning System block diagram of Figure 17 and two in the Motor Working Memory and Planning System block diagram of Figure 18. In the Visual Working Memory and Planning System, the STORE 2 network first codes the order of the items from the View-Invariant Category Learning module (Section A4.1). Thus, the short-term activity of the 3D Object Working Memory neurons encodes a stored list of 3D object categories. Only one invariant visual 3D object category is activated while attending to its visual cue. The invariant visual 3D object category inputs to the Object Working Memory module after

a match is computed and a volitional signal activates the Approach GO<sub>p</sub> control cell. Normalization within the STORE working memory network is tuned to generate a recency gradient, in which the item order is coded by magnitude, with the most recent item having the greatest activation.

Finally, the outputs of the Object Working Memory module (Section A4.2, equation (47)) provide parallel inputs to the Masking Field module (Section A4.3), which learns to code object list categories. The second working memory module, near the top of Figure 17, stores sequences of object list categories. The cells representing object list categories input to this Object List Working Memory (Section A4.4) during a reset event. The output of the Object List Working Memory makes parallel connections to the Gated Multipole module (Section 2.4.5) which fuses cognitive plans with information about the rewards they acquire.

The first working memory in the Motor Working Memory and Planning System (Figure 18) employs an algorithmic vector-to-map transform to convert the body movement vector, comprised of a distance and angle, into a spatial pattern (Section A7). Also assumed are winner-take-all dynamics which result in the activation of single map positions in response to distance and angle pairs. There are several networks capable of accomplishing this function. One such transform is the DODOG, or Difference of Difference of Gaussians (Gaudio and Grossberg, 1991) inputting into a recurrent competitive field (Grossberg, 1982a) with faster-than-linear feedback signal functions to choose the cell with the largest input. Whatever vector-to-map transform is used, the winning cell represents the current position of the body and is input to the Motor Map Working Memory after a head-orienting movement is complete.

The second working memory, near the top of Figure 18, stores sequences of motor list categories. The cells representing motor lists are printed to this Motor List Working Memory during a reset. The output of the Motor List Working Memory makes parallel connections to the Gated Multipole module (Section 2.4.5).

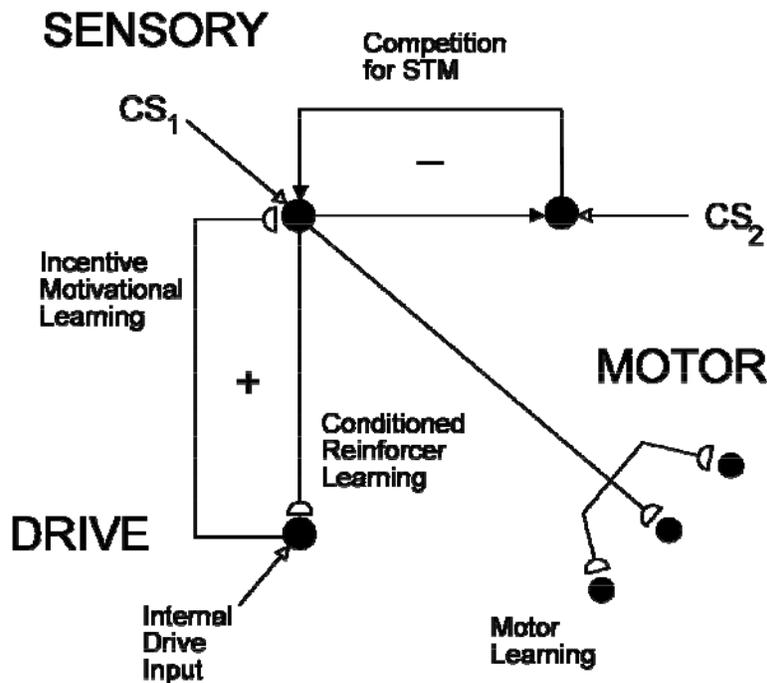
**2.6.2 Masking Field.** The Visual and Motor Working Memory and Planning Systems (Figures 17 and 18) use a modified masking field network (Cohen and Grossberg, 1986, 1987; Grossberg, 1978) to code lists of sensory cues and motor commands to reach a goal. The masking field is comprised of a network of input cells ( $F_1$ ) and a network of chunk cells ( $F_2$ ). Network  $F_1$  forms a spatial pattern of activation that stores both the items and their temporal order from the Object or Motor Map Working memory. As discussed above, the Object Working Memory stores sequences, or lists, of object categories, while the Motor Map Working Memory codes stores lists of motor map positions. The  $F_2$  field list chunk cells are selectively activated by sequences of  $F_1$  items (Section A4.3, equation (48)). List chunk activation is computed using bottom-up weighted signals from  $F_1$  and a competitive process between  $F_2$  list chunks that selects the winning list representations. There are several modifications hypothesized to the original masking field architecture in order to achieve robust performance in response to bottom-up signals from the Visual and Motor Working Memory modules. First, inputs to  $F_1$  are scaled by a size-dependent factor. Second, bottom-up adaptive weights are assumed to have achieved their steady-state values, subsequent to long-term memory (LTM) adaptation. Finally, inputs to  $F_1$  are subject to ART1 (Carpenter and Grossberg, 1987) category constraints via an additional dynamic process (Section A4.3, equations (50), (51) and (52)).

The masking field is duplicated for sequences of sensory items and motor items, shown in the block diagrams of Figures 17 and 18. The winning cells form object and motor list categories, which play the role of plans. Subsequences of each plan may also be coded by list chunks within the masking field. However, their activation is suppressed by plans representing

longer event sequences by competitive dynamics within the field of  $F_2$  cells. Active list chunks in  $F_2$  of each masking field are stored, during a reset, in the Object or Motor Map Working Memory, respectively. Stored sequences of winning list chunks in the Object List or Motor List Working Memory module are then processed by the Gated Multipole module (Section 2.4.5).

**2.6.3 Drive Representation and Sensory-Drive Heterarchy.** The Drive Representation module (Figures 2, 17 and 18) represents the animat's response to internal drive inputs, to external unconditioned stimulus (US) inputs, and to learned conditioned reinforcer inputs (Section A5, equation (67)). The Sensory-Drive Heterarchy module (Section A6, equations (68), (69)) chooses a winning sensory-drive combination in response to signals from the Drive Representation module. The Sensory-Drive Heterarchy module also influences the Gated Multipole module (Section 2.4.5) via learned incentive motivational signals. The combination of incentive signals from the Sensory-Drive Heterarchy module, that reflect the history of rewards, and bottom-up signals from the Object List and Motor List Working Memory modules (Figures 17 and 18), can bias the choice of output channels from the Gated Multipole module. The selected output channel can release a learned motor response via the Motivated WHAT and WHERE decision stages.

Both the Drive Representation and Sensory-Drive Heterarchy modules build upon the CogEM model. CogEM joins cognitive, emotional, and motor processes (Grossberg, 1982b, 1984a, 2000a, 2000b; Grossberg and Seidman, 2006) to model how emotional centers of the brain, such as the amygdala, interact with the sensory and prefrontal cortices – notably the ventral, or orbital, prefrontal cortex – to generate affective states, attend to motivationally salient sensory events, and elicit motivated behaviors. Activation of the feedback loop between cognitive and emotional centers is predicted in the brain to generate a cognitive-emotional resonance that can support conscious awareness of events happening in the world and how one feels about them.



**Figure 19.** The simplest *CogEM* (Cognitive-Emotional-Motor) model: Three types of interacting representations (sensory, drive and motor) that control three types of learning (conditioned reinforcer, incentive motivational, and motor) help to explain many reinforcement learning data. Sensory representations S temporarily store internal representations of sensory events in working memory. Drive representations D are sites where reinforcing and homeostatic, or drive, cues converge to activate emotional responses. Motor representations M control the read-out of actions. Conditioned reinforcer learning enables sensory events to activate emotional reactions to drive representations. Incentive motivational learning enables emotions to generate a motivational set that biases the system to process information consistent with that emotion. Motor learning allows sensory and cognitive representations to generate actions.

The diagram in Figure 19 summarizes the hypothesis that (at least) three types of internal representation interact during reinforcement learning: sensory and cognitive representations S, drive representations D, and motor representations M. Sensory representations S temporarily store internal representations of sensory events in working memory, Drive representations D are sites where reinforcing and homeostatic, or drive, cues converge to activate emotional responses. Motor representations M control the read out of actions. Three types of learning take place among these representations: Conditioned reinforcer learning (CRL) enables sensory events to activate emotional reactions at drive representations. Incentive motivational learning (IML) enables emotions to generate a motivational set that biases the system to process cognitive information consistent with that emotion. Motor learning allows sensory and cognitive representations to generate actions.

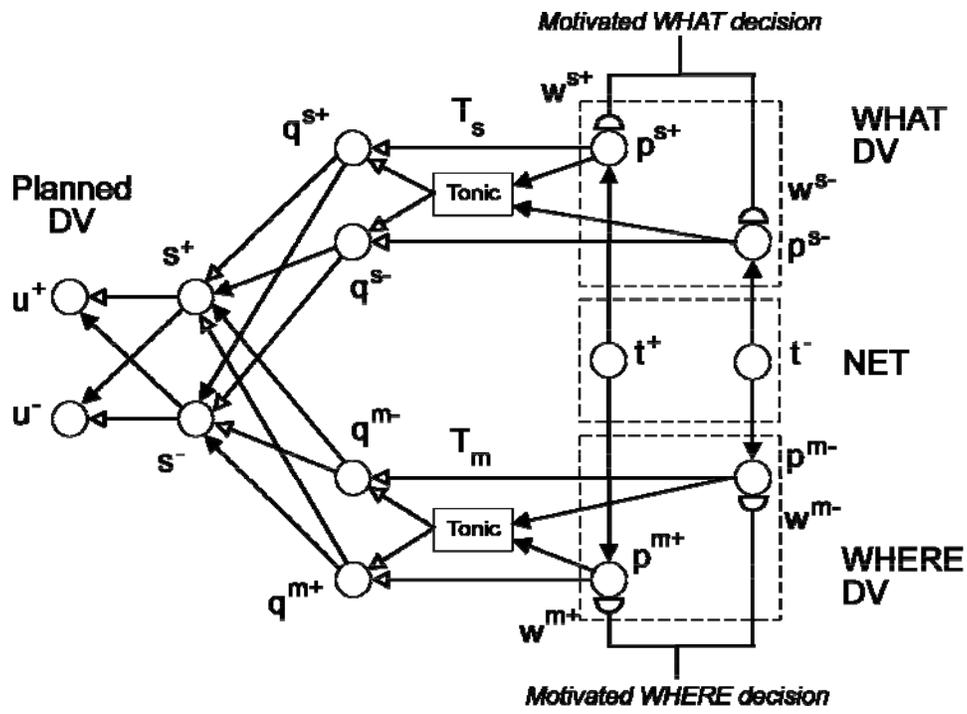
In particular, learning within the S→D conditioned reinforcer pathways converts a conditioned stimulus (CS) into a reinforcer when it activates its sensory representation just before the drive representation is activated by an unconditioned stimulus (US), or other previously conditioned reinforcer CSs. The ability of the CS to subsequently activate D via this learned pathway is one of its key properties as a conditioned reinforcer. As these S→D associations are being formed, incentive motivational learning within the D→S incentive motivational pathways also occurs, because of the same pairing of CS and US. Incentive motivational learning enables an activated drive representation to prime or modulate the sensory representations of all cues, including the CSs that have consistently been correlated with it. That is how activating D generates a “motivational set”: It primes all of the sensory and cognitive representations that have been associated with that drive in the past. These incentive motivational signals are a type of motivationally biased attention. The S→M motor, or habit, learning enables the sensory-motor maps, vectors, and gains that are involved in sensory-motor control to be adaptively calibrated, thereby enabling a CS to read out correctly calibrated movements. Taken together, these processes control the learning and recognition of sensory and cognitive memories, which are often classified as part of a “declarative memory” system (Mishkin, 1982, 1993; Squire and Cohen, 1984), and the performance of learned motor skills, which are often classified as part of a “procedural memory” system (Gilbert and Thach, 1977; Ito, 1984; Thompson, 1988).

Each drive representation obeys a polyvalent constraint whereby it can generate large incentive motivational output signals to sensory representations only if it gets a sufficiently large primary or conditioned reinforcer input at the same time that it gets a sufficiently large internal drive input. The internal drive input designates whether an internal drive, such as hunger, thirst

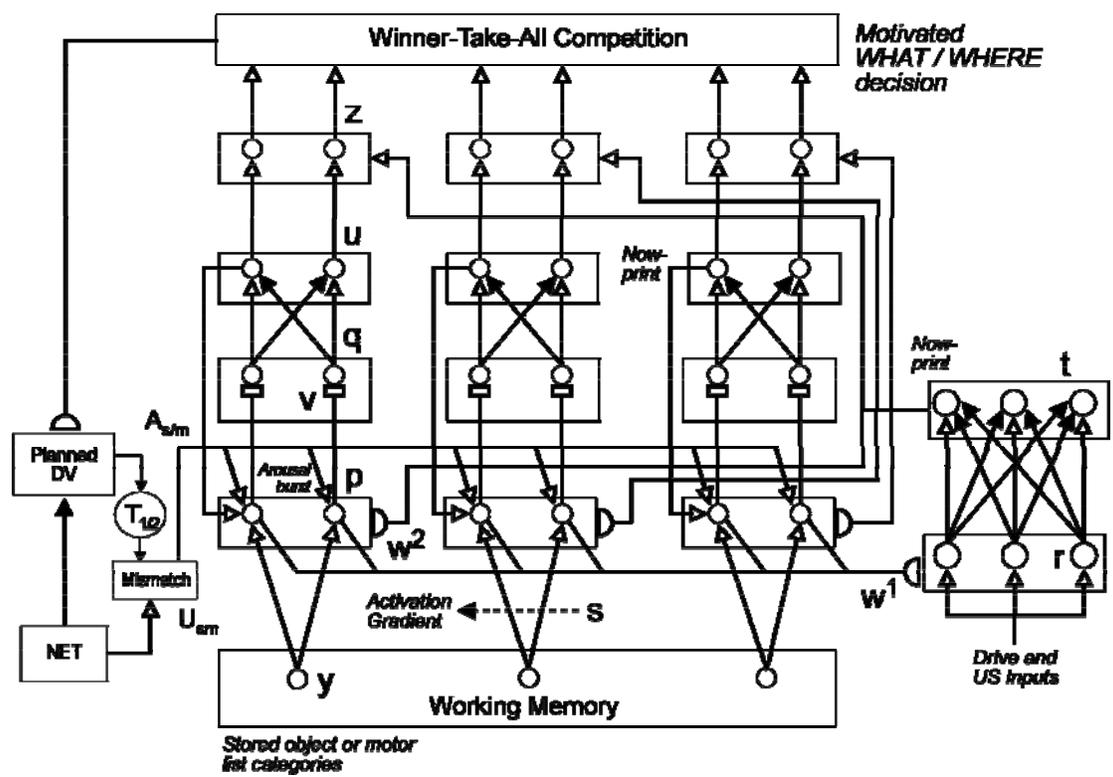
or sex, is high and in need of satisfaction. Different drive representations exist to represent these distinct internal homeostatic states. The polyvalent constraint means that a drive representation cannot fire vigorously unless it simultaneously receives a sufficiently large external sensory input and internal drive input. Because of the polyvalent constraint at the drive representation, an external reinforcing cue cannot activate strong incentive motivation and, with it, action to satisfy a drive that is already satisfied, because the drive input would be too small.

The Gated Multipole and Drive Representation circuits of Figures 17 and 18 build upon CogEM circuits to model how multiple cues can be conditioned to multiple drives, notably how a transfer can occur from behavior that is motivated by an exploratory drive to behavior that is motivated by a specific consummatory drive after it is rewarded and used to direct planned behavior.

**2.6.4 Sensory-Motor Top-Down Readout.** The Sensory-Motor Top-Down Readout module (Figure 20a) combines information from the Visual and Motor Working Memory and Planning Systems (Figures 17 and 18). These systems compute the Motivated WHAT and WHERE Decisions which comprise a spatial pattern representing a motor command. The Top-down Readout module combines the Motivated WHAT and WHERE decisions to form a composite expectation representing the next motor command. Both primes are required to distinguish between mazes with the same sequence of sensory cues, but different motor commands, and vice versa. The Sensory-Motor Top-Down Readout module (Figure 20a) computes the sensory-motor Planned DV activity (Section A8, equations (78), (79)) by combining learned signals from the Motivated WHAT and WHERE Decisions (Section A8, equations (80), (81)). Top-down readout activates the WHAT and WHERE DV, (Section A8, equations (70), (71)) which is inhibited by signals from the NET module (Section A10, equation (120)). Activity of the WHAT and WHERE DV fields are rescaled by tonic activation at second-layer cells (Section A8, equations (73), (74) and (75)). The sensory-motor convergence cells (Section A8, equations (76), (77)), compute a normalized sum of WHAT and WHERE DV activity which is tolerant to the absence of either top-down signal. Finally, Planned DV activity is computed via a differential competition from sensory-motor convergence cells.



(a)



(b)

**Figure 20.** (a) The Sensory-Motor Top-down Readout module combines spatial patterns comprising distance/angle motor doublets from the Motivated WHAT/WHERE Decision module. The resulting motor doublet forms a planned movement command in the Motor Approach and Orienting System. (b) The Gated Multipole module is a multi-layer neural network which combines bottom-up signals from either the Visual or Motor Memory System with learned motivational signals from the Sensory-Drive Heterarchy module to compute a motor response.

**2.6.5 Gated Multipoles.** The dynamics of the Masking Field and Working Memory modules (Figures 17 and 18) are insufficient to execute a plan memory in a motivationally appropriate way. In particular, it is known that rats chunk movement sequences in mazes by the type of reward (Dallal and Meck, 1990). List chunks of visual cues and motor commands are therefore associated with active drive representations within the model. Such lists, or chunks, are also capable of eliciting learned responses. During performance, previously rewarded chunks will be preferentially primed by active drives, ensuring that the selected plan is motivationally compatible.

The Gated Multipole module addresses the need to combine object and motor list categories, that are stored in the Working Memory module, with the current motivational state. The result is the choice of Motivated WHAT and WHERE decisions (Figures 17 and 18) which respond to joint cognitive and motivational contextual influences. A detailed wiring diagram of the Gated Multipole module appears in Figure 20b. The Working Memory represents stored object or motor list categories. Each category within the Working Memory activates a separate four-layer Gated Multipole whose pathways represent distinct drive channels. Gated Multipoles generalize gated dipoles, wherein only two opponent channels are available to represent the responses to bottom-up signals.

Activation of the Gated Multipole first-layer cells (Section A4.5, equation (54)) depend on bottom-up signals from the Working Memory and learned incentive signals, corresponding to different drives, (Section A4.5, equation (65)) from the Sensory-Drive Heterarchy (Section A6, equations (68), (69)). First-layer cells can, in turn, input to drive representations (Section A5, equation (67)) via learned conditioned reinforcer signals (Section A4.5, equation (62), (63) and (64)). The transmitters (rectangular synapses in Figure 20b; see Section A4.5, equation (57)), habituate (inactivate, depress) in response to output signals from first-layer cells. The activity of second-layer cells (Section A4.5, equation (58)) responds to signals from first-layer cells that are multiplicatively gated by transmitter level. Third-layer cells carry out a winner-take-all competition of signals from second-layer cells (Section A4.5, equations (59), (60)). Feedback from third-layer cells to first-layer cells achieves this choice by amplifying the largest signal in the feedforward pathway. The activity of fourth-layer cells (Section A4.5, equation (61)) depends on input from third-layer cells and a nonspecific excitatory signal from the Sensory-Drive Heterarchy. This nonspecific signal ensures that only plans which receive sufficient motivational support to fire Sensory-Drive Heterarchy outputs are allowed to release a top-down planned movement command and to potentially override a reactive movement command.

SOVEREIGN employs Gated Multipoles with three separate drive channels. This expands a typical gated dipole (Section 2.2). The rationale behind this choice is the need to represent exploratory and multiple consummatory drive states. Each object and motor list category stored in Working Memory can be associated with three distinct outcomes, one for each

channel. Furthermore, individual channels can be selected via a rebound mechanism similar to that of a gated dipole. This mechanism operates in four steps:

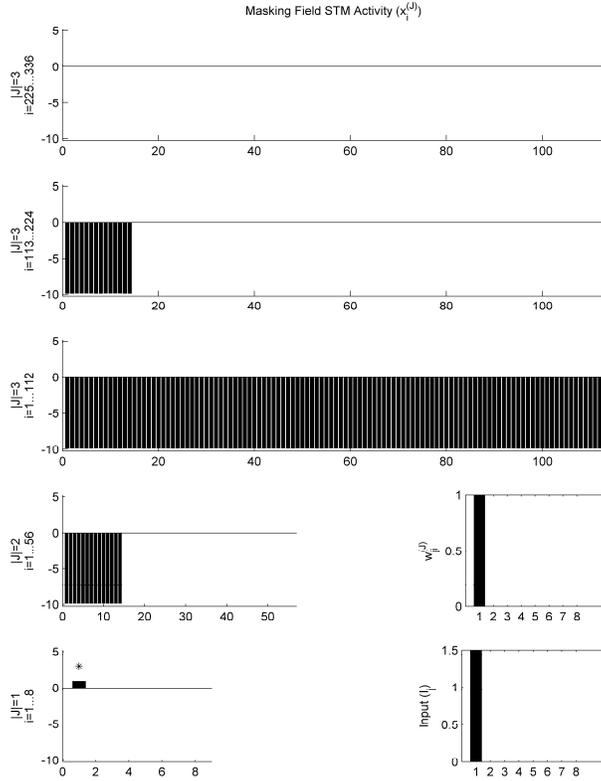
(1) Bottom-up signals from the Working Memory activate first-layer cells via a fixed spatial gradient of connection strengths across the channels. The drive channel with the strongest, or prepotent, bottom-up pathway receives the largest input. As noted above, first-layer cells then activate second-layer cells via signals that are gated by habituating transmitters. The transmitters immediately begin to deplete their transmitter levels in response to first-layer cell activity. Although the prepotent drive channel initially attains the highest activation, it also causes the most transmitter depletion. Competitive dynamics at third-layer cells then choose the second-layer cell of the prepotent channel, since it is the maximally activated channel.

(2) Third-layer cells initiate strong top-down on-center off-surround feedback to first-layer cells. The winning cell receives excitatory feedback, while the others are suppressed. The feedback is strong enough to maintain winning activity at third-layer cells, even after transmitter habituation occurs at the second-layer cell inputs. While the activity of the prepotent drive channel inactivates much of its transmitter, the other two channels, which are nearly inactive, have fully recovered transmitter levels.

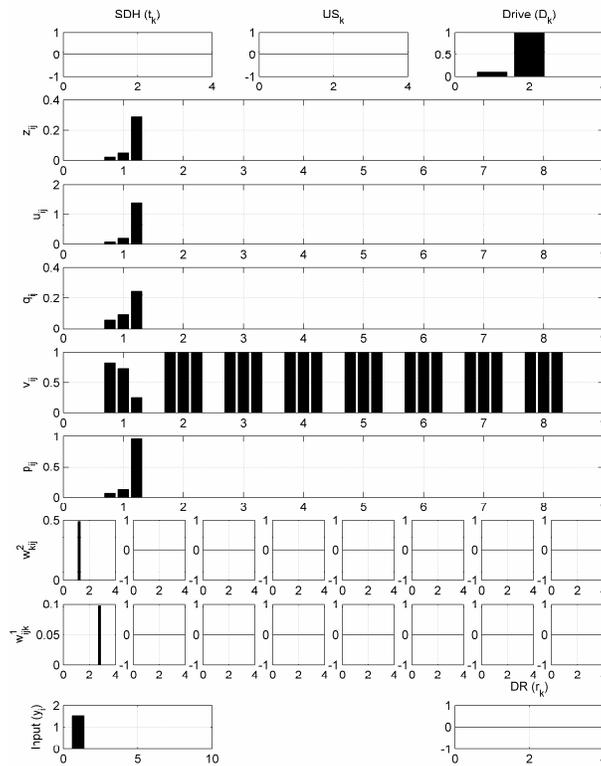
(3) The arrival of a burst of nonspecific arousal across all the channels can rebound the currently active drive channel. Such a burst occurs in response to an unexpected mismatch between learned top-down signals and the NET body displacement. The mismatch signals that a change in the currently active drive channel is needed to allow associations to be learned with a different motor response. When the arousal burst is sensed by the first-layer Gated Multipole cells, it causes a rapid increase in activity level at a time when the habituating gates of the second and third drive channels have nearly full transmitter levels. Thus, the corresponding second-layer cells respond strongly to the burst. As a result, third-layer cells choose the drive channel with least habituation. The third channel as the new winner. This happens because the third channel receives the smallest bottom-up signal, but the same top-down inhibition as the second drive channel. The second channel is slightly habituated due to a larger bottom-up input. As a result, the third channel attains the highest transmitter level. Under an arousal increment, second-layer cells in the third channel receive the largest increase to their input.

The third channel then triggers strong top-down on-center off-surround feedback to first-layer cells. The feedback strengthens the third drive channel, even though it has the smallest bottom-up signal, and inhibits the others. As the Gated Multipole activity equilibrates, the third drive channel maintains its activation level via this feedback loop, and can be associated with active drive representations and motor responses. Given the same working memory input and motivational state at a later time, this channel can be reactivated. Thus, this new cue-drive combination can selectively control its own response. If another arousal burst is applied, then the second drive channel can be activated, and thereby conditioned, using similar dynamics.

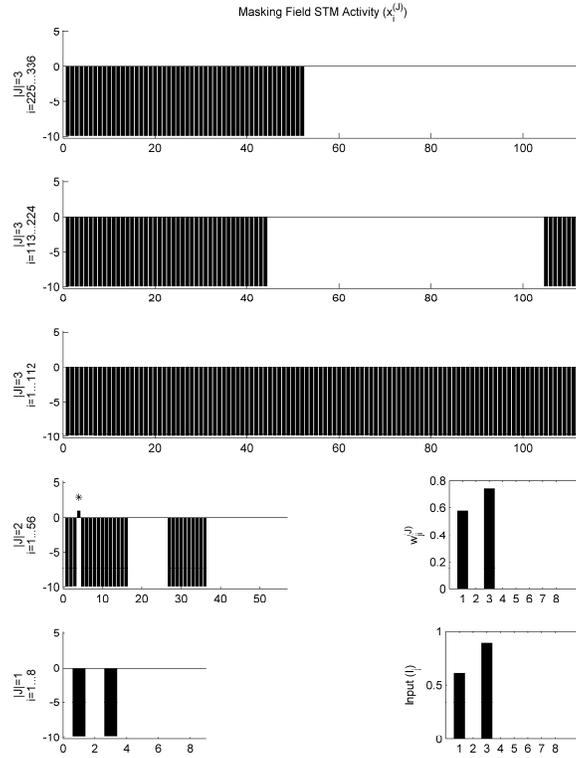
The Motivated WHAT and WHERE decisions (Figure 20b) are computed via a global winner-take-all process across all Gated Multipoles (Section A4.6, equation (66)). A single winning cell is selected, which can then read out (or learn) the activity of a Planned DV (see Figure 10). In the event of a mismatch between top-down readout and the NET body displacement, a nonspecific arousal signal to the first-layer cells can rebound all active Gated Multipoles, resulting in new Motivated WHAT and WHERE decisions and a different top-down readout (Section A8.1, equations (83)-(88)). Conditioning within these modules, at the appropriate time during exploration, leads to learning of a plan whose route is capable of reaching rewarded locations.



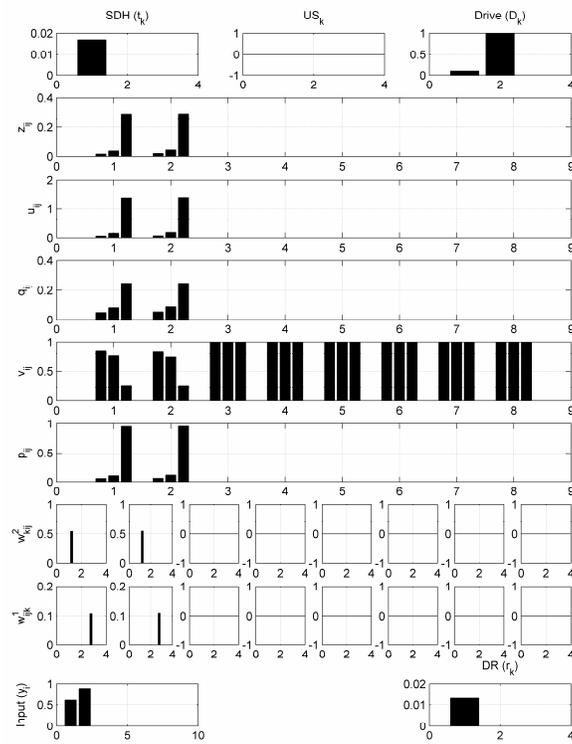
(a)



(b)



(c)



(d)

**Figure 21.** This series of diagrams depicts the response of several stages of the Visual Working Memory and Planning System under different conditions. **(a)** The 3D Object Working Memory response ( $I_i$ ) and Masking Field STM response ( $x_i^{(j)}$ ) when a single input, representing the triangle cue, is applied. Along the left side, the activity level for nodes with 1, 2 or 3 inputs are shown. In the lower right, the input and bottom-up ART weights ( $w_{ji}^{(j)}$ ) are shown. The most active node, representing the Triangle chunk, is shown by an asterisk (\*). In this case, the winning node receives one input, and inhibits the other nodes by masking field competitive dynamics. **(b)** The Object List Working Memory response and Gated Multipole STM response when a single input, representing the Triangle chunk, is applied. Each processing layer is labeled along the left side of the panel. The most active node appears at the output of the first Gated Multipole in the z field. After winner-take-all competition, top-down readout or learning may be activated in response to the Triangle chunk. **(c)** The 3D Object Working Memory response and Masking Field STM response when two inputs, representing the triangle and square cues, are applied. The most active node, representing the Triangle-Square chunk, is shown by an asterisk (\*). **(d)** The Object List Working Memory response and Gated Multipole STM response when two inputs, representing the Triangle and the Triangle-Square chunks are applied. The most active node appears at the output of the second Gated Multipole in the z field. After winner-take-all competition, top-down readout or learning may be activated in response to the Triangle-Square chunk.

The following sequence illustrates the flow of both spatial and motor information during a maze trial after learning. This trial is similar to the reactive exploratory trial summarized in Section 2.1. In that case, the animat learned the sequence Forward-Left-Forward to receive a reward when hungry. The animat now starts the maze along the centerline of the corridor, with its head facing the visual cue. Furthermore, the animat is again motivated by hunger, and will receive a food reward upon reaching the goal location. We will show how planned signals guide the animat toward the goal location and learned reinforcement signals strengthen the association between stored plan items and the current motivational state. The model initiates processing when a target in the visual field attracts the animat's attention, as follows:

Invariant preprocessing and learned ART categorization within the Visual Form and Motion System (Figure 5) encode a compressed representation of the triangle cue. The Visual Memory System (Figure 17) is updated by storing this category in the 3D Object Working Memory module and then computing the response of the remaining Visual Memory System modules. The Visual Memory System Masking Field module response is shown in Figure 21a. Furthermore, the Visual Memory System Gated Multipole module response is shown in Figure 21b. Finally, the NET body displacement in the Motor Approach and Orienting System (Figure 10) is reset by allowing the proprioceptive and vestibular and motor integrator activity to decay to zero (Figure 12). The head-centered azimuth and distance to the triangle cue are computed and converted to body-centered coordinates. The reactive movement command is relayed to the Motor Approach and Orienting System to begin reactive approach. Since the animat is facing the triangle cue, the Motor Approach and Orienting System computes a Motor Outflow command to initiate an approach movement toward the triangle cue.

The Sensory-Motor Top-down Readout module (Figure 20a) computes a Planned DV in response to the Motivated WHAT and WHERE Decisions from the Visual Memory System

(Figure 17). This planned command is relayed to the Motor Approach and Orienting System (Figure 10) and overrides the reactive movement command. The animat hereby approaches the triangle cue under planned control. During the approach movement, the Motor Approach and Orienting System updates the NET head and body displacement toward the visual target cue. The body-approach movement can continue until the NET approach signals zero the approach component of the Motor Outflow command and the animat reaches the triangle cue. Simultaneously, the Visual Form and Motion System (Figure 5) processes both form and motion signals which compete via learned interactions. A suitably strong motion signal in the periphery is unlikely to trigger a head-orienting movement away from the visual target cue since incentive signals from the Sensory-Drive Hierarchy module act to enhance attention on the visual target cue.

Once the approach movement is satisfied, the Motor Approach and Orienting System then computes a Motor Outflow command to initiate a planned orienting head turn to the left. The Motor Approach and Orienting System updates the NET body displacement during the head-orienting turn. The head-orienting movement continues until the NET head-orienting angle matches the planned angle and the square cue is centered in the visual field. Top-down signals from the Visual Working Memory and Planning System (Figure 17) corresponding to the Triangle chunk again learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System (Figure 10). The representation of the triangle cue is thereby associated with the exploratory drive and learns to predict the Forward-Left body movement. The Motor Memory System is updated by storing the Forward-Left body orienting and approach movement in the Motor Working Memory and Planning System (Figure 18). (The Motor Memory System Masking Field module and Gated Multipole module response is similar to the initial Visual Memory System response depicted in Figures 21a and 21b.)

Invariant preprocessing and learned ART categorization within the Visual Form and Motion System (Figure 5) encodes a compressed representation of the square cue. The Visual Memory System (Figure 17) is updated by storing this category in the 3D Object Working Memory module and then computing the response of the remaining Visual Memory System modules. The Visual Memory System Masking Field module response is shown in Figure 21c. Furthermore, the Visual Memory System Gated Multipole module response is shown in Figure 21d. Finally, the NET body displacement in the Motor Approach and Orienting System (Figure 10) is reset by allowing the proprioceptive and vestibular and motor integrator activity to decay to zero (Figure 12). Compute the head-centered azimuth and distance to the square cue and convert them to body-centered coordinates and relay this reactive movement command to the Motor Approach and Orienting System (Figure 10) to begin reactive approach. Since the animat is facing the square cue, the Motor Approach and Orienting System computes a Motor Outflow command to initiate an approach movement toward the square cue.

The Sensory-Motor Top-down Readout module (Figure 20a) computes a Planned Movement Command in response to the Motivated WHAT and WHERE Decisions from the Visual Memory System (Figure 17). This planned command is relayed to the Motor Approach and Orienting System and overrides the reactive movement command. The animat approaches the square cue under planned control. The animat's body is passively aligned with the head during an approach movement. During the approach movement, the Motor Approach and Orienting System updates the NET head and body displacement toward the visual target cue. The body-approach movement can continue until the NET approach signals zero the approach component of the Motor Outflow command and the animat reaches the square cue.

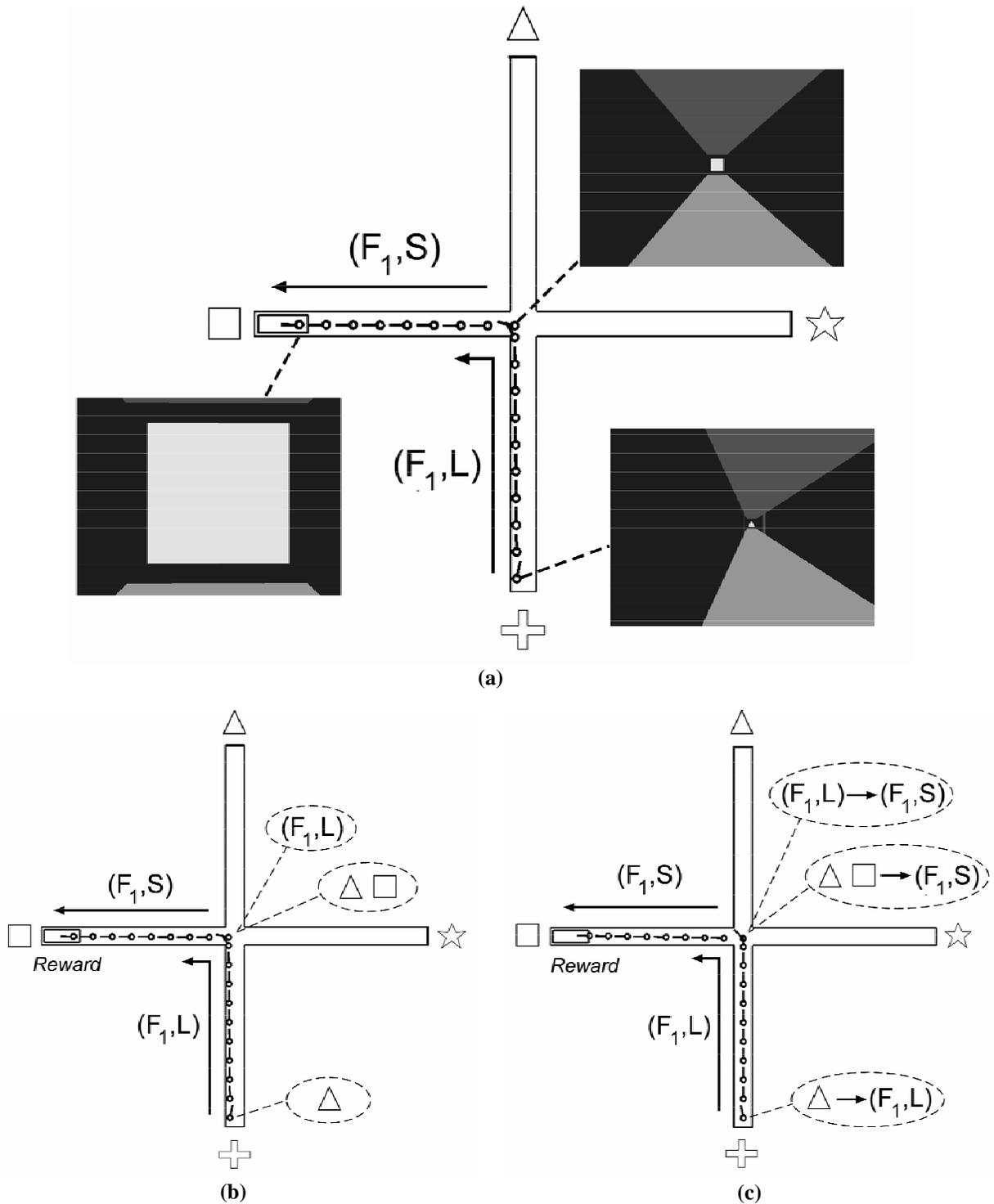
When the animat reaches the rewarded location, it receives a food reward. Plans stored in both the Visual and Motor Working Memory and Planning System are associated with the currently active hunger drive. Top-down signals from the Visual Working Memory and Planning System corresponding to the Triangle-Square chunk again learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System (Figure 10). The representation of the Triangle-Square chunk is thereby associated with the hunger drive and learns to predict the Forward-Straight body movement. Top-down signals from the Motor Working Memory and Planning System corresponding to the Forward-Left chunk learn the NET body orienting and approach movement computed by the Motor Approach and Orienting System (Figure 10). The representation of the Forward-Left chunk is thereby associated with the hunger drive and learns to predict the Forward-Straight body movement.

To summarize this maze trial, the animat is under planned guidance. A planned approach movement toward the triangle is followed by a head-orienting turn to the left. After approaching the square cue, the rewarded location is reached and adaptive weights are adjusted to strengthen the association between the currently active plan and the current motivational state. The Masking Field and Gated Multipole modules provide the dynamic plan representation which, coupled with learning at the appropriate time, leads to reinforcement of a successful planned response sequence.

### **3. End-to-end SOVEREIGN Simulations**

Three key SOVEREIGN simulations are demonstrated herein. In the Motivated Choice Experiment, the animat learns the route to two different goal locations under two different motivational states. In the Multiple Reward Experiment, the animat learns to repeatedly revisit two rewarded goal locations. Finally, in the Shortest Path Experiment, the animat learns the shortest path to the goal location, after repeated random exploration of the maze environment. Each of the simulation summaries contains the following types of information: (1) Explanation of the experimental setup; (2) movement trajectories; (3) a step-by-step description of model dynamics; (4) snapshots of visual input at key moments; (5) multi-trial learning; and (6) summary of the model properties demonstrated. The entire series of results is followed by a discussion of latent learning and how it might be explained by the SOVEREIGN model.

**3.1 Motivated Choice Experiment.** In this classic example of spatial learning, the animat learns the route to two different goal locations under two different motivational states. Specifically, the Forward-Left-Forward sequence when hungry leads to a food reward, whereas the Forward-Right-Forward sequence leads to a water reward when thirsty. Upon reaching the end of the goal arm, the animat is rewarded and long-term memory weights are updated using a slow learning rate. For the first five training trials of this sequence, the animat has a high hunger drive and is rewarded with food at the end of each trial. For the next five trials, the animat is thirsty and is rewarded with water at the end of each trial. The eleventh and final trial retests the response under the hunger drive to demonstrate that learning under each drive is preserved.

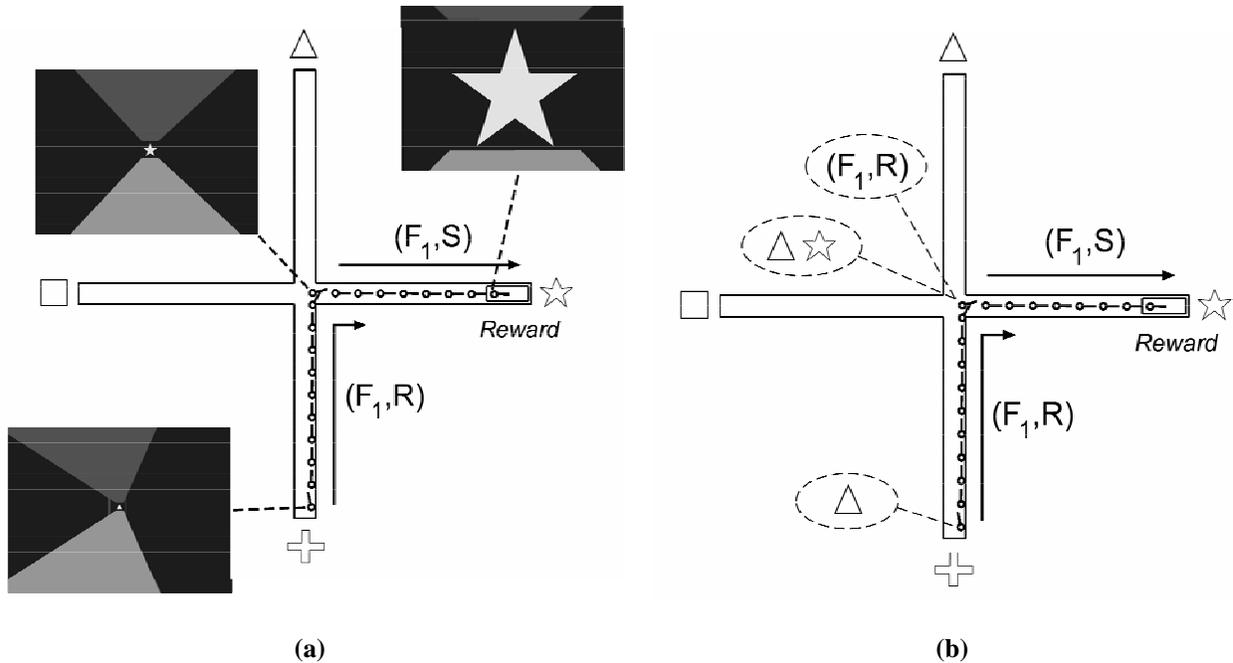


**Figure 22.** (a) Perspective views are shown for selected points during maze exploration toward the goal location in the left arm. (b) Each ellipse graphically depicts the short-term memory chunks represented in both the Visual and Motor Memory System during exploratory learning. (c) Each ellipse graphically depicts the

short-term memory chunks and associated motor responses in both the Visual and Motor Memory System after exploratory learning.

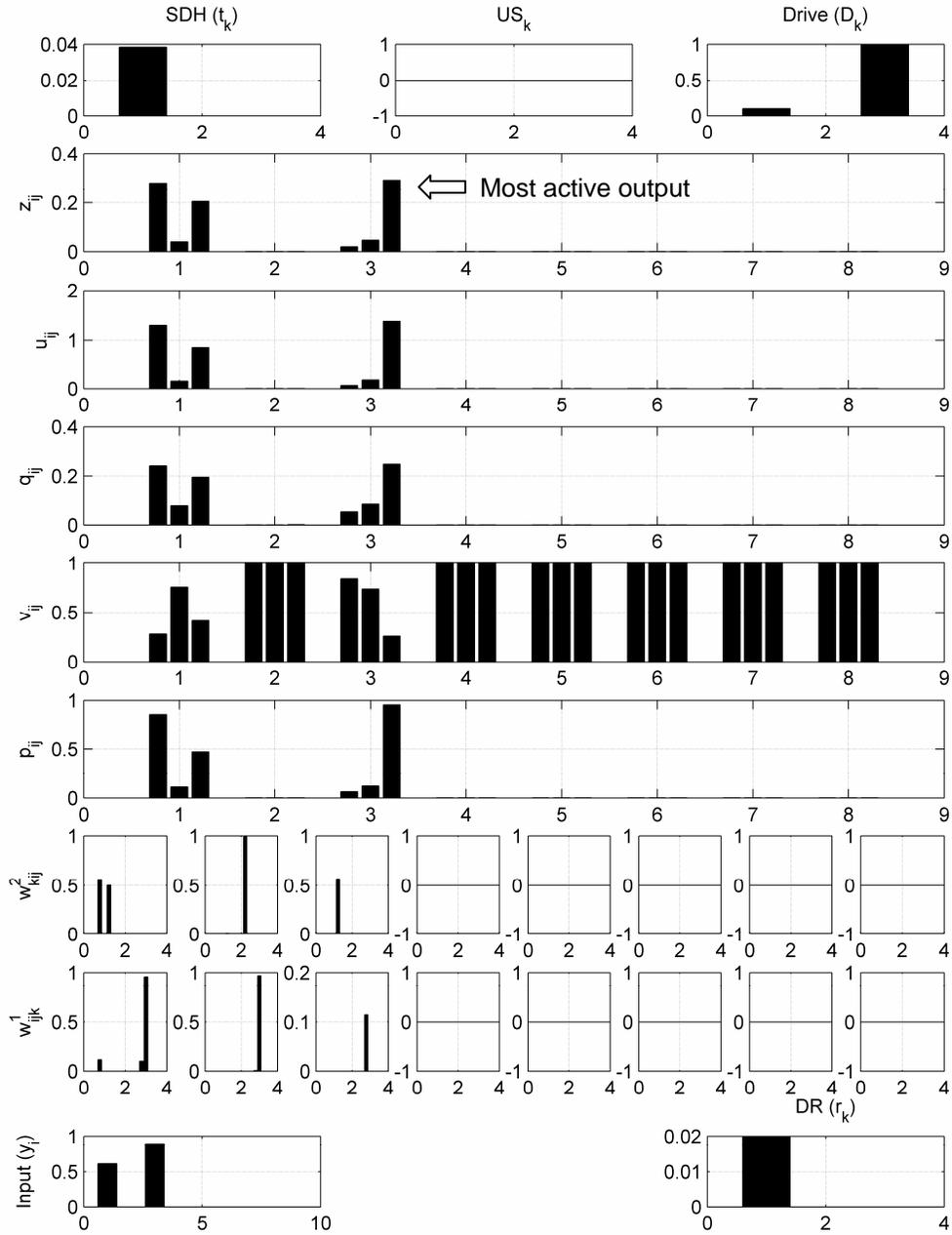
The diagram shown in Figure 22a shows the position, actions and local views seen by the animat during this training trial. Similarly, the diagram in Figure 22b shows the Visual and Motor Memory System plan chunks which are stored during this experimental trial. This simulation is similar to that presented in earlier examples. However, a summary is offered here for completeness. The animat starts this trial shifted to the left within the corridor, thereby increasing Magno signals in the left hemifield. The hungry animat learns to categorize the triangle cue and updates the Visual Memory System (Figure 17). It approaches the triangle cue under reactive control. As it nears the choice point, Magno signals trigger a head-orienting movement to the left bringing the square cue into view. The Triangle chunk is associated with the exploratory drive and can now sample the Forward-Left movement. The animat then learns to categorize the square cue. The Visual and Motor Memory System (Figures 17 and 18) are updated and the animat approaches the square cue under reactive control. Upon reaching the food reward, all active chunks are associated with the hunger drive and the Forward-Left and Triangle-Square chunks sample the Forward-Straight movement command. The initial training trial is complete.

The diagram in Figure 22c shows the learned plan chunks and their associated motor responses which are gradually strengthened during training. After several trials, the hungry animat starts this trial centered in the corridor, yielding Magno signals which are balanced between left and right. After learning to categorize the triangle cue and updating the Visual Memory System (Figure 17), the Triangle chunk can read out the command to go Forward-Left via top-down signals. The planned command overrides reactive signals and the animat moves forward and turns left. Both the approach speed and Parvo gain are increased because the previously rewarded plan element has been reactivated. After the turn is complete, the Triangle chunk is again associated with the exploratory drive. This learning is triggered by the exploratory learning signal in the absence of explicit reward, and is activated after a head turn is completed (Section A5). This model feature can explain the mechanism behind latent learning (Section 3.4). After learning to categorize the square cue, the Visual and Motor Memory System (Figures 17 and 18) are updated and the Forward-Straight command is directly read out via top-down Motivated WHAT and WHERE Decision signals. After approaching under planned control, the animat is rewarded with food and the active chunks are associated with the hunger drive. The Forward-Left and Triangle-Square chunks can sample the Forward-Straight movement command. The test trial is complete.



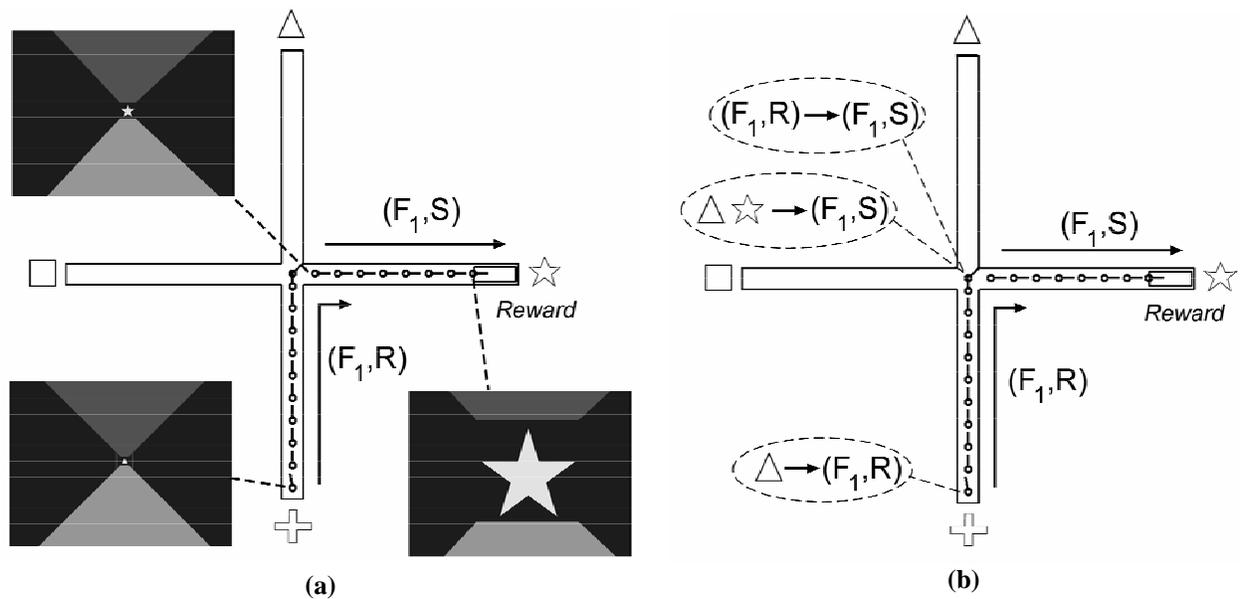
**Figure 23.** (a) Perspective views are shown for selected points of maze exploration during learning. (b) Each ellipse graphically depicts the short-term memory chunks represented in both the Visual and Motor Memory System during exploratory learning.

The diagram in Figure 23a shows the position, actions and local views seen by the animat during the second training trial. Similarly, the diagram in Figure 23b shows the Visual and Motor Memory System plan chunks which are stored during this experimental trial. In this training trial, the animat starts shifted to the right in the corridor, thereby increasing Magno signals in the right hemifield. The now-thirsty animat learns to categorize the triangle cue and then updates the Visual Memory System (Figure 17). The prepotent Forward-Left motor command is read out under the weakly active exploratory drive, but there is no Parvo gain or running speed increase. The animat moves toward the triangle cue under planned control. As it nears the choice point, Magno signals trigger a head-orienting movement to the right bringing the star cue into view. A top-down mismatch between the planned Forward-Left and NET Forward-Right movement triggers an arousal burst which selects an alternate channel in the Visual Memory System Gated Multipole module (Figure 17). The Triangle chunk is associated with the exploratory drive and can now sample the Forward-Right movement. The animat learns to categorize the star cue. The Visual and Motor Memory System (Figures 17 and 18) are updated and the animat approaches the star cue now under reactive control. The diagram in Figure 24 shows the Gated Multipole module response after the rebound and Visual Memory System update. Upon reaching the water reward, all active chunks are associated with the thirst drive and the Forward-Right and Triangle-Star chunks sample the Forward-Straight movement command. This training trial is complete.



**Figure 24.** The Working Memory and Gated Multipole short-term memory response when two inputs, representing the Triangle-Square chunk, are applied. After winner-take-all competition, top-down readout is activated in response to the Triangle-Square chunk, the most active node in the  $z_{ij}$  field, shown by the arrow. Top-down signals are processed by the Sensory-Motor Top-down Readout module. In the event that top-down signals don't match the actual NET movement, an arousal signal triggers a rebound in the Gated Multipole module. The rebound event allows a new channel to be selected for conditioning.

The diagram in Figure 25a shows the position, actions and local views seen by the animat during this experimental test trial. Similarly, the diagram in Figure 25b shows the Visual and Motor Memory System plan chunks which are stored during this experimental trial. After several learning trials, the thirsty animat starts this trial centered in the corridor, yielding Magno signals which are balanced between the left and right sides of the visual field. After learning to categorize the triangle cue and updating the Visual Memory System (Figure 17), the Triangle chunk is able to read out the command to go Forward-Right via top-down Motivated WHAT and WHERE Decision signals. The planned command overrides reactive signals and the animat moves forward and turns right. Both the approach speed and Parvo gain are increased because the previously rewarded plan element has been reactivated. After the turn is complete, the Triangle chunk is again associated with the exploratory drive. After learning to categorize the star cue, the Visual and Motor Memory System are updated and the Forward-Straight command is directly read out via top-down signals. After approaching under planned control, the animat is rewarded with water and the active chunks are associated with the thirst drive. The Forward-Right and Triangle-Star chunks can sample the Forward-Straight movement command. The test trial and this sequence of experiments are complete.



**Figure 25.** (a) Perspective views are shown for selected points during maze exploration under planned control. (b) Each ellipse graphically depicts the short-term memory chunks and associated motor responses in both the Visual and Motor Memory System after exploratory learning.

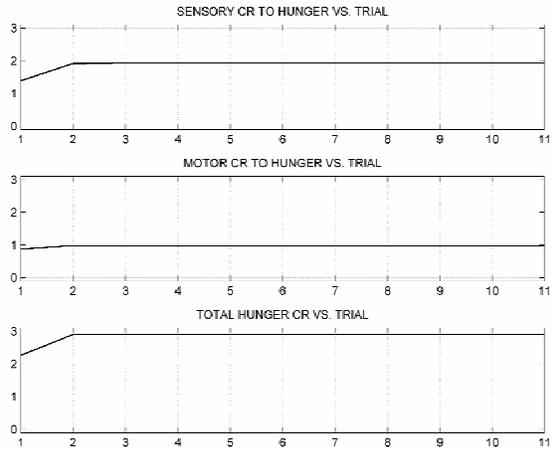
The next series of experiments demonstrate the acquisition of learned responses under different motivational states. Although the long-term memory (LTM) weights rapidly learn the appropriate responses, several trials are required for the weights to achieve asymptotic levels. The total level of conditioned reinforcement (CR) and incentive motivational learning is shown in the six panels of Figure 26. In the upper panel of Figure 26a, the CR adaptive weight from the Visual Memory System Gated Multipole module (Figure 17), denoted “SENSORY”, to the hunger Drive Representation module is plotted as a function of training trial. Recall that the first

five training trials comprise the Forward-Left-Forward movement sequence to receive a food reward while hungry. The next five training trials comprise the Forward-Right-Forward movement sequence to receive a water reward while thirsty. The eleventh, and final test trial demonstrates no loss of learning of the first sequence while hungry. In the middle panel of Figure 26a, the same CR data is shown for the Motor Memory System Gated Multipole module (Figure 18) to the hunger drive. In the bottom panel Figure 26a, the total learned CR versus training trial is shown. The CR levels increase and remain high as a consequence of doubly-gated learning (Section A4.5, equations (62)-(64)).

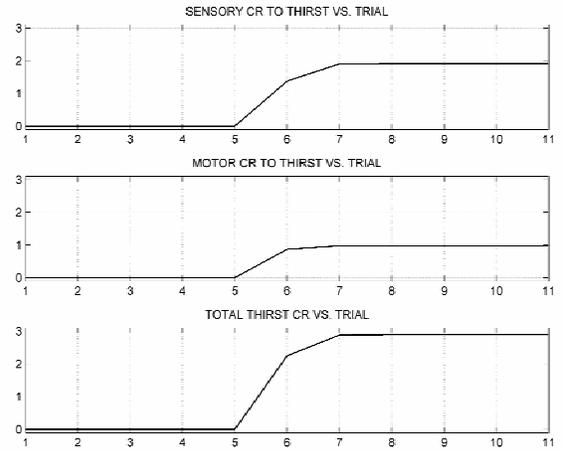
The learned CR to the thirst input of the Drive Representation module is similar to the hunger drive input, discussed above. In the upper panel, the learned CR from the Visual Memory System Gated Multipole module (Figure 17), denoted “SENSORY”, to the thirst input of the Drive Representation module is plotted as a function of training trial. In the middle panel, the same CR data is shown for the Motor Memory System Gated Multipole module (Figure 18) to the thirst drive. Finally, in the bottom panel, the total learned CR versus training trial is shown. The CR levels only increase after the fifth trial, when the thirst drive is active, and the animate receives a water reward for reaching the goal.

The exploratory drive performs an important role in spatial learning. The exploratory drive input to the Drive Representation module is inhibited by consummatory drive activity. However, it is never completely inactive, thereby allowing some conditioned reinforcement (CR) and incentive motivational learning to it. The exploratory drive input and its affect on the Drive Representation module is discussed in Section A5. In Figure 26c, upper panel, the learned CR from the Visual Memory System Gated Multipole module (Figure 17), denoted “SENSORY”, to the exploratory input of the Drive Representation module is plotted as a function of training trial. In the middle panel the same CR data is shown for the Motor Memory System Gated Multipole module (Figure 18) to the exploratory drive. Finally, in the bottom panel, the total learned CR versus training trial is shown. During the first five trials, the Star chunk and Forward-Left motor sequence were conditioned to the exploratory drive. This resulted in low CR levels for this first phase of training and a rapid increase in the CR levels after the fifth trial. The effect seen is that, during reactivation of the Star chunk, learned CR contributes to an increase in the Drive Representation module activity. This increase, in turn, is sampled by the Visual and Motor Memory System Gated Multipole modules.

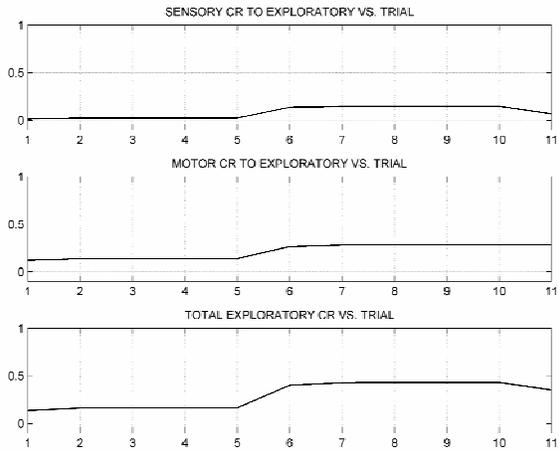
Incentive motivational signals aid in the selection of Visual and Motor Memory System Gated Multipole module channels. These signals are important for enabling a chunk to read out the correct motor command under different motivational states. The Gated Multipole module dynamics are discussed in Section A4.5. The total learned incentive motivation from the exploratory channel of the Sensory-Drive Heterarchy module is shown as a function of trial in Figure 26d. The upper, middle and lower panels are as discussed above. The acquisition of learned incentive motivation to the hunger and thirst drives is shown in Figures 26e and 26f. The upper, middle and lower panels are as above. The initially large incentive motivation signals under the hunger drive decrease later in the experiment when learning under the thirst drive. This results from the ongoing exploratory learning to the prepotent Gated Multipole channel which weakens its prior conditioning to the hunger drive. In summary, this experimental setup demonstrates the capacity for the SOVEREIGN model to learn different routes to a goal location under different motivational states.



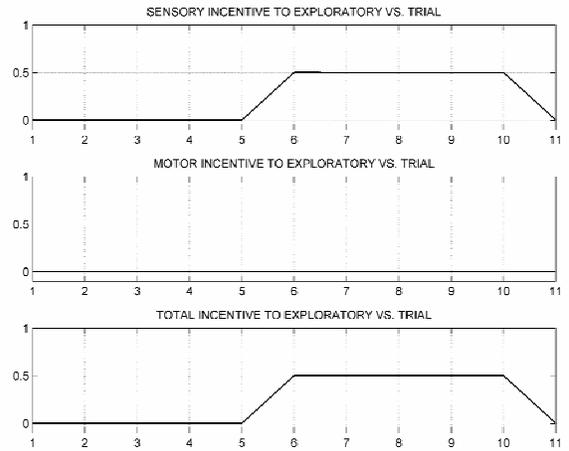
**(a)**



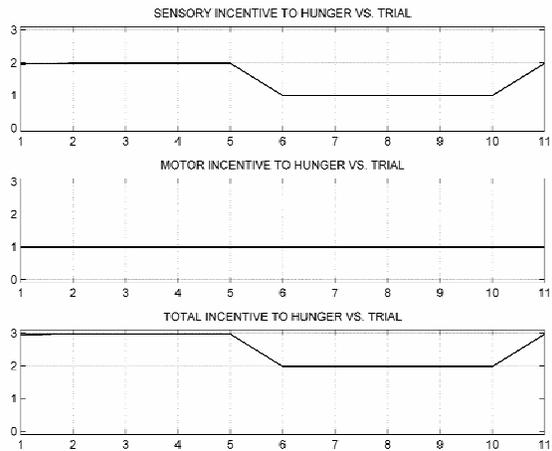
**(b)**



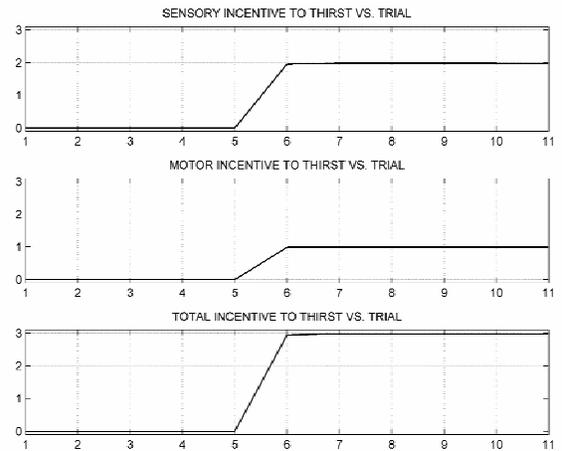
**(c)**



**(d)**



**(e)**



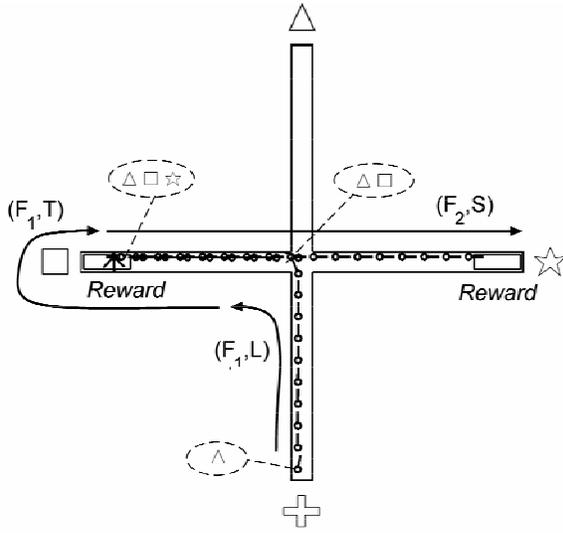
**(f)**

**Figure 26.** This series of plots show the acquisition of learned conditioned reinforcement and incentive motivation over a series of 11 experimental trials navigating to goal locations under either the hunger or thirst drives (*see text for details*). In each panel, the Visual Memory System (sensory), Motor Memory System (motor) and total learned associations are plotted vs. trial. **(a)** Acquisition of sensory and motor conditioned reinforcement to the hunger drive vs. trial. **(b)** Acquisition of sensory and motor conditioned reinforcement to the thirst drive vs. trial. **(c)** Acquisition of sensory and motor conditioned reinforcement to the exploratory drive vs. trial. **(d)** Acquisition of sensory and motor incentive motivation from the Sensory-Drive Heterarchy module exploratory channel vs. trial. **(e)** Acquisition of sensory and motor incentive motivation from the Sensory-Drive Heterarchy module hunger channel vs. trial. **(f)** Acquisition of vision and motor incentive motivation from the Sensory-Drive Heterarchy module thirst channel vs. trial.

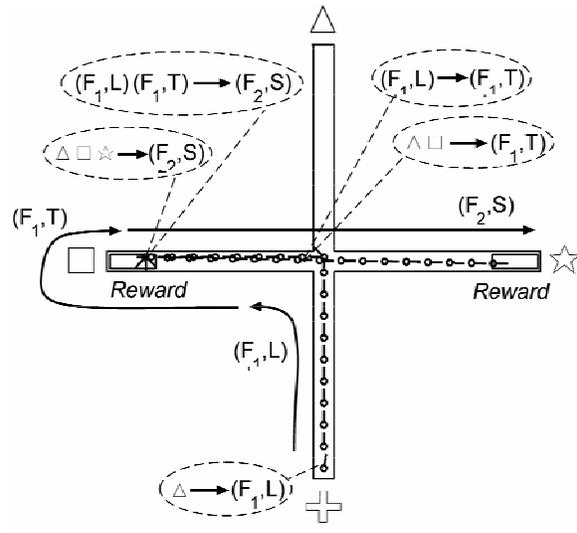
**3.2 Multiple Reward Experiment.** The second series of simulations demonstrate the ability to learn to visit multiple reward sites under a given motivational state. Shuttling between the rewarded locations continues until the Motor List Working Memory module capacity is reached. At that point, alternating high-level chunks can read out the appropriate motor response to continue. Results are presented in Figure 27 and discussed below.

In this trial, the animat visits multiple reward sites under reactive control. The layout of the experiment as well as the motor movements and Visual Memory System chunks are shown in Figure 27a. After visiting the second rewarded location, the animat should turn 180° and continue the shuttling behavior between the first and second locations to receive continued reward. The animat starts shifted to the left within the corridor, thereby increasing Magno signals in the left hemifield. The hungry animat learns to categorize the triangle cue and updates the Visual Memory System (Figure 17). It moves toward the triangle cue under reactive control. As it nears the choice point, Magno signals trigger a head-orienting movement to the left bringing the square cue into view. The Triangle chunk is associated with the exploratory drive and can now sample the Forward-Left movement. The animat learns to categorize the square cue. The Visual and Motor Memory System (Figures 17 and 18) are updated and the animat approaches the square cue under reactive control. Upon reaching the food reward, all active chunks are associated with the hunger drive.

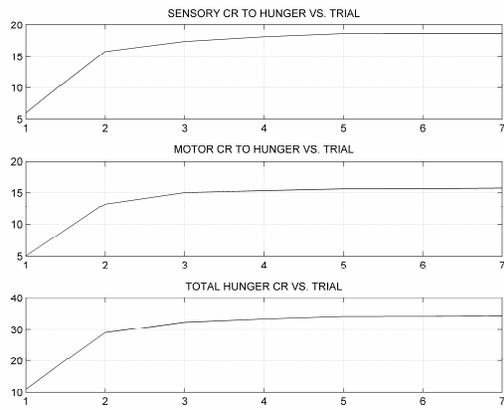
After a short time, the reward is removed and the ERG is disinhibited to trigger a random head-orienting movement. The ERG head turn brings the star cue into view. The Forward-Left and Triangle-Square chunks are associated with the exploratory drive and can both sample the Forward-Turnaround movement command. After learning to categorize the star cue, the Visual and Motor Memory System are updated. The animat approaches the star cue under reactive control. Incentive signals from the Sensory-Drive Heterarchy module enable an increase in running speed and attentional gain. The increase in attentional gain precludes a Magno attentional shift, even at the increased running speed. Upon reaching the second food reward, all active chunks are associated with the hunger drive. The Forward-Left, Forward-Turnaround chunk and Triangle-Square-Star chunk sample the Forward<sub>2</sub>-Straight movement command. The subscript, 2, denotes an approach distance of twice the arm length. Soon thereafter, the Forward<sub>2</sub>-Straight movement command is again selected under reactive control.



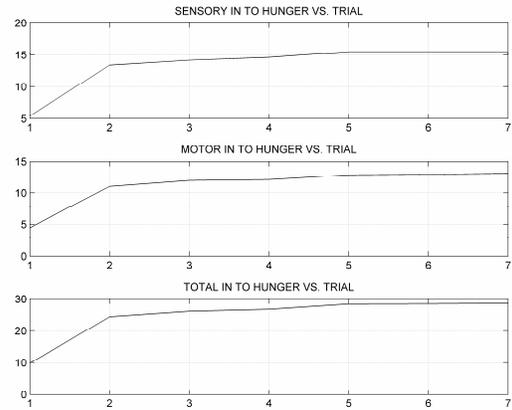
(a)



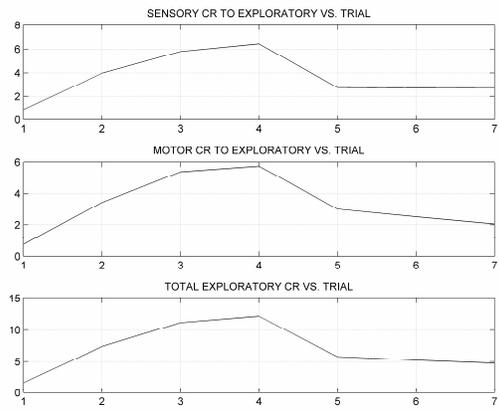
(b)



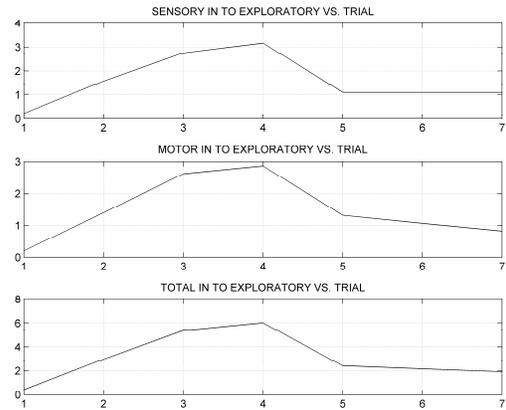
(c)



(d)



(e)



(f)

**Figure 27.** This series of figures show the key results after repeated shuttling to rewarded locations (*see text for details*). A symbolic representation of visual cues and motor commands during maze traversal is shown in the upper panels. This is followed by detailed plots of the acquisition of learned conditioned reinforcement and incentive motivation. In these lower panels, the Visual Memory System (sensory), Motor Memory System (motor) and total learned associations are plotted vs. trial. **(a)** Each ellipse graphically depicts the short-term memory chunks represented in the Visual Memory System along with the route during. **(b)** Each ellipse graphically depicts the short-term memory chunks and associated motor responses in both the Visual and Motor Memory System after exploratory learning. **(c)** Acquisition of sensory and motor conditioned reinforcement to the hunger drive vs. trial. **(d)** Acquisition of sensory and motor incentive motivation from the Sensory-Drive Heterarchy module hunger channel vs. trial. **(e)** Acquisition of sensory and motor conditioned reinforcement to the exploratory drive vs. trial. **(f)** Acquisition of sensory and motor incentive motivation from the Sensory-Drive Heterarchy module exploratory channel vs. trial.

In the second phase of learning, the animat visits multiple reward sites under planned control. The layout of the experiment as well as the motor movements and Visual Memory System chunks are shown in Figure 27b. The animat starts centered in the corridor, thereby balancing Magno signals in the left and right hemifields. The hungry animat learns to categorize the triangle cue and updates the Visual Memory System (Figure 17). The Triangle chunk can then read out the Forward-Left movement command via top-down signals. The animat approaches the triangle cue and turns left under planned control. The Triangle chunk is associated with the exploratory drive and can now sample the Forward-Left movement. The animat learns to categorize the square cue. The Visual and Motor Memory System (Figures 17 and 18) are updated and can read out the Forward-Turnaround command via learned top-down signals from the Motivated WHAT and WHERE Decisions. The animat thereby approaches the square cue under planned control.

Upon reaching the food reward, all active chunks are associated with the hunger drive. After a short time, the reward is removed and the learned Turnaround command triggers a head-orienting movement. The head turn brings the star cue into view. The Forward-Left and Triangle-Square chunks are associated with the exploratory drive and can both sample the Forward-Turnaround movement command. After learning to categorize the star cue, the Visual and Motor Memory System are updated. The Forward-Left, Forward-Turnaround and Triangle-Square-Star chunks read out the Forward<sub>2</sub>-Straight movement commands. The animat approaches the star cue under planned control. Learned incentive signals from the Sensory-Drive Heterarchy module enable an increase in running speed and attentional gain.

Upon reaching the second food reward, all active chunks are associated with the hunger drive. The Forward-Left, Forward-Turnaround chunk and Triangle-Square-Star chunk can sample the Forward<sub>2</sub>-Straight movement command. Soon thereafter, the Forward<sub>2</sub>-Straight movement command is again selected under reactive control. Eventually, an alternating combination of these chunks can take over and enable continuous shuttling for reward. The acquisition of learned conditioned reinforcement and incentive motivation to the hunger drive is shown in Figures 27c and 27d. The upper, middle and lower panels follow the convention

discussed above. Both conditioned reinforcement and incentive motivation increase during the experiment while learning under the hunger drive.

The acquisition of learned conditioned reinforcement and incentive motivation to the exploratory drive is shown in Figures 27e and 27f. Conditioned reinforcement and incentive motivation weights initially increase under the exploratory drive, and then decrease later in the experiment as learning to the hunger drive reaches asymptotic levels. In summary, this experimental setup demonstrates the capacity for the SOVEREIGN model to learn to repeatedly revisit goal locations while motivated under a constant drive state. As new shuttling continues, an alternating pair of chunks read out motor responses to receive reward.

**3.3 Shortest Path Experiment.** In the final series of simulations, the ability to randomly explore the maze from several starting locations and learn the shortest path to the goal is demonstrated. The experiment proceeds in three phases shown in Figures 28a-c and discussed below.

In the initial trial, the animat randomly explores the maze under reactive control with high exploratory and hunger drive inputs. The layout of the experiment as well as the motor movements and Visual Memory System chunks are shown in Figure 28a. Exploratory movements continue until reward is received. The animat starts shifted to the left within the corridor, thereby increasing Magno signals in the left hemifield. The hungry animat learns to categorize the triangle cue and updates the Visual Memory System (Figure 17). It approaches the triangle cue under reactive control. As it nears the choice point, Magno signals trigger a head-orienting movement to the left bringing the square cue into view. The Triangle chunk is associated with the exploratory drive and can now sample the Forward-Left movement. The animat learns to categorize the square cue. The Visual and Motor Memory System (Figures 17 and 18) are updated and the animat approaches the square cue under reactive control.

After a short time, the ERG is disinhibited to trigger a random head-orienting movement. The ERG head turn brings the star cue into view. The Forward-Left and Triangle-Square chunks are associated with the exploratory drive and can both sample the Forward-Turnaround movement command. After learning to categorize the star cue, the Visual and Motor Memory System are updated. The animat approaches the star cue under reactive control. As it nears the choice point, Magno signals trigger a head-orienting movement to the left bringing the triangle cue into view. The Forward-Left, Forward-Turnaround and Triangle-Square-Star chunks are associated with the exploratory drive and can both sample the Forward-Left movement command.

After learning to categorize the triangle cue, the Visual and Motor Memory System are updated. The animat approaches the triangle cue under reactive control. After a short time, the ERG is disinhibited to trigger a random head-orienting movement. The ERG head turn brings the plus cue into view. The Forward-Left, Forward-Turnaround, and Forward-Left motor category working memory items comprise the motor chunk. The Square-Star-Triangle object list working memory item comprises the sensory chunk. Both of these chunks are associated with the exploratory drive and can sample the Forward-Turnaround movement command. After learning to categorize the plus cue, the Visual and Motor Memory System are updated and the animat approaches the plus cue under reactive control. By chance, the Magno signals in the left and right hemifields are nearly balanced, thus no head-orienting movement is triggered as the animat nears the choice point. Upon reaching the plus cue, the animat stops.

After a short time, the ERG is disinhibited to trigger a random head-orienting movement. The ERG head turn brings the triangle cue into view. The Forward-Turnaround, Forward-Left, Forward-Turnaround and Star-Triangle-Plus chunks are associated with the exploratory drive

and can sample the Forward<sub>2</sub>-Turnaround movement command. Once the animat learns to categorize the triangle cue, the Visual and Motor Memory System are updated. The animat approaches the triangle cue under reactive control. As it nears the choice point, Magno signals trigger a head-orienting movement to the right bringing the star cue into view. The Forward-Left, Forward-Turnaround, Forward<sub>2</sub>-Turnaround and Triangle-Plus-Triangle chunks are associated with the exploratory drive and can sample the Forward-Right movement command. After learning to categorize the star cue, the Visual and Motor Memory System are updated and the animat approaches the star cue under reactive control. Upon reaching the food reward, all active chunks are associated with the hunger drive. At this point, there is a transfer of incentive motivational and conditioned reinforcer learning from the exploratory drive to the consummatory hunger drive. The Forward-Turnaround, Forward<sub>2</sub>-Turnaround, Forward-Right and Plus-Triangle-Star chunks can sample the Forward-Straight movement command. This completes the initial exploratory training trial.

In the second training trial, the animat starts shifted to the right in the corridor, thereby increasing Magno signals in the right hemifield. The layout of the experiment as well as the motor movements and Visual Memory System chunks are shown in Figure 28b. The hungry animat learns to categorize the triangle cue and updates the Visual Memory System (Figure 17). The prepotent Forward-Left motor command is read out under the weakly active exploratory drive. There is no Parvo gain or running speed increase since learned conditioned reinforcer inputs to the hunger and exploratory drives are insufficient to activate Sensory-Drive Heterarchy output (equations (67)-(69), (104)). The animat moves toward the triangle cue under planned control. As it nears the choice point, Magno signals trigger a head-orienting movement to the right, thereby bringing the star cue into view. A top-down mismatch between Forward-Left and Forward-Right triggers an arousal burst, which selects an alternate channel in the Visual Memory System Gated Multipoles (Figure 20b; Section 2.6.5, A4.5). The Triangle chunk is associated with the exploratory drive and can now sample the Forward-Right movement. The animat also learns to categorize the star cue. The Visual and Motor Memory System are updated and the animat approaches the star cue under reactive control. Upon reaching the food reward, all active chunks are associated with the hunger drive and the Forward-Right and Triangle-Star chunks sample the Forward-Straight movement command.

After several learning trials under a slow learning constraint, each starting from a random left or right corridor shift, the hungry animat starts this trial centered in the corridor, yielding Magno signals which are balanced between left and right. The layout of the experiment as well as the motor movements and Visual Memory System chunks are shown in Figure 28c. After learning to categorize the triangle cue and updating the Visual Memory System (Figure 17), the Triangle chunk is able to directly reactivate the alternate Gated Multipole channel, conditioned after the rebound, in the prior trial. This direct access to the alternate channel enables the top-down read out of the command to go Forward-Right. The planned command overrides reactive signals and the animat moves forward and then turns right. Both the approach speed and Parvo gain is increased because the previously rewarded plan element has been reactivated. After the turn is complete, the Triangle chunk is again associated with the exploratory drive. After learning to categorize the star cue, the Visual and Motor Memory System (Figures 17 and 18) are updated and the Forward-Straight command is directly read out via top-down signals. After approaching under planned control, the animat is rewarded with food and the active chunks are associated with the hunger drive. The Forward-Right and Triangle-Star chunks can sample the Forward-Straight movement command.



the short-term memory chunks represented in the Visual Memory System along with the route during exploratory learning while shifted to the right of the corridor. (c) Each ellipse graphically depicts the short-term memory chunks and associated motor responses in both the Visual and Motor Memory System after exploratory learning.

In this shortest path experiment, the SOVEREIGN model can demonstrate how repeated experience with the environment can selectively amplify the shortest path chunk, leading to plan optimization under a specific motivational state.

**3.4 Latent Learning.** The SOVEREIGN model explains aspects of latent learning by using its endogenously active exploratory drive (ERG). Three types of latent learning experiments have been tested using rats (Munn, 1950): (1) introduction of reward after non-rewarded training; (2) non-rewarded maze exploration followed by learning of the maze; and (3) Incidental learning of cues while being trained under the condition of another drive or while satiated.

The introduction of reward after non-rewarded training is the least contested of these paradigms. In such an experiment, rats are allowed to run a maze for several trials before a food reward is given. Thereafter, a food reward is given on each trial. The results show that there is little learning before food was introduced, followed by a rapid decrease in errors (Blodgett, 1929; Tolman and Honzik, 1930). Both groups of experimenters argued that the decrease can be explained by latent learning of the movements required to reach goal locations in the maze apparatus, which was not exploited until after reward was introduced. A similar experiment failed to reproduce these findings, however (Reynolds, 1945).

Non-rewarded maze exploration followed by learning of the maze is a second form of latent learning. It was first reported by Haney (1931), who trained two groups of rats on a maze after one group had explored the maze for eighteen hours on each of four days. In both errors and time the exploration group was superior (Haney, 1931; Daub, 1933). In order to reduce the amount of non-experimental reinforcement, a variation of this training was introduced (Buxton, 1940). The modified approach allowed the test rats to live in a multiple-T maze for a varying number of nights, while changing their entrance and exit points to the maze. During testing, each rat was taken to the food box to eat, and then placed at the entrance for each run. The errors for the experimental groups were much lower than for the controls. Thus, latent learning was demonstrated under reduced non-experimental reinforcement.

Incidental learning of cues while satiated or under a different drive condition is the most controversial of the three paradigms. It was successfully demonstrated by Leeper (1935) who trained rats to go to one place in a maze if hungry and to another place if thirsty. If they made an incorrect choice, the rats were allowed to backtrack to the correct location. The results were much stronger than in a similar study by Hull (1933) which did not allow backtracking, and so lacked immediate reinforcement on every trial. In contrast, there are several studies in which negative results were found (Spence and Lippett, 1946; Kendler, 1946, 1947a, 1947b; Grice, 1948 and Walker, 1948). This form of latent learning, while weak, has been experimentally demonstrated in a simple case.

These results are consistent with model properties. The proposed exploratory drive is one of the nominal drive inputs. The exploratory drive is inhibited by internal hunger and thirst drives and, though not explicitly modeled, assumed to be excited by an internal, habituated excitatory source. Such an excitatory source is assumed to be activated by novelty, or a general arousal level upon entering a new environment. This effect is emulated by directly manipulating the drive inputs to the model. Even if the animat is satiated for food and water, the exploratory

drive can excite the Motor Approach and Orienting System (Figure 10) to generate movements. If learning is allowed to occur with the exploratory drive as with the other drive representations, then a nonspecific motivational drive is created. For example, during exploratory, satiated runs, a small amount of learning to the exploratory drive can later be reactivated by a hunger or thirst state. This occurs since exploratory learning takes place after each orienting movement is complete. Sequences of movements are always conditioned to the exploratory drive. Only when reward is received, is the learning transferred to the consummatory drive state. This is sufficient to explain the first and second types of latent learning.

Finally, the third paradigm's data may be clarified by the fact that US inputs to the Drive Representation act as sampling signals in the model. Thus, if water is encountered during a run, the chunks of that plan resonate with the thirst Drive Representation. Reciprocal learning between the Gated Multipoles and the Drive Representation reinforces the currently active chunks. Active chunks gain conditioned reinforcer properties (equation (62)) as well as incentive motivational properties (equation (65)). At a later time, these chunks can influence the choice of motivational state by the Sensory-Drive Hierarchy, and the overall choice of chunk capable of top-down read out. This can be accomplished under either exploratory or consummatory drive states.

#### **4. General Discussion and Conclusions**

The SOVEREIGN architecture embodies a number of design principles whose mechanistic instantiation as neural circuits (Figure 2) enable incremental learning of planned action sequences to carry out route-based navigation towards a rewarded goal. SOVEREIGN circuits are based on neural models that have elsewhere been used to explain and predict many behavioral and brain data. Here the emphasis is on designing a neuromorphic controller that emphasizes behavioral competence.

The model has several notable strengths relative to other available models, including the following ones: First, it provides an end-to-end model that includes on-line vision, visual recognition learning and categorization, working memory storage of sequences of visual and motor categories, learning of sequential cognitive and motor plans, cognitive-emotional interactions whereby reinforcement learning can select plans that can attain a currently valued goal, and balancing of visually reactive exploratory vs planned movement decisions, based upon the relative salience of bottom-up and top-down information through time. Second, unlike various other models (e.g., Barto and Sutton, 1981; Dayan, 1987; Schmajuk, 1990), no explicit spatial goal gradient, proportional to the spatial distance from the goal, is required to guide goal-oriented sequential behavior in SOVEREIGN. Third, list chunks provide a compact context-sensitive code for learning plans to navigate a large number of different routes. Fourth, reliable, single-trial learning of a maze can occur if the animat happens to find the goal location during an exploratory trial. Fifth, the animat can respond to the same sequence of visual or motor events in different ways to achieve different goals when different drives are prepotent.

Whereas the detailed circuit realizations that are currently used in SOVEREIGN will doubtless be modified and further developed in the future, it embodies design principles that may need to be incorporated, in some form, into future autonomous adaptive controllers of navigational behaviors. One weakness in the current version of SOVEREIGN is that its navigational behaviors are all route-based. The model does not yet include mechanisms of spatial navigation (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978) such as the role of hippocampal place fields, head-direction cells, and the theta rhythm (e.g., Burgess et al., 1995). Such a development would require an understanding of how place fields form, notably the role of

entorhinal grid cells in their formation (e.g., Hafting, Fyhn, Molden, Moser, and Moster, 2005), which other modeling research is currently investigating (e.g., Fuhs and Touretzky, 2006; Gorchetchnikov and Grossberg, 2007).

## Appendix. Model Equations

SOVEREIGN is defined by interacting neurons, or neuron populations, whose activities, or membrane potentials,  $V(t)$ , obey a membrane, or shunting, equation:

$$C_m \frac{dV(t)}{dt} = -[V(t) - E_{leak}] \gamma_{leak} - [V(t) - E_{excit}] \gamma_{excit}(t) - [V(t) - E_{inhib}] \gamma_{inhib}(t) \quad (1)$$

(Grossberg, 1973; Hodgkin, 1964). In (1),  $C_m$  is the cell membrane electrical capacitance,  $\gamma_{leak}$  is the leakage conductance neuron, and  $\gamma_{excit}(t)$  and  $\gamma_{inhib}(t)$  are the time-varying excitatory and inhibitory conductances, respectively. In (1), the  $E$  terms represent reversal potentials that introduce shunting properties into cell responses. The excitatory and inhibitory conductance terms scale with the number, type and strength of neural synapses impinging on each cell. Without loss of generality, the reversal potentials are assigned as follows:  $E_{leak} = 0$ ,  $E_{excit} = B$  and  $E_{inhib} = -E$ . The cell membrane capacitance,  $C_m$ , and leakage conductance,  $\gamma_{leak}$ , are combined with other terms. The resulting equation defines the response of a neuron with shunting response:

$$\frac{dV}{dt} = -AV + [B - V] \gamma_{excit}(t) - [V + E] \gamma_{inhib}(t). \quad (2)$$

Reading from left to right, equation (2) includes a leakage or decay term, a shunting excitatory term, and a shunting inhibitory term. At steady-state,  $\frac{dV}{dt} = 0$ , and the response of the neuron is:

$$V = \frac{B\gamma_{excit} - E\gamma_{inhib}}{A + \gamma_{excit} + \gamma_{inhib}}. \quad (3)$$

Similarly, the potential in (2) approximates an additive response,

$$\frac{dV}{dt} = -AV + B\gamma_{excit}(t) - E\gamma_{inhib}(t) \quad (4)$$

when  $A \gg \gamma_{excit}$  and  $A \gg \gamma_{inhib}$ . In this regime, the cell responds linearly to excitatory and inhibitory inputs. Both shunting and additive equations are employed herein.

Model equations were simulated using the Matlab simulation environment and: (1) Runge-Kutta 4<sup>th</sup> order methods, (2) iterating the steady-state response, or (3) direct evaluation at steady-state. The method chosen depended upon the need for the exact system response and computational simplicity.

Concerning mathematical notation: each variable name is a lower-case letter, whereas constants and control signals are upper-case letters. A constant may have a subscript to denote various values in different contexts. Similarly, a variable may have subscripts to define the endpoints of its connections. For example,  $w_{ij}$ , is a variable representing the adaptive weight which makes connections between the  $i^{\text{th}}$  and  $j^{\text{th}}$  cells. Superscripts define the group to which a variable belongs. For example, variable,  $c_{ij}^d$ , represents the cell membrane potential with direction,  $d$  ( $d = L$  for left,  $R$  for right), which makes connections between the  $i^{\text{th}}$  and  $j^{\text{th}}$  cells.

### A1. Parvo processing

The Visual Form and Motion System (Figure 5) computes all visual input signals to the SOVEREIGN model. Each 3D view of the environment is generated in color by an OpenGL-based graphical simulator, implemented within the Matlab simulation environment. In the Matlab graphical simulator, a camera is defined which can be moved through a 3D virtual

environment, namely the plus-maze. Each component of the maze is defined by specifying the 3D coordinates of the vertices of a closed polygon. For example, the walls and floor comprise rectangles specified by the 3D coordinates of each vertex. Each visual target object, such as the triangle, square, star or plus is specified similarly. Furthermore, distinctive colors are chosen for each environmental component: floor, walls, choice point edges, and visual target objects. Since a flat shading model is chosen, a unique color, or grayscale value, corresponds to each target object. The lack of shading, texture or a sophisticated lighting model does not detract from the task of spatial learning. It merely simplifies the problem of locating and identifying visual landmarks in the environment. Such problems have been addressed by other lines of research.

The Matlab simulation environment and OpenGL graphics library define a 3D Cartesian coordinate system to describe locations within the virtual world. For convenience, each maze arm is defined to be 10,000 units long in world coordinates. The walls are 1000 units high and corridors are 1000 units wide. Thus, the viewing distance from animat to 3D visual target object is maximal when the animat is at the end of one arm and is looking at the visual target object at the far end of the corridor. In world coordinates, this distance equals the length of two arms ( $2 \times 10,000 = 20,000$ ) plus the width of the intervening corridor (1000) at the choice point. This sum is defined by the constant,  $D_{MAX} = 21,000$ , which is used to normalize distance inputs to the Spatial Attention module; see equation (30).

The camera resolution is 384 x 288 pixels, which is a tradeoff between computational throughput and accurate pattern recognition at long-range. Thus, the aspect ratio, or horizontal resolution/vertical resolution, is 4/3. The camera has a fixed field-of-view of 45 degrees horizontally. The vertical field-of-view is scaled by the aspect ratio and is therefore 37.5 degrees. The dimensions of the 3D environment are selected such that only one target object is visible at a time. This simplification again eases the visual processing task, without detracting from the model's ability to demonstrate spatial navigation. Although the camera rotates and translates along the X and Y axes, the camera's vertical position is fixed on the Z-axis at 500 units. This results in a constant height during maze traversal.

The animat makes discrete steps forward during body-approach movements and discrete steps during head-orienting turns to the left or right. After each head or body movement, the world coordinates of the animat are updated. During the initial model development, a minimum approach step size of 35 units was selected based on visual observation. In other words, after moving forward at least 35 units, there is a perceptible change in the scene. Multiples of this minimum step size were selected to define various approach speeds during exploratory reactive, or during goal-oriented behavior. The four multipliers: 3, 4, 5 and 6 correspond to step sizes of 105, 140, 175 and 210 units, respectively. These four approach speeds are denoted: very slow, slow, normal, and motivated.

When approaching under visually reactive or planned control, the animat moves at the normal speed of 175 units per step. When the Approach  $GO_P$  receives incentive motivational signals from the Sensory-Drive Heterarchy module, then the motivated speed of 210 units per step is used. A heuristic rule causes the animat to slow down as it approaches a target location. If the animat's body dynamics were modeled in further detail, this rule would reduce overshoot or collisions with the target object. The model is designed to reduce spurious motion signals during a planned approach. As the distance field of the  $DV_M$  (Figure 15d; Section 2.3.6) drops below an activity level equivalent to 750 units, the animat drops one speed level. Then, as the distance field of the  $DV_M$  drops below 375 units, the animat drops to a second, even more reduced speed

level. The animat continues to approach at the slow speed of 140 units per step. The animat does not stop entirely while slowing down under planned control.

The control of head-orienting movements has been chosen to complement the various body-approach speeds. When the Spatial Attention module (Figure 5) triggers a head-orienting turn, there is a two-step process which ensues. First, the animat's approach speed is decreased and a fast head-orienting movement begins. A baseline of 2.5 degrees per step of head orienting was selected because it is small enough for precise head orienting, but large enough to readily scale to rapid head orienting. The initial fast head-orienting speed is proportional to the animat's approach speed. For very slow or slow approach rates, defined above, a multiplier of 2 yields a turn rate of 5 degrees per step. For normal approach rates, a multiplier of 3 yields a turn rate of 7.5 degrees per step. Finally, for motivated approach rates, a multiplier of 6 yields a turn rate of 15 degrees per step. Such speed-sensitive head orienting emulates a system in which the head-orienting control signal is proportional to the speed of self-motion. The first phase of fast head orienting continues for 45 degrees of head rotation. This phase is designed to orient the animat away from the current visual target object. Once it is complete, the visual target object is out of the camera's field-of-view. In the second phase of head orienting, a slower head turn with a multiplier of 1.5, yielding a turn rate of 3.75 degrees per step, is used. The second phase of head-orienting slowly brings the new visual target object into view. The reactive head turn continues until the object is centered in the visual field. In short, these heuristics are chosen because they enable head orienting to reliably bring new visual target information into view over a range of approach speeds.

After each step of the simulation, the Matlab graphical environment updates the 2D rendering of the 3D perspective view. Then, a 2D "snapshot" of the 3D scene is copied into memory for further processing. This image in memory comprises the content of the image plane if acquired by an equivalent combination of camera and lens. By using 2D "snapshots" of the current 3D perspective view, and several stages of image processing and learned categorization, the Parvo processing stream computes the head-centered distance and azimuth to the target (Section A1.1) and an invariant 2D representation of the visual target object. The head-centered azimuth is converted to body-centered coordinates by adding the current head-to-body angle.

The body-centered target position is then sent to the Reactive Visual TPV module (Figure 5; equations (10) and (30)) which competes with motion signals to generate reactive head-orienting and body-approach movements. On the other hand, the Invariant Visual Target Map is categorized by unsupervised Fuzzy ART learning for later storage in the Visual Working Memory and Planning System (Figure 17). The stages of processing, similar to the VIEWNET model (Bradski and Grossberg, 1993), are implemented as follows:

1. Capture a 2D view (RGB color format) from the 3D visual front-end. Each 2D view has 384 x 288 pixels (width x height) of resolution.

2. Convert the 2D image to grayscale (YUV format) via linear combination of RGB values. The YUV model defines a color space in terms of one luminance and two chrominance components. Y stands for the luminance component (the brightness) and U and V are the chrominance (color) components. YUV signals are created from an original RGB (red, green and blue) source. The weighted values of R, G and B are added together to produce a single Y signal, representing the overall brightness, or luminance, of that spot. The U signal is then created by subtracting the Y from the blue signal of the original RGB and then scaling, while V is derived by subtracting the Y from the red, and then scaling by a different factor. The Y channel preserves nearly all of the information of an equivalent black-and-white camera. In this simulation, the Y

channel is sufficient for image segmentation, and is chosen for that reason. The relationship adopted for computing the Y channel response,  $y(i,j)$ , from RGB input is shown in the following equation,

$$y(i, j) = \frac{1}{2}r(i, j) + \frac{1}{3}g(i, j) + \frac{1}{6}b(i, j), \quad (5)$$

where  $r(i,j)$ ,  $g(i,j)$ , and  $b(i,j)$  denote the red, green, and blue components of the color image at pixel coordinates  $(i,j)$ . A unique grayscale value is guaranteed under these experimental conditions since uniform lighting ensures that each color will always have a unique luminance.

3. Compute figure-ground separation by scene segmentation using the unique grayscale value of visual target objects. This is performed by copying those pixels with a grayscale value corresponding to the yellow of the visual target objects, to another image map. Without a minimum number of pixels, it is difficult to reliably compute the Invariant Visual Target Map (Figure 5). Continue to step 4, only if the area of the visual target object in the resulting image map exceeds a minimum threshold. (The minimum area threshold is 25 pixels.) Otherwise, do not update the Parvo, or form, output signals of the Visual Form and Motion System.

4. Perform edge detection via a 2D Laplacian-of-Gaussian filter (Schalkoff, 1989). This step is implemented by performing a 2D convolution on the grayscale image using a  $3 \times 3$  kernel,  $h(n_1, n_2)$ , defined by:

$$h(n_1, n_2) = \frac{(n_1^2 + n_2^2 - 2\sigma^2)h_g(n_1, n_2)}{2\pi\sigma^6 \sum_{n_1} \sum_{n_2} h_g} \quad (6)$$

where

$$h_g(n_1, n_2) = e^{-(n_1^2 + n_2^2)/(2\sigma^2)}, \quad (7)$$

and the indices of the  $3 \times 3$  kernel are  $n_1 = 1, \dots, 3$  and  $n_2 = 1, \dots, 3$ . The variance parameter,  $\sigma = 0.1$ , is a constant which defines the sharpness of the filter. This parameter is chosen empirically to yield a closed contour of the target object, which is suitable for further invariance processing.

5. Find the centroid of the target object and center it by shifting horizontally and vertically. The horizontal position of the target centroid relative to the head direction is extracted for the Reactive Visual TPV.

6. Apply a log-polar transformation by replacing pixel coordinates  $(x,y)$  with  $\log_{10}(r)e^{i\theta}$ .

7. Find the centroid of the transformed image and center it by scaling and rotating. The resulting representation of the visual target object is invariant to 2D variations in position, size and rotation.

8. Update the Invariant Visual Target Map by coarse-coding the centered log-polar image. A coarse-coded image requires less memory, can help to correct 3D foreshortening effects, and can improve generalization when categorized. Coarse-coding is accomplished first by computing a 2D convolution of the centered log-polar image with a Gaussian kernel having standard deviation,  $\sigma = 12$ . The result is sampled on an evenly spaced grid, with a sample period of 24 pixels, yielding an Invariant Visual Target Map with  $12 \times 16$  pixels (height  $\times$  width) of resolution (Figure 6b). The standard deviation of the Gaussian kernel is chosen to be one-half of the sample period by the Nyquist criterion, which ensures that the original image is accurately represented by its samples, as was done in the VIEWNET model (Bradski and Grossberg, 1993).

9. Learn to categorize the Invariant Visual Target Map by unsupervised Fuzzy ART learning (Section A1.2). The input to the Fuzzy ART module (Figure 5) is formed from the output of the Invariant Visual Target Map as a 192-element vector (12 x 16), which is then complement-coded (see equation (11)). The resulting 384-element vector is categorized using the Fuzzy ART classifier.

10. The Fuzzy ART categories comprise the “What” component and the stored Reactive Visual TPV defines the “Where” component of the Parvo stream.

**A1.1 Perspective transformation.** In order to solve for the world coordinates of objects that appear in the image plane of the camera, the effects of the perspective transformation should be considered (Schilling, 1990). Consider, for example, the case in which a camera and lens are aimed directly at the centroid of a visual target object. Suppose that the image plane of the camera lies in the plane defined by the x- and y-axes. Therefore, the image plane is aligned such that the z-axis connects the center of the image plane and the centroid of the visual target object. The distance between the image plane and lens, or focal length, is unknown in this case. Just as in a real camera, the focal length must be computed in order to translate between image and world coordinates. A simplified expression for the effective focal length,  $f$ , of the camera is obtained by the following expression:

$$f = \frac{z_0 x_I}{x_I - x_0}, \quad (8)$$

where  $z_0$  is the distance to the visual target object in world coordinates,  $x_I$  is the target height in image coordinates, and  $x_0$  is the target height in world coordinates (Schilling, 1990). World coordinates are known and the image coordinates are chosen manually from the 2D snapshot of a visual target object.

Once the effective focal length,  $f$ , is known, then the expression in equation (8), can be solved for  $x_0$  to yield an expression for the visual target height in world coordinates at any time. This relationship is defined by the following equation:

$$x_0 = \frac{x_I (f - z_0)}{f}. \quad (9)$$

Using equation (9), any point in the image plane, such as the centroid of a newly seen visual target object, can be translated into world coordinates. If the visual target object does not lie in the center of the visual field, then its relative offset in world coordinates can be computed. The head-centered azimuth,  $\alpha$ , and body-centered azimuth,  $P^a$ , are obtained once such world coordinates are known, by:

$$\alpha = \arctan \frac{x_0}{z_0}; \quad P^a = \alpha + \alpha_{HB}, \quad (10)$$

where the head-centered azimuth is converted to body-centered coordinates by adding the current head-to-body angle,  $\alpha_{HB}$ . The body-centered target distance,  $P^d$ , is computed algorithmically by calculating the distance between the body position and the target position in world coordinates (see equation (30)). The body-centered target azimuth,  $P^a$ , and target distance,  $P^d$ , comprise the angle and distance of the Reactive Visual TPV module (Figure 5).

**A1.2 Fuzzy ART.** The Fuzzy ART model (Carpenter, Grossberg, and Rosen, 1991) is capable of unsupervised learning of recognition categories in response to binary or analog input patterns. This model is the simplest ART model that can categorize analog patterns such as those produced by the Invariant Visual Target Map. The control of Fuzzy ART dynamics is

accomplished by the selection of the choice parameter,  $\alpha$ , the vigilance parameter,  $\rho$  and the learning rate,  $\beta$ , discussed below.

Normalization of Fuzzy ART input patterns is necessary to prevent category proliferation. Let  $\mathbf{a} = (a_1 \dots a_M)$  denote the M-dimensional ON-cell response of the input pattern. Each input pattern is preprocessed using *complement coding*, which combines the responses to the input pattern of ON cells and the complementary responses of the corresponding OFF cells. Together, the total pattern of ON and OFF cell activities is normalized. The complement-coded OFF cell response vector,  $\mathbf{a}^c$ , is defined by:

$$\mathbf{a}^c = 1 - \mathbf{a}. \quad (11)$$

The total input pattern comprises both the ON and OFF responses:

$$I = (\mathbf{a}, \mathbf{a}^c) = (a_1 \mathbf{K} a_M, a_1^c \mathbf{K} a_M^c), \quad (12)$$

and has the  $L^1$  norm of M. Among other benefits, complement coding enables the categorization of input patterns that contain all zeros. In the SOVEREIGN model, this event might occur if no target object is visible at the beginning or end of a head orienting turn. In such an instance, the output of the log-polar and Gaussian processing stages is also zero.

Category activation is controlled by input pattern,  $I$ , as it is filtered by the learned weights,  $w$ . Let  $T = (T_1 \mathbf{K} T_N)$  be the N-dimensional category input vector. Then the input to the  $j^{\text{th}}$  category is:

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + w_j}, \quad (13)$$

where  $w_j$  is the weight vector whose synapses abut category  $j$  and  $\alpha$  is a choice parameter that influences how categories are searched. The winning category  $J$  is the one whose input is maximal:

$$T_J = \max \{T_j : j = 1 \mathbf{K} N\} \quad (14)$$

If multiple categories have the same input, the one with the lowest numbered index is chosen as the winner. The choice parameter  $\alpha = 0$  for all simulations because small values of  $\alpha$  tend to minimize category recoding during learning.

In (13),  $\wedge$  is the fuzzy AND operator (Zadeh, 1965) defined by:

$$(x \wedge y)_i \equiv \min(x_i, y_i). \quad (15)$$

The vector norm operator,  $|\cdot|$ , is defined by:

$$|x| \equiv \sum_{i=1}^M |x_i|. \quad (16)$$

Once the winning category is chosen, a vigilance test is performed to ensure that the match is sufficiently close. If it is, then resonance and learning ensue. If the match is not good enough, reset ensues which causes a memory search whereby another category index is chosen. The degree of match required is controlled by the vigilance parameter,  $\rho$ . If:

$$\frac{|I \wedge w_j|}{|I|} \geq \rho, \quad (17)$$

then resonance occurs, since the input is close to the learned weight,  $w_j$ . A high degree of accuracy is required to discriminate visual target objects at long-range. Thus the choice of  $\sigma = 0.995$  was made. Conversely, if

$$\frac{|I \wedge w_j|}{|I|} < \rho, \quad (18)$$

then reset occurs, leading to a memory search, or hypothesis testing, whereby a new category index,  $j$ , is chosen. This reset cycle continues until either a sufficiently good match occurs or an uncommitted cell is chosen. Once the resonance criterion is satisfied, the chosen adaptive weight vector is updated as follows:

$$w_j^{(new)} = \beta (I \wedge w_j^{(old)}) + (1 - \beta) w_j^{(old)}. \quad (19)$$

The learning rate,  $\beta$ , determines the speed at which the weights converge to the input values. By using the maximal learning rate  $\beta = 1$ , categories are formed after a single input presentation.

## A2. Magno processing

The Visual Form and Motion System (Figure 5) computes all of the visual inputs to the SOVEREIGN model. Each 3D view of the environment is generated in color by an OpenGL-based graphical simulator, implemented within Matlab. To simulate the Magno response due to self-motion, consecutive pairs of 2D “snapshots”, or doublets, of the current 3D perspective view are sampled during body-approach movements. Since approach movements are separated into discrete steps for ease of computation, these steps are a convenient way to define the spatial sampling of static frames for Magno processing. During a body-approach movement, the first frame of a doublet is selected after a fixed step distance in world coordinates. In contrast, the second frame of a doublet is selected after a step distance which is proportional to the animat’s approach speed. The constant first frame spacing is chosen equal to twice the just noticeable difference in visual input, or  $35 \cdot 2$  units in world coordinates. The intra-doublet spacing is chosen equal to  $35 \cdot \beta$ , where  $\beta = 1, 2$  or  $3$ , while approaching at the slow, normal and motivated approach speeds, respectively. Consecutive snapshot pairs, or doublets, form the transient inputs to a dynamic neural network, discussed below.

The transient cells are designed to respond to variable speeds. The following equations are a subset of the circuitry derived for this purpose by Grossberg, Mingolla and Viswanathan (2001). The Directional Transient Cell Network (Figure 7) is defined by three stages of cells. The first stage comprises the unidirectional transient cells, which compute the transient response to the consecutive frames of a doublet. The unidirectional cell response,  $b_{ij}$ , at position  $(i,j)$  is defined by:

$$b_{ij} = \begin{cases} U[g_{ij}(t+K) + g_{ij}(t)] & \text{if } t < \alpha \\ U[g_{ij}(t+K) - g_{ij}(t)] & \text{if } t \geq \alpha, \end{cases} \quad (20)$$

where the transient response is bounded by a step function,  $U[x] = 2$  for  $x > 0$ , and 0 otherwise. The stimulus is presented for the duration  $\alpha = 0.2$  seconds. The activities of unidirectional cell input,  $g_{ij}(t)$ , at position  $(i,j)$  and time  $t$ , is defined as follows:

$$g_{ij}(t) = \begin{cases} 2 & \text{if } h_{ij}(t) > \varepsilon \\ 0 & \text{otherwise,} \end{cases} \quad (21)$$

where  $h_{ij}(t)$  is the grayscale value of the 2D “snapshot” at position (i,j) and time t. The threshold parameter  $\varepsilon = 0.5$ . Likewise,  $g_{ij}(t + K)$  is the response to the input at the same location, but at a later time,  $t + K$ . The fixed time interval, K, is approximated by choosing the approach speed-sensitive spacing of consecutive pairs of doublets using the method described above. Thus,  $g_{ij}(t)$  and  $g_{ij}(t + K)$  correspond to the thresholded, saturating responses to the 2D “snapshots” of the first and second frames of a doublet, respectively.

In the second stage, directional interneurons perform a time-average of unidirectional transient cell activities. The cell activity,  $c_{ij}^d$ , of the directional interneuron at position (i,j) with directional preference  $d$  ( $d = L$  for left,  $R$  for right) is defined by:

$$\frac{dc_{ij}^d}{dt} = -c_{ij}^d + b_{ij} - 10 \left[ c_{xy}^D \right]^+, \quad (22)$$

where each cell receives excitatory input,  $b_{ij}$ , from an unidirectional cell at the same location, and inhibitory input,  $c_{xy}^D$ , from the directional interneuron tuned to the opposite direction and spatially offset one unit from (i,j) along the preferred direction. For example, a directional interneuron that is sensitive to leftward motion, receives inhibition from a directional interneuron sensitive to rightward motion, but offset one unit to the left. In the third stage, the activity of directional transient cell,  $e_{ij}^d$ , at position (i,j) with directional preference  $d$ , is defined by:

$$\frac{de_{ij}^d}{dt} = 10 \left( -e_{ij}^d + b_{ij} - 10 \left[ c_{xy}^D \right]^+ \right), \quad (23)$$

where the unidirectional cell response,  $b_{ij}$ , at position (i,j) and complementary directional interneuron response,  $c_{xy}^D$ , is defined in (22). There is no need for vertical directional interneurons or transient cells, since the output of this motion circuitry provides only a left/right orienting signal. The directional interneuron activities,  $c_{ij}^d$ , and directional transient cell activation,  $e_{ij}^d$ , are integrated for a period of 0.5 seconds to obtain robust response over a range of body approach speeds.

The activity of the directional transient cells,  $e_{ij}^d$ , at position (i,j) with directional preference  $d$  comprise the input to the Motion Direction Hemifield Map (Figure 5). This map is partitioned into a left and right hemifield, each of which is sensitive to the difference between the activities of directional transient cells of opposite directional tuning. For example, map cells in the left hemifield,  $e^L$ , are sensitive to the difference between left and right directional transient cell activation, whereas cells in the right hemifield,  $e^R$ , are sensitive to the difference between right and left directional transient cell activation. The Motion Left/Right Decision module yields the total normalized left hemifield response,  $d^L$ , using the following equation:

$$d^L = \frac{1}{N} \frac{2}{M} \sum_{i=1}^{N/4} \left( \sum_{j=1}^{M/2} \left[ \left[ e_{4i,j}^L \right]^+ - \left[ e_{4i,j}^R \right]^+ \right]^+ \right) \quad (24)$$

and the total normalized right hemifield response,  $d^R$ , using the following equation:

$$d^R = \frac{1}{N} \frac{2}{M} \sum_{i=1}^{N/4} \left( \sum_{j=(M/2+1)}^M \left[ \left[ e_{4i,j}^R \right]^+ - \left[ e_{4i,j}^L \right]^+ \right]^+ \right), \quad (25)$$

where  $N$  is the number of rows, and  $M$  is the number of columns in the Motion Direction Hemifield Map. To reduce computational demands, the vertical index,  $i$ , is evaluated for only  $\frac{1}{4}$  of the total number of rows,  $N$ , in the 2D grayscale snapshot. The normalized quantity,  $d^L$ , is sensitive to the total difference between left-tuned and right-tuned directional transient cells in the left hemifield. Conversely,  $d^R$ , is sensitive to the total difference between right-tuned and left-tuned directional transient cells in the left hemifield. Both quantities tend to increase in absolute value as the total motion signals increase. Differences in the two signals,  $d^L$  and  $d^R$ , relative to Parvo signals, are used in the Spatial Attention module (Figure 5) and can command the Head-Orienting Movement module to turn away from the visual target cue.

### A3. Spatial Attention module

The Spatial Attention module (Figure 5) implements a competitive choice mechanism between visually-guided approach toward a visual target object, or motion-induced orienting away from a visual target object, or random head orienting in the absence of form or motion signals. The circuit is defined by a two-stage feedforward competitive network. In the first processing stage, cell activities  $l_1$  (Left),  $p_1$  (Parvo, or Straight) and  $r_1$  (Right) compute a time-averaged response to their inputs:

$$\frac{dl_1}{dt} = -Al_1 + G^L d^L, \quad (26)$$

$$\frac{dp_1}{dt} = -Ap_1 + G^P d^P, \quad (27)$$

and

$$\frac{dr_1}{dt} = -Ar_1 + G^R d^R, \quad (28)$$

where the decay constant  $A = 0.5$  and inputs  $d^L$  and  $d^R$  are obtained from the output of the Motion Left/Right Decision module, defined in equations (24) and (25). The Magno input gain parameters are constant; namely,  $G^L = G^R = 640$ . In contrast, the Parvo input gain parameter,  $G^P$ , is emotionally modulated by the output of the Sensory-Drive Heterarchy module,  $t_k$ , as follows:

$$G^P = \begin{cases} 1.15 & \text{if } \sum_k [t_k]^+ > \Gamma_t \\ 0.6 & \text{otherwise,} \end{cases} \quad (29)$$

where the  $k^{\text{th}}$  output,  $t_k$ , of the Sensory-Drive Heterarchy is defined in Section A6, and the threshold parameter  $\Gamma_t = 0.5$ . The normalized Parvo input,  $d^P$ , represents the visual target object salience:

$$d^P = \frac{(D_{MAX} - P^d)}{D_{MAX}}, \quad (30)$$

where  $D_{MAX} = 21,000$  is the maximum viewing distance from the animat to any 3D visual target object in world coordinates. (World coordinates refer to the spatial coordinate system which defines the location of points within the 3D graphical front-end.) The distance in world coordinates to the visual target object,  $P^d$  (see Section A1.1), comprises the distance field of the Reactive Visual TPV module (Figure 5).

The second processing stage implements competitive interactions between the cell activities,  $l_1$  (Left),  $p_1$  (Parvo, or Straight) and  $r_1$  (Right):

$$\frac{dl_2}{dt} = -A_1 l_2 + (B_1 - l_2) l_1 - (l_2 + E_1)(r_1 + p_1), \quad (31)$$

$$\frac{dp_2}{dt} = -A_1 p_2 + (B_1 - p_2) p_1 - (p_2 + E_1)(l_1 + r_1), \quad (32)$$

and

$$\frac{dr_2}{dt} = -A_1 r_2 + (B_1 - r_2) r_1 - (r_2 + E_1)(l_1 + p_1), \quad (33)$$

where  $A_1 = 0.1$ ,  $B_1 = 1$ , and  $E_1 = 0.5$ . The decay constant  $A_1$  is chosen small enough to allow the network response to be integrated over a relatively long time period; namely, several steps forward during a body-approach movement within the 3D virtual environment. The three decision signals,  $l_2^+$  (Left),  $p_2^+$  (Parvo, or Straight) and  $r_2^+$  (Right), are:

$$l_2^+ = [l_2 - \Gamma_c]^+, \quad (34)$$

$$p_2^+ = [p_2 - \Gamma_c]^+, \quad (35)$$

and

$$r_2^+ = [r_2 - \Gamma_c]^+, \quad (36)$$

where threshold  $\Gamma_c = 0.12$ .

Inertia in the response of the Motor Plant (Figure 2) to movement commands is modeled by an additional stage of processing that improves robust performance over variable approach speeds. This inertial effect also reduces false alarms due to noise in the environment. One source of noise is from variations in the visual scene resulting from the discrete steps used during body-approach movements. The final output is thus controlled by a heuristic Motor Plant,  $P_M$ . If  $P_M$  exceeds a variable threshold,  $\Gamma_M$ , then a head-orienting movement, or attentional shift, is initiated. After each step forward during a body-approach movement, the output is computed by the following algorithm:

1. If no decision signal ( $l_2^+$ ,  $p_2^+$ , or  $r_2^+$ ) is active, then decrement the value of  $P_M$  by one, stopping at zero.
2. If the Parvo output,  $p_2^+$ , is active, then decrement the value of  $P_M$  by one, stopping at zero.
3. If either Magno decision signal ( $l_2^+$  or  $r_2^+$ ) is active, then increment the value of  $P_M$  by one.
4. If the value of  $P_M$  exceeds a variable threshold,  $\Gamma_M$ , then an attention shift to the left or right is initiated in the direction commanded by  $l_2^+$  or  $r_2^+$ . The Head-Orienting Movement module commands a reactive head-orienting movement in the direction indicated and the value of  $P_M$  is set to zero. The threshold  $\Gamma_M = 2$  for the nominal approach speed and  $\Gamma_M = 1$  for the motivated walking speed. (The motivated walking speed is twice the nominal speed.) More details on the body approach speeds are provided in Section A9.3.
5. If no head-orienting movement is indicated, then the parvo decision signal ( $p_2^+$ ) commands the Reactive Visual TPV module (Figure 5; equations (10) and (30)) to load the Reactive Visual TPV Storage module (Figure 5; equation (116)), and then activates the

Approach or Orienting  $GO_p$  control signals (Figure 10; equations (101) and (103)) to initiate head-orienting or body-approach movements toward the target object. When the Approach or Orienting  $GO_p$  control signals are active, further update of the Reactive Visual TPV Storage module is inhibited.

6. If no decision signal ( $l_2^+$ ,  $p_2^+$ , or  $r_2^+$ ) is active, and the distance field, or measure of linear displacement, of the Approach GO signal,  $GO_d^p$  in the parvo stream (see equation (103)) is inactive, then no planned or reactive movement command is present. The Endogenous Random Generator, or ERG (see equation (105)), is then disinhibited to initiate exploration with a random head-orienting or body-approach movement. Further details on the ERG are provided in Section A9.3.

#### **A4. Visual Working Memory and Planning System**

The Visual Working Memory and Planning System, called henceforth the Visual Memory System for simplicity (Figure 17), computes motivationally-reinforced representations of sequences of 3D object categories. These sequences are categorized into goal-oriented list chunks, or plans, which are learned and read out under various exploratory or consummatory motivational states. The Visual Memory System comprises six different components.

**A4.1 View-invariant category learning.** The View-Invariant Category Learning module (Figure 17) links multiple view categories that are learned by Fuzzy ART into a view-invariant category representation of a visual object. Even with log-polar and Gaussian coarse-coding, there are enough changes in the Invariant Visual Target Map (Figure 5) during an approach movement that several dozen categories are formed in the Fuzzy ART Classifier module, as details of the visual target object are resolved within the 2D “snapshot”. The dynamics of view-invariant category learning have been described elsewhere (Bradski and Grossberg, 1995, Carpenter and Ross, 1993, Fazl, Grossberg and Mingolla, 2007). Here the process is realized algorithmically, for simplicity:

1. Learn a Fuzzy ART category ( $A_C$ ) in response to signals from the Invariant Visual Target Map module.
2. Determine if an associative link already exists from  $A_C$  to a view-invariant visual 3D object category ( $I_C$ ). If not, add  $A_C$  to a list mapping  $A_C$  to  $I_C$ .
3. When the Spatial Attention module signals an attentional shift and head-orienting movement away from the visual target object, create a new invariant visual 3D object category ( $I_C$ ) and associated list mapping.
4. View-invariant visual 3D object categories,  $I_C$ , comprise the inputs to the 3D Object Working Memory module.

**A4.2 3D Object working memory.** The 3D Object Working Memory module (Figure 17) stores lists of 3D object categories in a short-term working memory. Item-and-order information is represented by a spatial gradient of cell activation; that is, the most active item representations are performed first. We use the STORE 2 working memory model (Bradski, Carpenter and Grossberg, 1992, 1994). Subsequent to STORE 2 processing, an additional processing stage thresholds and normalizes the working memory output. This normalization maintains consistent pattern activity after the nonlinear thresholding operations. At most, three output cells have positive outputs with the chosen parameters. This limit provides a sufficient temporal context to disambiguate object sequence contexts in the current study. For simplicity, a list-based storage algorithm is employed to represent repeated items. A total of four items, repeated no more than twice, can be represented in any order within the 3D Object Working Memory module. Additional presentations of a given 3D object category will overwrite earlier items in the list.

The STORE 2 model equations generate a stable representation of event order that minimizes dependence on input presentation duration. The model dynamics also allow tuning of the shape of the stored activity gradient. A recency gradient is formed if the parameters are chosen such that the most recently active items have the highest activity levels.

Input,  $I$ , which represent individual items, are presented for the duration,  $\alpha$ , as follows:

$$I_i(t) = \begin{cases} 1 & \text{if } \alpha_i - t_i < t < t_i \\ 0 & \text{otherwise.} \end{cases} \quad (37)$$

The activity of input cells,  $x_i$ , respond to total input activity,  $I$ , and to stored memory cell activity,  $y_i$ , as follows:

$$\frac{dx_i}{dt} = (AI_i + y_i - x_i x - Bx_i)I, \quad (38)$$

where  $x = \sum_k x_k$  and  $I = \sum_k I_k$ . Term  $x$  helps to normalize the  $x_i$  response. Term  $I$  allows input cells to respond only when some input is on. The activity,  $y_i$ , of stored memory cells, tracks the activity,  $x_i$ , of input cells, as follows:

$$\frac{dy_i}{dt} = (x_i - y_i)I^C \quad (39)$$

where  $I^C \equiv 1 - I$ . These cells respond to the frozen  $x_i$  values when all inputs are off. Initially,  $x_i(0) = y_i(0) = 0$ .

Equations for the total input cell and stored memory cell activities are found by summing equations (38) and (39). By defining  $y \equiv \sum_k y_k$ , the following equations are obtained:

$$\frac{dx}{dt} = (AI + y - x^2 - Bx)I \quad (40)$$

and

$$\frac{dy}{dt} = (x - y)I^C. \quad (41)$$

Let the total input cell activity, after the  $i^{\text{th}}$  input shuts off at  $t_i$ , equal  $S_i \equiv x(t_i) = \sum_k x_k(t_i)$ . By (41),  $y(t_i) \cong x(t_{i-1}) \equiv S_{i-1}$  and by (40),

$$A + S_{i-1} - S_i^2 - BS_i \cong 0. \quad (42)$$

The total activity,  $S_1$ , during the first time interval,  $t_1$ , therefore obeys:

$$S_1 \equiv x(t_1) = 0.5 \left( -B + \sqrt{B^2 + 4A} \right) \quad (43)$$

where  $S_0 \equiv 0$ . Similarly, for subsequent time intervals,  $t_i$ , the total F<sub>1</sub> activity,  $S_i$ , obeys:

$$S_i \equiv x(t_i) = 0.5 \left( -B + \sqrt{B^2 + 4(A + S_{i-1})} \right) \quad i > 1 \quad (44)$$

where each cell activity,  $x_k(t_i)$  at position  $k$ , and time  $t_i$  of the working memory F<sub>1</sub>, obeys:

$$x_1(t_1) = S_1; \quad x_k(t_i) = \frac{x_k(t_{i-1})}{S_i + B}; \quad x_k(t_k) = \frac{A}{S_i + B} \quad (45)$$

These iterative STORE 2 working memory equations are employed within the SOVEREIGN model for computational simplicity. The decay parameter  $A = 0.01$ , and the working memory decay constant  $B = 1.45$  were selected to generate a gradual recency gradient.

The output of the STORE 2 network is supplemented by additional processing steps. The STORE 2 output is thresholded:

$$z(t_i) = \begin{cases} x(t_i) & \text{if } x(t_i) > \Gamma_w \\ 0 & \text{otherwise,} \end{cases} \quad (46)$$

with threshold  $\Gamma_w = 0.0022$ , chosen so that no more than three cells are active at once. Second, the output is then normalized so that the total activation  $\sum_i H(t_i) = 1.5$ :

$$H(t_i) = 1.5 \frac{z(t_i)}{\sum_k z(t_k)} \quad (47)$$

Normalized working memory activity,  $H(t_i)$ , is stored using a simplified list-based algorithm to represent repeated items. During animat exploration, visual target objects may be seen more than once. This leads to multiple instances of 3D object categories stored in the working memory. The ability to represent repeated items may be necessary to reach a goal location. Since repeated items are overwritten by new instances of the same 3D object category. The STORE 3 model (Bradski, Carpenter and Grossberg, 1994) provides a method for storing repeated working memory items. Here, a simple list-based algorithm was adopted instead, that is defined by the following steps:

1. At the start of a trial, create a list with eight zero entries to store enumerated order information for four view-invariant visual 3D object categories,  $I_C$  (see Section A4.1), each repeated at most twice. The list is represented by a 2D array with dimensions: (*category, repeat*).

2. To add a new object category to the list, search the list for non-zero entries. Entries in the list are numbered from 1...N. The highest numbered entry, N, designates both the most recent stored object category and total number of stored items. Check list location ( $I_C, 1$ ) to determine if  $I_C$  has already been stored once. Then, complete one of the following:
  - 2a. If  $I_C$  has not been stored once, then increment N, and store the result, N + 1, in list location ( $I_C, 1$ ).
  - 2b. If  $I_C$  has been stored once, then increment N, and store the result in list location ( $I_C, 2$ ).

3. To read out the contents of the working memory, and to convert them to STORE 2 activity levels, search the list for non-zero entries. Such a conversion provides inputs to the Masking Field module. Compute the activities,  $H(t_i)$ , of the limited-length working memory, defined by (47), after presenting N inputs at times  $t_i$ , where  $i=1, \dots, N$ . Form an eight-dimensional (4 categories x 2 repeats) output vector,  $I$ , by using the enumerated category order information from the list array to index into the N locations of  $H(t_i)$ . The resulting output vector,  $I$ , is an ordered copy of STORE 2 activities,  $H(t_i)$ , which form inputs to the Masking Field module (Section A4.3; equations (48)-(50),(52)).

**A4.3. Masking field.** The Masking Field module is a short-term memory which activates a single object list category, or plan, in response to combinations of one, two, or three active items in the 3D Object Working Memory (Cohen and Grossberg, 1986, 1987; Grossberg, 1978; Grossberg and Myers, 2000). Each plan cell receives permutations of subsets of bottom-up input signals and the cell with the closest matching bottom-up signals inhibits the activity of other plan cells (Section 2.6.2). Specifically, the activity,  $x_i^{(J)}$ , of the  $i^{\text{th}}$  plan cell, that receives the inputs  $J$ , where  $|\cdot|$  denotes the size of set  $J$ , obeys:

$$\begin{aligned} \frac{dx_i^{(J)}}{dt} = & -x_i^{(J)} + (B_1 - x_i^{(J)}) \left\{ \left( \frac{11}{|J|+8} \right) \sum_{m \in J} J_m w_{mi}^{(J)} z_{mi}^{(J)} + D_1 f(x_i^{(J)}) \right\} \\ & - F_2 (x_i^{(J)} + E_1) \left\{ \frac{\sum_{m,K} f(x_m^{(K)}) K (1 + |K \cap J|)}{\sum_{m,K} |K| (1 + |K \cap J|)} \right\}. \end{aligned} \quad (48)$$

In (48), the excitatory saturation level  $B_1 = 1$ , the inhibitory gain  $F_2 = 500$ , and the inhibitory saturation level  $E_1 = 10$ . The choice  $E_1 \gg B_1$  ensures that plan cells whose bottom-up signals  $J_m$  do not closely match their weights  $w_{mi}^{(J)}$  are strongly inhibited.

The excitatory input term in (48) comprises two parts. The first part is the filtered sum over all bottom-up inputs to plan cell,  $x_i^{(J)}$ :

$$\left( \frac{11}{|J|+8} \right) \sum_{m \in J} I_m w_{mi}^{(J)} z_{mi}^{(J)} \quad (49)$$

where the normalizing term  $\frac{11}{|J|+8}$  decreases with increasing cell size,  $|J| = \{1, 2, 3\}$ ,  $I_m$  is the activity of the  $m^{\text{th}}$  input from the 3D Object Working Memory,  $w_{mi}^{(J)}$  represents the strength of the learned weight connecting the  $m^{\text{th}}$  input in set  $J$  from the 3D Object Working Memory to the  $i^{\text{th}}$  plan cell, and the fixed connection weights,  $z_{mi}^{(J)}$ , cause slightly different initial inputs to different plan cells with the same number of inputs,  $|J|$ . When  $|J|=1$ , the coefficients  $z_{mi}^{(J)} = \{1.0\}$ . When  $|J|=2$ ,  $z_{mi}^{(J)} = \{0.95, 1.05\}$ , and when  $|J|=3$ ,  $z_{mi}^{(J)} = \{0.95, 1.0, 1.05\}$ . These fixed connections are similar to the random normalized growth rule of Cohen and Grossberg (1987). In the present case, random coefficients are not necessary, since permutations of the fixed, deterministic coefficients,  $z_{mi}^{(J)}$ , are sufficient to select the plan cells.

The adaptive weight,  $w_{mi}^{(J)}$ , obeys the ART 1 category learning law (Carpenter and Grossberg, 1987):

$$\frac{dw_{mi}^{(J)}}{dt} = f(x_i^{(J)}) \left[ (1 - w_{mi}^{(J)}) L I_m - K w_{mi}^{(J)} \sum_{n \neq m} I_n \right], \quad (50)$$

where learning is gated on and off by the sigmoid signal function:

$$f(x_i^{(J)}) = \frac{\left( [x_i^{(J)}]^+ \right)^2}{1 + \left( [x_i^{(J)}]^+ \right)^2} \quad (51)$$

The positive constants  $L$  and  $K$  control the normalizing competition between the inputs,  $I_n$ , in allocating learned weight. At steady-state,

$$w_{mi}^{(J)} = \frac{LI_m}{(L-K)I_m + K \sum_{n=1}^N I_n} \quad (52)$$

where the constants  $L = 2.0$ ,  $K = 1.0$  and  $N = 8$ .

The second term,  $D_1 f(x_i^{(J)})$ , of the total excitatory input in (48) is a sigmoid positive feedback signal with  $D_1 = 50$ , and enables plan cells with more inputs to suppress those plan cells that represent fewer inputs. The total inhibitory input term in (48) obeys:

$$\frac{\sum_{m,K} f(x_m^{(K)}) |K| (1 + |K \cap J|)}{\sum_{m,K} |K| (1 + |K \cap J|)}, \quad (53)$$

where the inhibitory signals from competing plan cells are scaled by a size-dependent strength term  $|K|$  and by the size  $|K \cap J|$  of the shared set of cells.

**A4.4. Object list working memory.** The Object List Working Memory module (Section 2.6.5; Figure 17), stores sequences of object list categories, stored one-by-one from the winner-take-all Masking Field module. Object list categories accumulated here provide context before motivational association in the next level, the Gated Multipole module (Section A4.5). This working memory activity gradient is also stored by a STORE 2 network, as in the 3D Object Working Memory module (Section A4.2). Again, only three list category cells can have a positive output with the chosen parameters. A list-based algorithmic stage (see Section A4.2) allows the encoding of as many as eight items, each repeated at most twice. Items are stored in the list when the volitional signal to approach a target object is released (Section A9.3). The stored list is reset at the beginning of each trial. The parameters and other details are the same as in the 3D Object Working Memory module (Section A4.2).

**A4.5. Gated multipoles.** The Gated Multipole module (Figures 17, 18 and 20b), computes emotionally modulated object list categories. This module enables motivationally-compatible plan items to learn and to read out head-orienting and body-approach commands. There is one three-channel Gated Multipole module for each cell in the Object List Working Memory module. Each channel corresponds to a distinct motivational state that is modulated by conditioned reinforcement and incentive motivational learning. When signals from the Object List Working Memory module, representing the stored list categories, reach the Gated Multipole module, the signals are broadcast across all three drive channels via a hard-wired spatial gradient. The prepotent channels receive the largest bottom-up signals, and rapidly win the feedforward competition. Top-down feedback within each gated multipole (Figure 20b) strengthens this choice and results in strong habituation in the feedforward pathway (see rectangular synapses in Figure 20b). The winning channel from among all of the gated multipoles can then elicit top-down signals from the winner-take-all computation (Figure 20b) which can learn or read out motor commands. In the event that such a top-down command does not match the current NET head-orienting or body-approach movement, a mismatch arousal burst can rebound the gated multipole and select a different drive channel. The bottom-up choice and mismatch cycle continues until an uncommitted drive channel is able to sample the NET head-orienting or body-approach movement. In this way, each plan can come under the learned control of any drive. After repeated experience with a particular reward, learning trials based upon reactive behaviors gradually give way to planned responses.

The Gated Multipole modules comprise a four-layer feedback network. The activity,  $p_{ij}$ , of first layer cells in the  $j^{\text{th}}$  drive channel of the  $i^{\text{th}}$  Gated Multipole are defined by:

$$\begin{aligned} \frac{dp_{ij}}{dt} = & -A_1 p_{ij} + (B_1 - p_{ij}) \left( C_1 y_i S_j (1 + D_1 f_u(u_{ij})) + E_1 \sum_k [t_k]^+ w_{kij}^{(2)} + L \right) \\ & - F_1 p_{ij} \sum_{k \neq j} f_u(u_{ik}). \end{aligned} \quad (54)$$

The decay constant  $A_1 = 1$ , the excitatory saturation level  $B_1 = 1$ , the excitatory input gain  $C_1 = 2$ , and the bottom-up spatial gradient across drive channels  $S_j = \left\{ \frac{1}{6}, \frac{1}{3}, \frac{1}{2} \right\}$ . The inputs,  $y_i$ , are the sigmoid output signals from the stored object list category cells:

$$y_i = \frac{\left( [x_i^{(j)}]^+ \right)^2}{0.1 + \left( [x_i^{(j)}]^+ \right)^2}. \quad (55)$$

The superscripts  $(j)$  in  $p_{ij}$  are not shown for simplicity. In (54), top-down attentional feedback from the final choice stage of the multipole provides excitatory modulation of the input,  $y_i$ , via term  $(1 + D_1 f_u(u_{ij}))$  that is balanced by lateral inhibition via term  $-F_1 p_{ij} \sum_{k \neq j} f_u(u_{ik})$ . The inhibitory gain  $F_1 = 3$ . The sigmoid excitatory feedback signal:

$$f_u(u_{ij}) = \frac{\alpha^N \left( [u_{ij}]^+ \right)^N}{\alpha^N + \theta \left( [u_{ij}]^+ \right)^N}, \quad (56)$$

where  $\alpha = 1.7$ ,  $N = 5$ ,  $\theta = 9.8$  and the excitatory feedback gain  $D_1 = 18$ . These parameters were chosen following numerical search and simulation, although others are possible, owing to the robustness and flexibility of the model. Term  $E_1 \sum_k [t_k]^+ w_{kij}^{(2)}$  alters the plan choice using learned incentive motivational signals from the Sensory-Drive Heterarchy (Section A6). The excitatory incentive gain  $E_1 = 2$ . The arousal burst  $L$  ( $L=10$  when ‘‘on’’) is a signal arising from a mismatch between the Planned DV and NET modules (Figure 20b). The onset of arousal rebounds, and thereby resets, all active Gated Multipoles, allowing them to gain conditioned reinforcement and incentive properties under a different motivational state.

Habituated transmitter levels (rectangular synapses in Figure 20b) are depleted in response to activity of the first-layer cells,  $p_{ij}$ . The transmitter levels,  $v_{ij}$ , in the  $j^{\text{th}}$  drive channel of the  $i^{\text{th}}$  Gated Multipole, obey:

$$\frac{dv_{ij}}{dt} = (1 - v_{ij}) C_2 - F_2 v_{ij} p_{ij} \quad (57)$$

(Grossberg 1972, 1980), where the accumulation rate  $C_2 = 0.5$  and the linear depletion rate  $F_2 = 1.5$ . Such a habituated mechanism ensures that the most active drive channel of each first-layer cell is weakened by its own activity. Conversely, less active drive channels habituate relatively little. At the onset of an arousal burst, the other drive channels can therefore

immediately elicit a strong excitatory feedforward burst of activity (Grossberg, 1972, 1984a), and potentially win the competition at the fourth-layer (see equation (59)).

The activity of second-layer cells,  $q_{ij}$ , in the  $j^{\text{th}}$  drive channel of the  $i^{\text{th}}$  Gated Multipole, obey:

$$\frac{dq_{ij}}{dt} = -A_2 q_{ij} + C_3 p_{ij} v_{ij}, \quad (58)$$

where the signals from the first-layer cells,  $p_{ij}$ , are multiplicatively gated by the habituated transmitter,  $v_{ij}$ . The decay constant  $A_2 = 1$  and the input gain  $C_3 = 1$ .

The activity of third-layer cells,  $u_{ij}$ , obey:

$$\frac{du_{ij}}{dt} = -A_3 u_{ij} + (B_2 - u_{ij}) g(q_{ij}) - F_3 u_{ij} \sum_{k \neq j} g(q_{ik}), \quad (59)$$

where the decay constant  $A_3 = 7.5$ , and the excitatory saturation level  $B_2 = 170$ . These cells implement a feedforward competition to choose the most active drive channel, given the momentary combination of bottom-up list category signals, learned incentives, and the effect of arousal,  $L$ . Excitatory inputs from second-layer cells,  $q_{ij}$ , are amplified by a faster-than-linear feedback function,  $g(q_{ij})$ , defined by:

$$g(q_{ij}) = \left( [q_{ij}]^+ \right)^2 \quad (60)$$

and balanced by the inhibitory off-surround term,  $-F_3 u_{ij} \sum_{k \neq j} g(q_{ik})$ , where the inhibitory gain  $F_3 = 0.5$ . The winning cell then provides strong top-down support to first-layer cells,  $p_{ij}$ .

The activity of fourth-layer cells,  $z_{ij}$ , obey:

$$\frac{dz_{ij}}{dt} = -A_4 z_{ij} + (1 - z_{ij}) C_4 [u_{ij}]^+ \left( \varepsilon_1 + D_3 \sum_k [t_k]^+ \right), \quad (61)$$

where the decay constant  $A_4 = 1$  and the input gain  $C_4 = 0.5$ . Fourth-layer cells respond to excitatory signals from the third-layer choice cells. Activities of these fourth-layer cells,  $z_{ij}$ , are amplified by learned incentive motivational signals. The baseline gain  $\varepsilon_1 = 0.1$ , which provides some excitatory input, even when the nonspecific incentive signals,  $D_3 \sum_k [t_k]^+$ , are zero. The incentive gain  $D_3 = 0.5$ .

First-layer cells in (54) acquire conditioned reinforcer properties by sampling the drive representation activities,  $r_k$  (see Section A5), via doubly-gated outstar learning:

$$\frac{dw_{ijk}^{(1)}}{dt} = \eta N_{ij}^{(2)} p_{ij} \left( N_k^{(1)} r_k - w_{ijk}^{(1)} \right), \quad (62)$$

where the learning rate  $\eta = 1$ . The learned weights,  $w_{ijk}^{(1)}$ , where  $i$  is the Gated Multipole index,  $j$  is the drive channel, and  $k$  is the drive representation index, track the gated activity of the drive representation,  $N_k^{(1)} r_k$ . Both  $N_k^{(1)}$  and  $N_{ij}^{(2)}$  are gating signals which must be on for learning to occur. In particular,

$$N_k^{(1)} = \begin{cases} 1 & \text{if } t_k > 0 \\ 0 & \text{otherwise,} \end{cases} \quad (63)$$

which is triggered by Sensory-Drive Heterarchy cell activity,  $t_k$ , ensuring that the weights do not sample a zero quantity, and

$$N_{ij}^{(2)} = \begin{cases} 1 & \text{if } u_{ij} > 1 \\ 0 & \text{otherwise,} \end{cases} \quad (64)$$

which allows only the drive channel which wins the competition at third-layer cells, to acquire conditioned reinforcer properties. Learning is doubly-gated by activity at first-layer cells,  $p_{ij}$ , representing the outstar sampling signal. Initially, all conditioned reinforcement weights,  $w_{ijk}^{(1)}$ , are zero. Once the learned weights are strong enough, they can bias the activity at the Drive Representation module (see Section A5) and influence the selection of the most active motivational state.

First-layer cells (see equation (54)) can also acquire incentive motivational properties via gated instar learning as follows:

$$\frac{dw_{kij}^{(2)}}{dt} = \eta N_{ij}^{(2)} p_{ij} \left( [t_k]^+ - w_{kij}^{(2)} \right), \quad (65)$$

where again the learning rate  $\eta = 1$ . Adaptive weights,  $w_{kij}^{(2)}$ , track the activity,  $[t_k]^+$ , of the Sensory-Drive Heterarchy module. The term,  $N_{ij}^{(2)}$  gates learning on (equation (64)) when it is activated by the winning channel at the fourth-layer competition. Learning is further gated by activity at first-layer cells,  $p_{ij}$ , representing the instar sampling signal. The initial strength of incentive motivation weights,  $w_{kij}^{(2)}$ , is zero. When these learned weights are strong enough, they can bias the activity at first-layer cells,  $p_{ij}$ , potentially changing the feedforward signals through the Gated Multipole module, thus activating a different drive channel.

The ability of Gated Multipole modules to activate different drive channels after learning is demonstrated by simulation (Section 3) and summarized here. For example, in the Motivated Choice Experiment (Section 3.1), the animat learns to move Forward-Left-Forward when hungry to receive a food reward, and Forward-Right-Forward when thirsty to receive a water reward. In both instances, the animat initially responds to the Triangle cue (Figure 1a) under different motivational states, and learns to respond differently. The mechanism employed is the conditioning of two different drive channels in the Gated Multipole to the same compressed representation of the Triangle cue. Specifically, the prepotent channel is conditioned to the Forward-Left command under the hunger drive, whereas the second channel is conditioned to the Forward-Right command under the thirst drive. During maze exploration, the second drive channel of the Gated Multipole cell was conditioned to the Drive Representation (Section A5) thirst channel. Furthermore, the Sensory-Drive Heterarchy (Section A6) thirst channel was conditioned to the second channel of first-layer Gated Multipole cells. The effect of such a learned feedback loop is the direct reactivation of the second drive channel, when the Triangle cue is seen under the same motivational state.

After learning the Motivated Choice Experiment task, the animat can repeat the experiment from the beginning while thirsty. When the Visual Memory System (Figure 17) is updated at the start of the trial, a compressed category representation of the Triangle cue is stored

in each stage, up to the Object List Working Memory (Section A4.4). The activity,  $p_{ij}$  (equation (54)), of first-layer Gated Multipole cells begin to respond to Object List Working Memory signals,  $y_i$ , which are broadcast across all three drive channels via a hard-wired spatial gradient,  $S_i$ . The prepotent channel receives the largest bottom-up input, followed by the second and third channels. Subsequently, the activity,  $r_k$ , (equation (67)) of Drive Representation cells (Section A5) respond to learned conditioned reinforcer inputs,  $\sum_{jmn} p_{mn} w_{kmn}^{(1)}$ , from first-layer Gated Multipole cells.

Thus, the Drive Representation thirst channel receives both a drive input and a learned conditioned reinforcer input. In the absence of other more salient inputs, this combination can win the competition at the Sensory-Drive Heterarchy module (Section A6). Thus, the activity,  $t_k$ , of the Sensory-Drive Heterarchy cells can influence the activity first-layer Gated Multipole cells via learned incentive motivational input,  $E_1 \sum_k [t_k]^+ w_{kij}^{(2)}$  (equation (54)). These learned incentive

inputs are sufficient to overcome the bias implicit in the bottom-up spatial gradient,  $S_i$ . The second drive channel Gated Multipole cell receives more total input than either the prepotent or third drive channels. Consequently, the second drive channel can win the competition at the third-layer cells (equation (59)), which sends on-center off-surround top-down feedback to first-layer cells. This feedback enhances the activity of the first-layer cell receiving the largest bottom-up input, and suppresses the others. The second drive channel is thereby selected over the prepotent channel, and can again read out the command to move Forward-Right.

**A4.6. Winner-take-all competition.** The Winner-Take-All Competition module (Figures 17, 18 and 20b) computes a global choice, also called a Motivated WHAT and WHERE decision, across each set of Gated Multipoles. These Motivated WHAT and WHERE decisions correspond to the output of the Visual and Motor Memory Systems, respectively. The maximally active fifth-layer cells,  $z_{ij}$  (equation (61)), from the Gated Multipole module is computed by a steady-state approximation to a high-gain shunting feedback competitive network:

$$T(z_{ij}) = \frac{\left([z_{ij} - \Gamma_2]^+\right)^N}{\alpha^N + \sum_m \sum_n \left([z_{mn} - \Gamma_2]^+\right)^N}, \quad (66)$$

where the threshold  $\Gamma_2 = 0.1$ , exponent  $N = 240$ , and slope parameter  $\alpha = 0.05$ . The winning cell can activate top-down learning and read out of head-orienting and body-approach movements to the Top-Down Readout module (Section A8).

To select parameters, Gated Multipole activities,  $z_{ij}$ , under various conditions during maze exploration, are computed and processed by (66) using nominal values for  $\Gamma_2$ ,  $N$ , and  $\alpha$ . The amplitude of the winner,  $T$ , relative to other cells is then maximized by systematically varying each of the parameters in turn. This optimization process realizes good performance under all operating conditions.

## A5. Drive representation

The Drive Representation module (Figures 17, 18 and 20b), computes the internal drive state of the model by combining drive inputs,  $\bar{D}_k$ , unconditioned stimulus inputs,  $\bar{U}_k$  and conditioned reinforcer inputs,  $\sum_{Jmn} p_{mn} w_{kmn}^{(1)}$ :

$$\frac{dr_k}{dt} = -A_5 r_k + (B_3 - r_k) \left[ (\varepsilon_2 + \bar{D}_k) \left( \bar{U}_k + \sum_{Jmn} p_{mn} w_{kmn}^{(1)} \right) - \Gamma_3 \right]^+, \quad (67)$$

where the decay constant  $A_5 = 1$ , the excitatory saturation level  $B_3 = 1.25$ , and the baseline excitatory input gain  $\varepsilon_2 = 0.01$ . The drive inputs,  $\bar{D}_k$ , are on at the beginning of a learning trial and remain active until the end of the trial. The model can accommodate up to three drive inputs, denoted by  $\bar{D}_1$ ,  $\bar{D}_2$  and  $\bar{D}_3$ . The “exploratory” drive,  $\bar{D}_1$ , is assumed to be endogenously active. Consummatory drive inputs, like hunger,  $\bar{D}_2$ , or thirst,  $\bar{D}_3$ , can vary under different learning conditions. There are also unconditioned stimulus inputs,  $\bar{U}_k$ , which correspond to these exploratory or consummatory drive states. The “exploratory” learning signal,  $\bar{U}_1$ , is activated when head orienting is completed. Such a learning signal is the primary means of effecting learning without an explicit reward, or latent learning (see Section 3.4). In the model, this internally generated learning signal is triggered at the offset of the head orienting GO signal,  $\overline{GO}_a^p$  (equations (82) and (101)). Activation of  $\bar{U}_k$  can yield a strong excitatory response in (67) if the corresponding drive inputs,  $\bar{D}_k$ , are large. Even if the drive inputs are small, a large  $\bar{U}_k$ , can generate a positive response resulting from the multiplicative baseline gain,  $\varepsilon_2$ , provided that the input threshold  $\Gamma_3 = 0.025$  is exceeded. Similar conditions hold for the consummatory  $\bar{U}_k$ , which are activated when a rewarded location is reached. Finally, activation of any unconditioned stimulus reinforces the plan represented by Gated Multipole activity (equations (62) and (65)). Learned conditioned reinforcement signals,  $\sum_{Jmn} p_{mn} w_{kmn}^{(1)}$ , can eventually supplant  $\bar{U}_k$ , after repeated explorations that lead to reward.

#### A6. Sensory-drive heterarchy

The Sensory-Drive Heterarchy module (Figures 17, 18 and 20b), implements a competitive feedforward network whose winning cell represents the motivational state which has the largest total internal and external support. The activity of Sensory-Drive Heterarchy cells,  $t_k$ , are defined by the following shunting on-center off-surround network:

$$\frac{dt_k}{dt} = -t_k + (1 - t_k) g(r_k) - F_3 \sum_{m \neq k} g(r_m), \quad (68)$$

where signals from the Drive Representation module,  $r_k$ , are amplified by the faster-than-linear transfer function,  $g(\omega)$ , as follows:

$$g(r_m) = 100 \left( [r_m]^+ \right)^2, \quad (69)$$

and the inhibitory gain  $F_3 = 0.5$ . Sensory-Drive Heterarchy module activities,  $t_k$ , gates several types of learning, modulate the animat’s approach speed, and prime the Gated Multipole signals to the Winner-Take-All Competition module (Section A4.6).

### A7. Motor Working Memory and Planning System

The Motor Working Memory and Planning System, or Motor Memory System (Figure 18), computes motivationally-reinforced representations of sequences of motor map positions. These sequences represent goal-oriented motor plans, which are learned and read out under various consummatory or exploratory motivational states. The Motor Memory System comprises six different components. The Visual and Motor Memory System has five of these components in common. The only difference is that the Motor Memory System uses a heuristic Self-Organizing Map module to convert NET motor displacements, which are approach-orient pairs, into map positions which are suitable for encoding by STORE 2 dynamics into the Motor Map Working Memory. The other layers are described in the Visual Memory System sections.

### A8. Sensory-Motor Top-Down Readout

The Sensory-Motor Top-Down Readout module (Figure 20a), combines top-down signals from the Visual and Motor Memory System (Figures 17 and 18) to enable the learning and read out of motor displacement commands. The convergence of sensory and motor information is necessary because, under some circumstances, one type of signal may be unavailable. For example, at the beginning of a learning trial, the Visual Memory System is updated, leading to top-down readout. However, before a motor command is performed, no Motor Memory System readout is available. With the Top-Down Readout module, the system can accommodate this situation. This module comprises a four-layer feedforward network which represents distance and angular motor commands as opponent vectors. In the first layer, the WHAT and WHERE DVs compute the difference between top-down readout and net motor displacement. The distance (d) or angle (a) field ON-channel activities,  $p_{d/a}^{s+}$ , and OFF-channel activities,  $p_{d/a}^{s-}$ , of first-layer sensory cells are defined by:

$$\frac{dp_{d/a}^{s+}}{dt} = -p_{d/a}^{s+} + f(T)w_{d/a}^{s+} - t_{d/a}^{+} \quad (70)$$

and

$$\frac{dp_{d/a}^{s-}}{dt} = -p_{d/a}^{s-} - f(T)w_{d/a}^{s-} + t_{d/a}^{-}, \quad (71)$$

where  $T$  is the Motivated WHAT Decision from the Winner-Take-All Competition module in the Visual Memory System (Figure 17; equation (66)). The sigmoidal signal function,  $f(\omega)$ , is defined by:

$$f(\omega) = \begin{cases} 1 & \text{if } \omega > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (72)$$

The adaptive weights in the ON-channel,  $w_{d/a}^{s+}$ , and OFF-channel,  $w_{d/a}^{s-}$ , from the WHAT decision to the read-out module, encode the planned top-down body movement command. Top-down adaptive weight strengths,  $w_{d/a}^{s+}$  and  $w_{d/a}^{s-}$ , are initially zero. The ON-channel activity,  $t_{d/a}^{+}$ , and OFF-channel activity,  $t_{d/a}^{-}$  activities represent the calibrated NET body distance and angle displacement (Section A10; equation (120)). The activities,  $p_{d/a}^{m+}$  and  $p_{d/a}^{m-}$ , of first-layer motor cells are defined similarly for top-down signals from the Motor Memory System (Figure 18). The second layer amplifies and adds a tonic input to first-layer cell activities. If first-layer cells

have zero activity, then second-layer cells have constant tonic responses. The activities,  $q_{d/a}^{s+}$  and  $q_{d/a}^{s-}$ , of second-layer sensory cells are defined by:

$$\frac{dq_{d/a}^{s+}}{dt} = -q_{d/a}^{s+} + 2p_{d/a}^{s+} + T_s \quad (73)$$

and

$$\frac{dq_{d/a}^{s-}}{dt} = -q_{d/a}^{s-} + 2p_{d/a}^{s-} + T_s. \quad (74)$$

The tonic cell response,  $T_s$ , which is activated by first-layer cell activities,  $p_{d/a}^{s+}$  and  $p_{d/a}^{s-}$ , is defined by:

$$T_s = f\left(\left[p_{d/a}^{s+}\right]^+ + \left[p_{d/a}^{s-}\right]^+\right) \quad (75)$$

and the signal function,  $f(\omega)$ , is defined in equation (72). The activities of second-layer motor cells,  $q_{d/a}^{m+}$  and  $q_{d/a}^{m-}$ , are defined similarly. The third layer combines sensory and motor signals from the second-layer cell activities. Competitive interactions at third layer cells normalize the response to sensory and motor inputs to enable accurate learning, and read out even if one is missing. The activities,  $s_{d/a}^+$  and  $s_{d/a}^-$ , of third-layer sensory-motor cells are thus defined by:

$$\frac{ds_{d/a}^+}{dt} = -As_{d/a}^+ + (1-s_{d/a}^+)(q_{d/a}^{s+} + q_{d/a}^{m+}) - s_{d/a}^+(q_{d/a}^{s-} + q_{d/a}^{m-}) \quad (76)$$

and

$$\frac{ds_{d/a}^-}{dt} = -As_{d/a}^- + (1-s_{d/a}^-)(r_{d/a}^{s-} + r_{d/a}^{m-}) - s_{d/a}^-(r_{d/a}^{s+} + r_{d/a}^{m+}), \quad (77)$$

where the decay constant  $A = 0.01$ . The activities of fourth-layer cells,  $u_{d/a}^+$ , and  $u_{d/a}^-$ , define the Planned DV in the Motor DV System (Figure 14):

$$\frac{du_{d/a}^+}{dt} = -u_{d/a}^+ + s_{d/a}^+ - s_{d/a}^- \quad (78)$$

and

$$\frac{du_{d/a}^-}{dt} = -u_{d/a}^- + s_{d/a}^- - s_{d/a}^+. \quad (79)$$

These cells thus receive opponent inputs from the third layer to generate a differential response consistent with the NET body distance and angle displacement (equation (120)). Finally, top-down learning from the Visual Memory System (Figure 17) utilizes the VAM learning rule, whereby the adaptive weights,  $w_{d/a}^{s+}$  and  $w_{d/a}^{s-}$ , learn to zero the activities of first-layer cells,  $p_{d/a}^{s+}$  and  $p_{d/a}^{s-}$  (equations (70) and (71)). By doing so, the top-down weights,  $w_{d/a}^{s+}$  and  $w_{d/a}^{s-}$ , learn the NET body distance and angle displacement (Section A10; equation (120)), which can be read out during later exploration. These adaptive weights,  $w_{d/a}^{s+}$  and  $w_{d/a}^{s-}$ , are updated as follows:

$$\frac{dw_{d/a}^{s+}}{dt} = \eta f(T) \overline{GO}_a^p (-p_{d/a}^{s+} - w_{d/a}^{s+}) \quad (80)$$

and

$$\frac{dw_{d/a}^{s-}}{dt} = \eta f(T) \overline{GO}_a^p (-p_{d/a}^{s-} - w_{d/a}^{s-}), \quad (81)$$

where the learning rate,  $\eta = 1$ ,  $T$  is the Motivated WHAT Decision from the Winner-Take-All Competition module in the Visual Memory System (Figure 17) and the transfer function,  $f(\omega)$ , is defined in equation (72). The last term,  $\overline{GO}_a^p$ , is the activity of the learning gate derived from the Motor DV System (Figure 14), which turns on when head orienting is completed. This learning gate triggers learning (e.g., latent learning) in the absence of reward (see Section A5). It obeys:

$$\overline{GO}_a^p = 1 - f(GO_a^p), \quad (82)$$

where  $GO_a^p$  is the Parvo GO signal and  $f(\omega)$  is again defined in equation (72). Activation of  $\overline{GO}_a^p$  occurs at the offset of the  $GO_a^p$ . The  $GO_a^p$ , which commands head orienting movements (Figure 15a,c), shuts off when both reactive and planned head orienting signals are zero (see equation (101)). Top-down adaptive weights,  $w_{d/a}^{m+}$  and  $w_{d/a}^{m-}$ , from the Motor Memory System (Figure 18) are defined similarly.

**A8.1. Control signals.** The Top-down Readout Mismatch signal (Figures 10 and 20b) compares the Sensory-Motor NET Movement (Section A10) with the WHAT and WHERE DVs (Section A8). A sufficiently large difference between the NET body displacement and learned top-down weights leads to a mismatch. Mismatch leads to arousal, which rebounds and thereby resets active Visual and Motor Memory System Gated Multipoles (Figures 17 and 18).

The arousal calculation is accomplished in three stages. First, the activity,  $T_1$ , of the sensory tonic cell, is inhibited by the activities,  $p_d^{s+}$  and  $p_d^{s-}$ , (equations (70) and (71)) of first-layer Top-Down Readout sensory, or WHAT DV, cells:

$$T_1 = \left[ 2 - g(p_d^{s+}) - g(p_d^{s-}) - g(p_a^{s+}) - g(p_a^{s-}) \right]^+ . \quad (83)$$

Likewise, the activity,  $T_2$ , of the motor tonic cell, is inhibited by the activities,  $p_d^{m+}$  and  $p_d^{m-}$  of first-layer Top-Down Readout motor, or WHERE DV, cells:

$$T_2 = \left[ 2 - g(p_d^{m+}) - g(p_d^{m-}) - g(p_a^{m+}) - g(p_a^{m-}) \right]^+ , \quad (84)$$

where the sigmoidal signal function:

$$g(\omega) = \frac{([\omega]^+)^2}{0.25 + ([\omega]^+)^2} . \quad (85)$$

Next, the activity,  $U_{sm}$ , of the mismatch cell, which responds to the activities,  $t_{d/a}^+$  and  $t_{d/a}^-$ , (see equation (120)) of the Sensory-Motor NET Movement, is defined by:

$$U_{sm} = g(t_d^+) + g(t_d^-) + g(t_a^+) + g(t_a^-) . \quad (86)$$

Finally, the activities of sensory and motor arousal signals are defined by the following:

$$A_s = \begin{cases} 1 & \text{if } U_{sm} - T_1 > 0.15 \\ 0 & \text{otherwise} \end{cases} \quad (87)$$

and

$$A_m = \begin{cases} 1 & \text{if } U_{sm} - T_2 > 0.15 \\ 0 & \text{otherwise.} \end{cases} \quad (88)$$

Generally, the activities,  $T_1$  and  $T_2$ , of tonic cells increase when the activities,  $p_{d/a}^{s+}$  and  $p_{d/a}^{s-}$ , of WHAT DV cells and the activities,  $p_{d/a}^{m+}$  and  $p_{d/a}^{m-}$ , of WHERE DV cells decrease, respectively. In the extreme case, if WHAT and WHERE DV cells are zero, then the tonic cells are disinhibited, the activity,  $U_{sm}$ , of the mismatch cell is zero, and the activities,  $A_s$  or  $A_m$ , of arousal cells are zero. WHAT and WHERE DV activity results from differences between the top-down motor commands and the NET body displacement. If the activities,  $A_s$  or  $A_m$ , of the arousal cells exceed the fixed threshold (equation (87) and (88)), as is the case when then the WHAT and WHERE DV cells are sufficiently active, then the cells trigger an arousal burst in the Vision or Motor Memory System Gated Multipoles (Section A4.5), respectively.

### A9. Motor Approach and Orienting System

The Motor Approach and Orienting System (Figure 10) directs head orienting and body approach movements in response to reactive and planned motor commands. The body-centered distance and angle coordinates of the visual target object are computed in the Reactive Visual TPV module. The visual target coordinates flow into the Reactive Visual TPV Storage module and activates the Reactive DV module, which initiates head and body movements via motor processing downstream. The Reactive DV distance (d) or angle (a) ON-channel activity,  $v_{da}^+$ , obeys:

$$\frac{dv_{d/a}^+}{dt} = -v_{d/a}^+ + y_{d/a}^+ - t_{d/a}^+ w_{d/a}^{n+}, \quad (89)$$

where the excitatory input term,  $y_{da}^+$ , is the Reactive Visual TPV Storage module distance (d) or angle (a) ON-channel activity, and the inhibitory input term,  $t_{d/a}^+ w_{d/a}^{n+}$ , is the NET body displacement activity,  $t_{d/a}^+$ , scaled by learned adaptive weights,  $w_{d/a}^{n+}$  to make the two terms dimensionally consistent. The OFF-channel response of the Reactive DV module distance or angle fields,  $v_{d/a}^-$ , is defined similarly. The NET module can learn to zero the activity of the Reactive DV by modifying adaptive weights,  $w_{d/a}^{n+}$ . Weight updates obey VAM learning:

$$\frac{dw_{d/a}^{n+}}{dt} = \eta f(x_{d/a}^+) [t_{d/a}^+]^+ (-v_{d/a}^+ - w_{d/a}^{n+}), \quad (90)$$

where  $\eta = 1$ ,  $v_{d/a}^+$  is the Reactive DV distance or angle field ON-channel activity, and  $x_{d/a}^+$  is the Reactive Visual TPV ON-channel activity. The activities,  $x_{d/a}^+$  and  $x_{d/a}^-$ , code the present distance and angle to the target object.  $f(w)$  is the gating signal defined by:

$$f(w) = \begin{cases} 1 & \text{if } w < 0.025 \\ 0 & \text{otherwise,} \end{cases} \quad (91)$$

which activates learning when the Reactive Visual TPV is small, or upon reaching the target object. The threshold value  $w = 0.025$  corresponds to approximately 500 distance units in world coordinates. VAM learning in (90) is also gated by NET module distance or angle ON-channel activity,  $[t_{d/a}^+]^+$ , to preclude learning if the animat has not moved. A more detailed description of the NET module is provided in Section A10.

The Motor Approach and Orienting System also processes top-down planned motor commands. Motivated WHAT and WHERE Decisions (Figures 10, 17, 18 and 20) learn to read out planned head-orienting and body-approach movements. Top-down learning is gated by the Orienting  $GO_p$  offset,  $\overline{GO}_a^p$  (equations (82) and (101)), which is activated when a head-orienting movement is completed. Top-down commands are computed in the Sensory-Motor Top-Down Readout module (Figure 20a), which can initiate head and body movements by further processing in the Motor DV module (Section A9.1). The NET module (Figure 10) inhibits activation of the Planned DV, thus ending a planned movement. The Top-down Readout Mismatch signal (Figure 10 and 20b; Section A8.1) compares the activity of the Planned DV and the NET modules. A sufficiently large discrepancy can elicit an arousal burst to select a different top-down channel from the Motivated WHAT and WHERE Decision circuit. A more detailed description of top-down readout is provided in Section A8.

**A9.1 Motor DV system.** The Motor DV system (Figure 14) computes the difference between target position and body displacement, and controls the execution of reactive and planned head-orienting and body-approach movements. Head-orienting and body-approach movements are complete when the DV reaches zero. The Motor DV system comprises a two-stage feedforward network with reactive and planned streams. The output stage is the result of three types of competitive interactions:

1. The angle field of the Reactive DV module inhibits the distance fields of the Reactive and Planned  $DV_C$  modules. This competition ensures that the animat reactively orients toward a visual target object before releasing reactive or planned body-approach.

2. The distance field of the Planned DV module inhibits the angle field of the Planned  $DV_C$  module. This interaction ensures that any planned body-approach movement is completed before a planned head-orienting movement. This enables read out of approach-orient movement pairs.

3. The distance field of the Planned  $DV_C$  module nonspecifically inhibits the distance and angle fields of the Reactive  $DV_C$  module. This enables planned body-approach commands to override reactive movement commands.

These competitive interactions between the first and second stage of the Motor DV system are codified by the following equations. The activity,  $t_d^+$ , of the Reactive  $DV_C$  distance field ON-channel obeys:

$$\frac{dt_d^+}{dt} = -t_d^+ + 2v_d^+ \left[ 1 - f(s_d^+) - f(s_a^+) - f(v_a^+) \right]^+, \quad (92)$$

where  $v_d^+$  is the Reactive DV distance field ON-channel activity (see equation (89)),  $v_a^+$  is the Reactive DV angle field ON-channel activity,  $s_d^+$  and  $s_a^+$ , are the Planned  $DV_C$  distance (d) and angle (a) field ON-channel activities, respectively (see equation (76)), and the signal function:

$$f(w) = \begin{cases} 1 & \text{if } w > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (93)$$

The activity,  $t_a^+$ , of the Reactive  $DV_C$  angle field ON-channel obeys:

$$\frac{dt_a^+}{dt} = -t_a^+ + 2v_a^+ \left[ 1 - f(s_d^+) - f(s_a^+) \right]^+, \quad (94)$$

where  $v_a^+$  is the Reactive DV angle field ON-channel activity, and the other variables are defined as in equation (93). The OFF-channel activities of the Reactive DV<sub>C</sub> module distance and angle fields,  $t_d^-$  and  $t_a^-$ , are defined similarly.

The activity,  $s_d^+$ , of the Planned DV<sub>C</sub> module distance field ON-channel obeys:

$$\frac{ds_d^+}{dt} = -s_d^+ + u_d^+ [1 - f(v_a^+)]^+, \quad (95)$$

where  $u_d^+$  is the Planned DV distance field on-channel activity (see equation (77)),  $v_a^+$  is the Reactive DV angle field ON-channel activity and the signal function,  $f(w)$ , is defined by (93).

The activity,  $s_a^+$ , of the Planned DV<sub>C</sub> module angle field ON-channel obeys:

$$\frac{ds_a^+}{dt} = -s_a^+ + u_a^+ [1 - f(u_d^+)]^+, \quad (96)$$

where  $u_a^+$  is the activity of the Planned DV angle field ON-channel,  $u_d^+$  is the Planned DV distance field ON-channel activity, and  $f(w)$ , is again defined by (93). The activities,  $s_d^-$  and  $s_a^-$ , of the Planned DV<sub>C</sub> module distance and angle field OFF-channels are defined similarly.

**A9.2 Reactive/Planned Orienting and Approach Systems.** The Reactive/Planned Orienting and Approach Systems (Figure 15) translate head-orienting and body-movement commands into outflow signals which move the head or body. This translation effects the change from sensory to motor coordinates. In particular, signals from the Reactive DV<sub>C</sub> and Planned DV<sub>C</sub> modules converge within the Reactive/Planned Orienting and Approach Systems. First, the signals are combined into a single Reactive/Planned DV<sub>S</sub>, or difference vector, in sensory coordinates. The DV<sub>S</sub> computes the difference between target position and net body displacement, thereby representing the distance to approach and angle to turn to reach the target. Then, via an innate sensory-motor transform, the Reactive/Planned DV<sub>S</sub> signals project directly to the Reactive/Planned DV<sub>M</sub> module. The Reactive/Planned DV<sub>M</sub> module output is gated by the Orienting and Approach GO<sub>P</sub> that are volitionally modulated. The gated output of the Reactive/Planned DV<sub>M</sub> module generates outflow motor commands that control stance and gait. The Reactive/Planned DV<sub>S</sub> and the Reactive/Planned DV<sub>M</sub> modules are defined in the following equations.

The activity,  $x_d^+$ , of the Reactive/Planned DV<sub>S</sub> distance field ON-channel obeys:

$$\frac{dx_d^+}{dt} = -x_d^+ + t_d^+ + s_d^+, \quad (97)$$

where  $t_d^+$  is the activity of the Reactive DV<sub>C</sub> distance field ON-channel, and  $s_d^+$  is the activity of the Planned DV<sub>C</sub> distance field ON-channel.

The activity,  $x_a^+$ , of the Reactive/Planned DV<sub>S</sub> angle field ON-channel obeys:

$$\frac{dx_a^+}{dt} = -x_a^+ + t_a^+ + s_a^+ + \Phi, \quad (98)$$

where  $t_a^+$  is the activity of the Reactive DV<sub>C</sub> angle field ON-channel,  $s_a^+$  is the activity of the Planned DV<sub>C</sub> angle field ON-channel, and  $\Phi$  is a left or right head orienting burst, from the output of the Head-Orienting Movement module (Figure 15a), computed in the Spatial Attention Module (Figure 5; Section A3). If the Spatial Attention Module, which computes a choice

between visually-guided approach and motion-induced orienting, commands the Head-Orienting Movement module to initiate a head turn, then the burst signal,  $\Phi \gg 1$ , active for the duration of the head turn, is applied to either the DV<sub>S</sub> ON-channel,  $x_a^+$ , for a left turn, or to the DV<sub>S</sub> OFF-channel,  $x_a^-$ , for a right turn. The activities,  $x_d^-$  and  $x_a^-$ , of the Reactive/Planned DV<sub>S</sub> distance and angle field OFF-channels are defined similarly.

The activity,  $y_d^+$ , of the Reactive/Planned DV<sub>M</sub> distance field ON-channel obeys:

$$\frac{dy_d^+}{dt} = -y_d^+ + GO_d^p x_a^+, \quad (99)$$

where  $x_a^+$  is the activity of the Reactive/Planned DV<sub>S</sub> distance field ON-channel, and  $GO_d^p$  is the activity of the Approach GO<sub>P</sub> (see equation (103)).

The activity,  $y_a^+$ , of the Reactive/Planned DV<sub>M</sub> angle field ON-channel obeys:

$$\frac{dy_a^+}{dt} = -y_a^+ + GO_a^p x_a^+, \quad (100)$$

where  $x_a^+$  is the activity of the Reactive/Planned DV<sub>S</sub> angle field ON-channel, and  $GO_a^p$  is the activity of the Orienting GO<sub>P</sub> (see equation (101)). The activities,  $y_d^-$  and  $y_a^-$ , of the Reactive/Planned DV<sub>M</sub> distance and angle field OFF-channels are defined similarly.

**A9.3 GO<sub>P</sub>, GO<sub>M</sub> and ERG.** Both reactive movement signals from the Visual Form and Motion System (Figure 5), and planned movement signals from the Visual and Motor Working Memory and Planning Systems (Figures 17 and 18), need to be performed at the correct times. Release of such signals is known to be under volitional control in vivo. Such volitional circuits have been modeled elsewhere (Brown, Bullock, and Grossberg, 2004; Gaudiano and Grossberg, 1991). To that end, GO signals gate the output of motor commands control the timing of the animat's movements. As seen above, these signals are also important to gate learning and to control other model functions, such as Vision or Motor Memory System updates. In the absence of either reactive or planned movement commands, an ERG, or endogenous random generator, is released to initiate exploratory random head-orienting movements.

The activity of the Orienting GO<sub>P</sub>, or  $GO_a^p$ , is defined by:

$$GO_a^p = f(t_a^+) + f(s_a^+) + GO_a^m, \quad (101)$$

where  $t_a^+$  is the activity of the Reactive DV<sub>C</sub> angle field ON-channel,  $s_a^+$  is the activity of the Planned DV<sub>C</sub> angle field ON-channel, and the signal function,  $f(\omega)$ , incorporates volitional modulation, such that:

$$f(\omega) = \begin{cases} 1 & \text{if } \omega > 0 \text{ and volitional gating is ON} \\ 0 & \text{otherwise.} \end{cases} \quad (102)$$

Volitional gating is controlled heuristically within the SOVEREIGN simulations since circuitry for volitional control is beyond the scope of this research. For simplicity, volitional input in (102) is assumed to be always on. A more sophisticated control mechanism will be used in future research. Finally,  $GO_a^m$  represents the activity of the Orienting GO<sub>M</sub>, controlled by the Head-Orienting Movement module (Figure 5). The activity of the Approach GO<sub>P</sub>, or  $GO_d^p$ , is defined by:

$$GO_d^p = \left( f(t_d^+) + f(s_d^+) \right) V^{nearby} G_d^{speed}, \quad (103)$$

where  $t_d^+$  is the activity of the Reactive DV<sub>C</sub> distance field ON-channel,  $s_d^+$  is the activity of the Planned DV<sub>C</sub> distance field ON-channel and the signal function,  $f(\omega)$ , is defined by (102). The  $V^{nearby}$  term is a visual signal which responds to the salience of the visual target object:  $V^{nearby} = 1$  if the apparent size of the visual target object is greater than 65% of the total size of the object's 2D snapshot, in either the vertical or horizontal dimension. This fixed ratio was selected by examining the behavior of the animat under controlled conditions. The activity of  $GO_d^p$  is also multiplicatively scaled by an incentive motivation speed signal,  $G_d^{speed}$ . This speed signal is sensitive to activities,  $[t_k]^+$ , of the Sensory-Drive Heterarchy, defined by (68). The general idea is that a motivated animat will increase its running speed in anticipation of reward, in order to receive the reward sooner. The speed signal,  $G_d^{speed}$ , is defined by:

$$G_d^{speed} = \begin{cases} 2 & \text{if } \left( \sum_{k=2}^3 [t_k]^+ - [t_1]^+ \right) > 0.25 \\ 1 & \text{otherwise,} \end{cases} \quad (104)$$

where  $N = 3$  is the total number of drives, and  $[t_k]^+$  is the output of the sensory-drive heterarchy. Note that  $[t_2]^+$  and  $[t_3]^+$  correspond to the consummatory drive channels, hunger and thirst, whereas  $[t_1]^+$  is the exploratory drive channel. If sufficiently large, the total consummatory activities of the Sensory-Drive Heterarchy module can double the approach speed. This increase in speed is analogous to that observed experimentally (Hull, 1932, 1934).

The ERG is modeled as an endogenously active noise source (Gaudio and Grossberg, 1991). When the ERG is disinhibited, it generates an orienting burst of activity that is sufficient to initiate a head orienting turn in a random direction. The ERG activation continues until the distance field of the Approach GO<sub>P</sub> is reactivated. The ERG obeys:

$$ERG = 1 - GO_d^p. \quad (105)$$

The ERG allows the animat to continue random exploration, even after the visual target object is reached, or when Parvo and Magno signals may be unavailable. Although not part of the current SOVEREIGN simulations, the ERG could also be used to initiate a random body approach movement.

#### **A10. Sensory-Motor NET Movement Module**

The Sensory-Motor NET Movement Module (Figure 12), learns to compute a robust representation of body displacement using a combination of visual, vestibular and motor copy signals. The primary output from this module, the NET vector, provides signals which are used to zero the Reactive and Planned DV fields, as well as for encoding by the Motor Working Memory and Planning System, shown in Figure 18. One input to this module is the vestibular Present Position Vector, or PPV<sub>V</sub>, a vector which codes the head and body displacement using vestibular signals. A second input is the motor Present Position Vector, or PPV<sub>M</sub>, a vector which also codes the head and body displacement, but uses motor outflow, or corollary discharge, signals. The PPV<sub>V</sub> and PPV<sub>M</sub>, are computed by the Motor Integrator Module (Section A10.1).

The activities,  $u_d^+$  and  $u_a^+$ , of the M-V Mismatch module distance (d) and angle (a) field ON-channels obey:

$$\frac{du_{d/a}^+}{dt} = -u_{d/a}^+ + p_{d/a}^+ w_{d/a}^{1+} - q_{d/a}^+, \quad (106)$$

where  $p_{d/a}^+$  is the activity of the PPV<sub>V</sub> distance or angle field ON-channels, obtained from the integration of vestibular feedback and  $q_{d/a}^+$  is the activity of the PPV<sub>M</sub> distance or angle field ON-channels, obtained from the integration of proprioceptive feedback. The strength of adaptive weights,  $w_{d/a}^{1+}$ , obey VAM learning:

$$\frac{dw_{d/a}^{1+}}{dt} = \eta f(p_{d/a}^+) (-u_{d/a}^+ - w_{d/a}^{1+}), \quad (107)$$

where  $\eta = 0.001$ ,  $u_d^+$  and  $u_a^+$  are the activities of the M-V Mismatch distance and angle field ON-channels,  $p_d^+$  and  $p_a^+$  are the activities of the PPV<sub>V</sub> distance and angle field ON-channels, and the signal function,  $f(\omega)$ , is defined by:

$$f(\omega) = \begin{cases} 1 & \text{if } \omega > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (108)$$

The adaptive weights,  $w_{d/a}^{1+}$ , are updated continuously, albeit with a slow learning rate. The learning rate was chosen to be large enough to calibrate differences between the PPV<sub>V</sub> and PPV<sub>M</sub>, but small enough to preclude interference from learning if vestibular signals temporarily deviate from motor copy signals. A short-term error between the signals might occur when the animat is traversing slippery terrain. Because VAM learning is employed, the activities,  $u_d^+$  and  $u_a^+$ , of the M-V Mismatch module ON-channels are almost always near zero.

The activities,  $r_d^+$  and  $r_a^+$ , of the NET<sub>MV</sub> distance and angle field ON-channels obey:

$$\frac{dr_{d/a}^+}{dt} = -r_{d/a}^+ + q_{d/a}^+ - q_{d/a}^- + 2u_{d/a}^+, \quad (109)$$

where  $q_d^+$  and  $q_a^+$  are the activities of the PPV<sub>M</sub> distance and angle field ON-channels,  $q_d^-$  and  $q_a^-$  are the activities of the PPV<sub>M</sub> distance and angle field OFF-channels and  $u_d^+$  and  $u_a^+$  are the activities of the M-V Mismatch distance and angle field ON-channels. The activities,  $r_d^-$  and  $r_a^-$ , of the NET<sub>MV</sub> distance or angle field OFF-channels are defined similarly.

The activities,  $s_d^+$  and  $s_a^+$ , of the S-MV Mismatch module distance and angle field ON-channels obey:

$$\frac{ds_{d/a}^+}{dt} = -s_{d/a}^+ + \bar{r}_{d/a}^+ w_{d/a}^{2+} - \bar{z}_{d/a}^+ [1 - GO_a^m]^+, \quad (110)$$

where  $\bar{r}_d^+$  and  $\bar{r}_a^+$  are the normalized activities of the NET<sub>MV</sub> distance and angle field ON-channels,  $\bar{z}_d^+$  and  $\bar{z}_a^+$  are the normalized activity of the NET<sub>S</sub> distance and angle field ON-channels and  $GO_a^m$  is the activity of the Magno GO. The normalization of  $\bar{r}_{d/a}^+$  and  $\bar{r}_{d/a}^-$  is defined by:

$$\bar{r}_{d/a}^+ = \frac{1+r_{d/a}^+}{2}; \quad \bar{r}_{d/a}^- = \frac{1-r_{d/a}^-}{2}. \quad (111)$$

Similarly, the normalization of  $\bar{z}_{d/a}^+$  and  $\bar{z}_{d/a}^-$  is defined by:

$$\bar{z}_{d/a}^+ = \frac{1+z_{d/a}^+}{2}; \quad \bar{z}_{d/a}^- = \frac{1-z_{d/a}^-}{2}. \quad (112)$$

This type of normalization derives from the following shunting competitive system:

$$\frac{dc}{dt} = -\varepsilon c + (1-c)(1 \pm a) - c(1 \mp a) \quad (113)$$

which defines the activity,  $c$ , that results from push-pull inputs,  $a$ . At steady-state, the activity,  $c$ , is defined by:

$$c = \frac{1 \pm a}{2}, \quad (114)$$

as in equations (111) and (112). The adaptive weights,  $w_d^{2+}$  and  $w_a^{2+}$ , obey VAM learning:

$$\frac{dw_{d/a}^{2+}}{dt} = \eta f(r_{d/a}^+) (-s_{d/a}^+ - w_{d/a}^{2+}), \quad (115)$$

where  $\eta = 0.001$ ,  $r_d^+$  and  $r_a^+$  are the activities of the NET<sub>MV</sub> distance or angle field ON-channels,  $f(\omega)$  is the signal function in (108) and  $s_d^+$  and  $s_a^+$  are the activities of the S-MV Mismatch module distance and angle field ON-channels, and. The adaptive weights,  $w_d^{2+}$  and  $w_a^{2+}$ , are updated continuously, albeit with a slow learning rate. Because VAM learning is employed, the activities,  $s_d^+$  and  $s_a^+$ , of the S-MV Mismatch module distance or angle field ON-channel are almost always near zero.

The activities,  $y_d^+$  and  $y_a^+$ , of the Reactive Visual TPV Storage module distance and angle field ON-channels obey:

$$\begin{aligned} \frac{dy_{d/a}^+}{dt} = & -A_1 y_{d/a}^+ + (1 - y_{d/a}^+) (5y_{d/a}^+ + x_{d/a}^+ (GO_d^p \text{ onset})) \\ & - y_{d/a}^+ (5y_{d/a}^- + x_{d/a}^- (GO_d^p \text{ onset})), \end{aligned} \quad (116)$$

where the decay term  $A_1 = 0.001$ . The excitatory input,  $5y_{d/a}^+ + x_{d/a}^+ (GO_d^p \text{ onset})$ , comprises two parts. First, there is strong excitatory feedback via  $5y_{d/a}^+$ . Second, there is a gated input from the activities,  $x_d^+$  and  $x_a^+$ , of the Reactive Visual TPV module distance and angle field ON-channels. The onset of the Approach GO<sub>P</sub> (Section A9.3; equation (103)), or  $GO_d^p \text{ onset}$ , marks the beginning of a body-approach movement, and expresses the Reactive Visual TPV input,  $x_{d/a}^+$ . This onset signal lasts only briefly, while the Reactive Visual TPV Storage module tracks the Reactive Visual TPV ON-channel and OFF-channel activities,  $x_{d/a}^+$  and  $x_{d/a}^-$ . After the  $GO_d^p \text{ onset}$ , the activity of the Reactive Visual TPV Storage module is maintained via strong self-feedback. The inhibitory terms,  $5y_{d/a}^- + x_{d/a}^- (GO_d^p \text{ onset})$ , also comprise two parts, which function similarly and serve to normalize the positive and negative opponent cells. The activities,

$x_d^-$  and  $x_a^-$ , of the Reactive Visual TPV Storage module distance field OFF-channels, are described similarly.

The activities,  $z_d^+$  and  $z_a^+$ , of the NET<sub>S</sub> distance and angle field ON-channels obey:

$$\frac{dz_{d/a}^+}{dt} = -z_{d/a}^+ + y_{d/a}^+ - x_{d/a}^+ w_{d/a}^{3+}, \quad (117)$$

where  $y_d^+$  and  $y_a^+$  are the activities of the Reactive Visual TPV Storage module distance and angle field ON-channels and  $x_d^+$  and  $x_a^+$  are the activities of the Reactive Visual TPV module distance and angle field ON-channels. The adaptive weights,  $w_{d/a}^{3+}$ , obey the VAM learning rule as follows:

$$\frac{dw_{d/a}^{3+}}{dt} = \eta f(x_{d/a}^+) (GO_d^p \text{ onset}) (-z_{d/a}^+ - w_{d/a}^{3+}), \quad (118)$$

where  $\eta = 1$ ,  $z_d^+$  and  $z_a^+$  are the activities of the NET<sub>S</sub> module distance and angle field ON-channels,  $x_d^+$  and  $x_a^+$  are the activities of the Reactive Visual TPV module distance and angle field ON-channels,  $f(\omega)$  is the signal function in (108), and  $GO_d^p \text{ onset}$  is active at the onset of the Approach GO<sub>P</sub> (Section A9.3; equation (103)). Since fast VAM learning,  $\eta = 1$ , is employed, the activity of the distance and angle field of the NET<sub>S</sub> module is zero shortly after the onset of the Approach GO<sub>P</sub>,  $GO_d^p \text{ onset}$ , gates learning.

Fast VAM learning is used to calibrate target position information in the NETs module (Figure 12) for the following reason: After calibration, the NETs module has zero activity, but is gradually disinhibited as the animat approaches the target. A high learning rate enables calibration to be completed before the movement commences. Fast learning is also used in the Top-down Readout module to learn the NET body displacement before the start of a new movement.

In contrast, slow VAM learning is used in the S-MV Mismatch module to calibrate differences between visual and proprioceptive signals. These differences are typically small, thus easily handled by slow learning. A short-term error might occur if visual signals were shut off entirely (i.e., in the dark). The M-V Mismatch module accomplishes a similar calibration between vestibular and proprioceptive signals. Again, slow learning can maintain calibrated adaptive weights, which are unaffected by short-term deviations in input (i.e., on slippery terrain).

The combined activities,  $v_d^+$  and  $v_a^+$ , of the distance and angle field of the NET<sub>S</sub> and S-MV Mismatch ON-channels are defined by:

$$\frac{dv_{d/a}^+}{dt} = -v_{d/a}^+ + 2z_{d/a}^+ [1 - GO_a^m]^+ + 2s_{d/a}^+, \quad (119)$$

where  $z_d^+$  and  $z_a^+$  are the activities of the NET<sub>S</sub> distance and angle field ON-channels,  $GO_a^m$  is the activity of the Magno GO and  $s_d^+$  and  $s_a^+$  are the activities of the S-MV Mismatch distance and angle field ON-channels. The activities of the NET module distance and angle field ON-channels obey:

$$\frac{dt_{d/a}^+}{dt} = -A_2 t_{d/a}^+ + (1 - t_{d/a}^+) (v_{d/a}^+ + T) - t_{d/a}^+ (v_{d/a}^- + T), \quad (120)$$

where  $v_d^+$  and  $v_a^+$  are the activities of the combined NET<sub>S</sub> and S-MV Mismatch distance and angle field ON-channels, and  $T=1$ , is a tonically-active input cell (Figure 12). The resulting normalized ON-channel activities,  $t_{d/a}^+$ , and OFF-channel activities,  $t_{d/a}^-$ , (see equations (113) and (114)) are defined by:

$$\bar{t}_{d/a}^+ = \frac{1+t_{d/a}^+}{2}; \quad \bar{t}_{d/a}^- = \frac{1-t_{d/a}^-}{2}. \quad (121)$$

**A10.1 Motor integrators.** The Motor Integrator module (shown at the bottom of Figure 12), computes the activities,  $p$ , of the PPV<sub>V</sub>, or vestibular Present Position Vector, a measure of body displacement in vestibular coordinates, and the activities,  $q$ , of the PPV<sub>M</sub>, a similar signal based on proprioceptive, or motor copy, signals. A heuristic implementation of an integrator with linear response over a long time period is assumed, and implemented using a counter, which adds the head turn angle and body approach distance after each step. These angle and distance inputs to the vestibular (F) and proprioceptive (C) counters are also emulated heuristically. The distance or angle fields of the vestibular and motor copy integrators are reset by a burst of activity at the onset of the Approach GO<sub>P</sub>, or  $GO_d^p$  onset (Section A9.3; equation (103)).

## References

- Baloch, A. A. and Waxman, A. M. (1991). Visual learning, adaptive expectations, and behavioral conditioning of the mobile robot MAVIN. *Neural Networks*, 4, 271–302.
- Barto, A. G. and Sutton, R. S. (1981). Landmark learning: An illustration of associative search. *Biological Cybernetics*, 42, 1–8.
- Berzhanskaya, J., Grossberg, S. and Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, in press.
- Blodgett, H. C. (1929). An analysis of the behavior units in maze learning. *Proceedings of the 9th International Congress on Psychology*, 9, 85–86.
- Bradski, G., Carpenter, G. A., and Grossberg, S. (1992). Working memory networks for learning temporal order with application to three-dimensional visual object recognition. *Neural Computation*, 4, 270–286.
- Bradski, G., Carpenter, G.A., and Grossberg, S. (1994). STORE working memory networks for storage and recall of arbitrary temporal sequences. *Biological Cybernetics*, 71, 469-480.
- Bradski, G. and Grossberg, S. (1995). Fast learning VIEWNET architectures for recognizing 3-D objects from multiple 2-D views. *Neural Networks*, 8, 1053-1080.
- Bullock, D., Cisek, P. and Grossberg, S. (1998). Cortical networks for control of voluntary arm movements under variable force conditions. *Cerebral Cortex*, 8, 48-62.
- Bullock, D. and Grossberg, S. (1988). Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review*, 95, 49-90.
- Bullock, D. and Grossberg, S. (1991). Adaptive neural networks for control of movement trajectories invariant under speed and force rescaling. *Human Movement Science*, 10, 3-53.
- Bullock, D., Grossberg, S., and Guenther, F. H. (1993). A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm. *Journal of Cognitive Neuroscience*, 5, 4, 408–435.
- Burgess, N., Recce, M., and O’Keefe, J. (1995). Hippocampus – Spatial models. In *The Handbook of Brain Theory*. Bradford Books / MIT Press, Cambridge, MA.
- Buxton, C. E. (1940). Latent learning and the goal gradient hypothesis. *Contr. Psychol. Theor.*, 2, 2, 75.
- Cao, Y. and Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. *Spatial Vision*, 18, 515-578.
- Carpenter, G. A. and Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, 37, 54–115.
- Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J.H., and Rosen, D.B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, 3, 698-713.
- Carpenter, G. A., Grossberg, S., and Rosen, D. B. (1991). Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, 4, 759–771.
- Carpenter, G.A. and Ross, W.D. (1993). ART-EMAP: A neural network architecture for learning and prediction by evidence accumulation. *Proceedings of the World Congress on Neural Networks, (WCNN- 93)*, III-649-656. Technical Report CAS/CNS-TR-93-015, Boston University.

- Carr, H. and Watson, J. B. (1908). Orientation of the white rat. *Journal of Comparative Neurology and Psychology*, 18, 27–44.
- Cohen, M.A. and Grossberg, S. (1986). Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short-term memory. *Human Neurobiology*, 5, 1-22.
- Cohen, M. A. and Grossberg, S. (1987). Masking fields: A massively parallel neural architecture for learning, recognizing, and predicting multiple groupings of patterned data. In Grossberg, S. (Ed.), *Neural Networks and Natural Intelligence*, Chapter 7, pp. 317–367. The MIT Press, Cambridge, MA.
- Contreras-Vidal, J.L., Grossberg, S., and Bullock, D. (1997). A neural model of cerebellar learning for arm movement control: Cortico-spino-cerebellar dynamics. *Learning and Memory*, 3, 475-502.
- Dallal, N. L. and Meck, W. H. (1990). Hierarchical structures: Chunking by food type facilitates spatial memory. *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 1, 69–84.
- Daub, C. T. (1933). The effect of doors on latent learning. *Journal of Comparative Psychology*, 15, 49–58.
- Dayan, P. (1987). Navigating through temporal difference. In *Neural Information Processing Systems*, 3, pp. 464–470. Morgan Kaufmann Publishers, San Mateo, CA.
- Dennis, W. (1932). Multiple visual discrimination in the block elevated maze. *Journal of Comparative and Physiological Psychology*, 13, 391–396.
- Fang, L. and Grossberg, S. (2007). From stereogram to surface: How the brain sees the world in depth. Technical Report CAS/CNS TR-06-013, Boston University. Submitted for publication.
- Fazl, A., Grossberg, S., and Mingolla, E. (2007). View-invariant object category learning, recognition and search: How spatial and object attention are coordinated using surface-based attentional shrouds. Technical Report CAS/CNS TR-07-011, Boston University. Submitted to *Cognitive Psychology*, April, 2007.
- Fiala, J.C., Bullock, D., Grossberg, S. (1996). Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye blink response. *Journal of Neuroscience*, 16, 3760-3774.
- Fuhs, M.C. and Touretzky, D.S. (2006). A spin glass model of path integration in rat medial entorhinal cortex. *Journal of Neuroscience*, 26, 4266 – 4276.
- Gaudio P. and Grossberg S. (1991). Vector associative maps: Unsupervised real-time error-based learning and control of movement trajectories. *Neural Networks*, 4, 147-183.
- Gilbert, P. F. C. and Thach, W. T. (1977). Purkinje cell activity during motor learning. *Brain Research*, 128, 309-328.
- Gnadt, W. and Grossberg, S. (2005a). SOVEREIGN: *A Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation system*. In Douglas Blank and Lisa Meeden (Eds.), *Developmental Robotics: Papers from the 2005 Spring Symposium*, March 21-23, pp. 106-110. American Association for Artificial Intelligence, Menlo Park, California.
- Gnadt, W. and Grossberg S. (2005b). *SOVEREIGN: A Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation system*. Dynamical Neuroscience XIII: Computational Cognitive Neuroscience satellite symposium of the annual Society for Neuroscience meeting, November 10-11, 2005.

- Gnadt, W. and Grossberg S. (2006). *SOVEREIGN: A Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation system*. Tenth International Conference on Cognitive and Neural Systems, Boston, Massachusetts, May 2006.
- Goodale, M. A. and Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15, 20–25.
- Gorchetchnikove, A. and Grossberg, S. (2007). Space, time, and learning in the hippocampus: How fine spatial and temporal scales are expanded into population codes for behavioral control. *Neural Networks*, 20, 182 – 193.
- Greve, D., Grossberg, S., Guenther, F., and Bullock, D. (1993). Neural representations for sensory–motor control, I: Head–centered 3–D target positions from opponent eye commands. *Acta Psychologica*, 82, 115–138.
- Grice, G. R. (1948). An experimental test of the expectation theory of learning. *Journal of Comparative and Physiological Psychology*, 41, 137–143.
- Grossberg, S. (1972). A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences*, 15, 253-285.
- Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen and F. Snell (Eds.), *Progress in theoretical biology, Volume 5*. New York: Academic Press, pp. 233-374.
- Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, 87, 1-51.
- Grossberg, S. (1982a). *Studies of Mind and Brain: Principles of Learning, Perception, Development, Cognition and Motor Control*. D. Reidel Publishing Co., Boston, MA.
- Grossberg, S. (1982b). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. *Psychological Review*, 89, 529-572.
- Grossberg, S. (1984a). 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. (1984b). Some normal and abnormal behavioral syndromes due to transmitter gating of opponent processes. *Biological Psychiatry*, 19, 1075–1118.
- Grossberg, S., Guenther, F., Bullock, D., and Greve, D. (1993). Neural representations for sensory-motor control, II: Learning a head-centered visuomotor representation of 3-D target position. *Neural Networks*, 6, 43-67.
- Grossberg, S. (2000a). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences*, 4, 233-246.
- Grossberg, S. (2000b). The imbalanced brain: From normal behavior to schizophrenia. *Biological Psychiatry*, 48, 81-98.
- Grossberg, S., Guenther, F., Bullock, D., and Greve, D. (1993). Neural representations for sensory–motor control, II: Learning a head–centered visuomotor representation of 3–D target position. *Neural Networks*, 6, 43–67.
- Grossberg, S. and Merrill, J.W.L. (1992). A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research*, 1, 3-38.
- Grossberg, S. and Merrill, J.W.L. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience*, 8, 257-277.
- Grossberg, S., Mingolla, E. and Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, 41, 2521-2553.

- Grossberg, S. and Myers, C.W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, 107, 735-767.
- Grossberg, S. and Seidman, D. (2006). Neural dynamics of autistic behaviors: Cognitive, emotional, and timing substrates. *Psychological Review*, 113, 3, 483-525.
- Grossberg, S. and Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research*, 45, 1725-1743.
- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102, 594-621.
- Guenther, F. H., Bullock, D., Greve, D., and Grossberg, S. (1994). Neural representations for sensory-motor control, III: Learning a body-centered representation of 3-D target position. *Journal of Cognitive Neuroscience*, 6, 341 - 358.
- Hafting, T., Fyhn, M., Molden, S., Moser, M.H., and Moser, E.I. (2005). Microstructure of the spatial map in the entorhinal cortex, *Nature*, 436, 801 – 806.
- Haney, G. W. (1931). The effect of familiarity on maze performance of albino rats. *University of California Publications on Psychology*, 4, 319-333.
- Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review*, 39, 25-43.
- Hull, C. L. (1933). Differential habituation to internal stimuli in the albino rat. *Journal of Comparative Psychology*, 16, 255-273.
- Hull, C. L. (1934). The rat's speed of locomotion gradient in the approach of food. *Journal of Comparative Psychology*, 17, 393-422.
- Ito, M. (1984). *The cerebellum and neural control*. New York: Raven Press.
- Kelly, F.J. and Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Perception and Psychophysics*, 62, 1596-1619.
- Kendler, H. H. (1946). The effect of change of drive on t-maze performance. *American Psychologist*, 1, 246.
- Kendler, H. H. (1947a). A comparison of learning under motivated and satiated conditions in the white rat. *Journal of Experimental Psychology*, 37, 545-549.
- Kendler, H. H. (1947b). An investigation of latent learning in the t-maze. *Journal of Comparative and Physiological Psychology*, 40, 265-270.
- Leeper, R. (1935). The role of motivation in learning: A study of the phenomenon of differential motivational control of the utilization of habits. *Journal of Genetic Psychology*, 46, 3-40.
- Leonard, B. and McNaughton, B. L. (1990). Spatial representation in the rat: Conceptual, behavioral, and neurophysiological perspectives. In Kesner, R. P. and Olton, D. S. (Eds.), *Neurobiology of Comparative Cognition*, Chapter 13, pp. 363-422. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London, Series B*, 298, 85-95.
- Mishkin, M. (1993). Cerebral memory circuits. In T. A. Poggio and D. A. Glaser (Eds.), *Exploring brain functions: Models in neuroscience*, pp. 113-125. New York: Wiley.
- Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6, 414-417.
- Munn, N. L. (1950). *Handbook of Psychological Research on the Rat*. Houghton Mifflin Co., Boston, MA.

- O'Keefe, J.M. and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34, 171 – 175.
- O'Keefe, J. and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford University Press, New York, NY.
- Pribe, C., Grossberg, S., and Cohen, M.A. (1997). Neural control of interlimb oscillations, II: Biped and quadruped gaits and bifurcations. *Biological Cybernetics*, 77, 141-152.
- Raizada, R. and Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, 13, 100-113.
- Reynolds, B. (1945). A repetition of the blodgett experiment on “latent learning”. *Journal of Experimental Psychology*, 35, 504–516.
- Russell, J. T. (1932). Depth discrimination in the rat. *Journal of Genetic Psychology*, 40, 136–159.
- Saltzman, E. (1979). Levels of sensorimotor representation. *Journal of Mathematical Psychology*, 20, 91–163.
- Schalkoff, R. J. (1989). *Digital Image Processing and Computer Vision*. John Wiley and Sons, Inc., New York, NY.
- Schilling, R. J. (1990). *Fundamentals of Robotics: Analysis and Control*. Prentice Hall, Englewood Cliffs, NJ.
- Schmajuk, N. A. (1990). Role of the hippocampus in temporal and spatial navigation: An adaptive neural network. *Behavioral Brain Research*, 39, 205–229.
- Spence, K. W. and Lippett, R. (1946). An experimental test of the sign–gestalt theory of trial and error learning. *Journal of Experimental Psychology*, 36, 491–502.
- Squire, L. R. and Cohen, N. J. (1984). Human memory and amnesia. In G. Lynch, J. McGaugh, and N. M. Weinberger (Eds.), *Neurobiology of learning and memory*, pp. 3-64. New York: Guilford Press.
- Stoltz, S. P. and Lott, D. F. (1964). Establishment in rats of a persistent response producing net loss of reinforcement. *Journal of Comparative and Physiological Psychology*, 57.
- Thompson, R. F. (1988). The neural basis of basic associative learning of discrete behavioral responses. *Trends in Neurosciences*, 11, 152-155.
- Tolman, E. C. and Honzik, C. H. (1930). “Insight” in rats. *University of California Publications in Psychology*, 4, 14, 215–232.
- Ungerleider, L. G. and Mishkin, M. (1982). Two cortical visual systems: Separation of appearance and location of objects. In Ingle, D. L., Goodale, M. A., and Mansfield, R. J. W. (Eds.), *Analysis of Visual Behavior*, pp. 549–586. The MIT Press, Cambridge, MA.
- Walker, E. L. (1948). Drive specificity and learning. *Journal of Experimental Psychology*, 38, 39–49.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8, 338-353.