Neuromodulation as a Robot Controller: A Brain Inspired Strategy for Controlling Autonomous Robots

Brian R. Cox and Jeffrey L. Krichmar
Department of Cognitive Sciences
University of California, Irvine
Irvine, CA 92697-5100

Introduction

We present a strategy for controlling autonomous robots that is based on principles of neuromodulation in the mammalian brain. Neuromodulatory systems signal important environmental events to the rest of the brain, causing the organism to focus its attention on the appropriate object, ignore irrelevant distractions, and respond quickly and appropriately to the event [1]. There are separate neuromodulators that alter responses to risks, rewards, novelty, effort, and social cooperation. Moreover, the neuromodulatory systems provide a foundation for cognitive function in higher organisms; Attention, emotion, goal-directed behavior, and decision-making all derive from the interaction between the neuromodulatory systems, and brain areas such as the amygdala, frontal cortex, and hippocampus. Therefore, understanding neuromodulatory function may provide control and action selection algorithms for autonomous robots that effectively interact with the environment.

Neuromodulatory Systems

Neuromodulators are chemical transmitters in the brain that can have a strong and lasting effect on an animal's behavior. The neuromodulatory systems include noradrenergic, serotonergic, dopaminergic, and cholinergic projections from below the cerebral cortex to broad areas of the central nervous system [2]. The origins of these systems are small pools of neurons (on the order of thousands in the rodent and tens of thousands in the human) located below the cortex.

Despite the different origination and chemical signatures of these neuromodulatory systems, there are several commonalities among them:

1. Each of these neuromodulatory systems originates below the cerebral cortex and projects broadly to all regions of the brain.
2. Each of these neuromodulatory systems is reciprocally connected with cognitive areas of the brain such as the amygdala, frontal cortex and the hippocampus [2].
3. The effect of each of these neuromodulatory systems on downstream neuronal targets is similar. That is, they cause target neural networks to sharpen, resulting in a winner-take-all response [1, 3, 4].

A computational framework for applying neuromodulatory systems to the control of autonomous robots can be based on the following premises:

1. The common effect of the neuromodulatory systems is to drive an organism to be decisive when environmental conditions call for such actions, and to allow the organism to be more exploratory when there are no pressing events [1, 5].
2. The main difference between neuromodulatory systems is the environmental stimuli that activate them. The serotonergic system responds to risks and threats [6], the cholinergic system sets a level of attentional effort [7], the dopaminergic system drives reward anticipation [8], and the noradrenergic system responds to novel and salient objects [9].

From the evidence, it appears that the common effect of the neuromodulatory system is to focus attention on important objects in the environment by increasing the signal to noise ratio of neuronal responses [1, 5]. Indeed, the major targets of the neuromodulators are areas noted for driving behavior, conditioning responses, focusing attention, and making decisions [2]. The means by which neuromodulatory systems focus an animal's attention is through short bursts of activity in response to important events occurring in its surroundings. During phasic neuromodulation, information from sensory systems (e.g. visual, auditory, etc) is amplified relative to recurrent or associational information [1, 3, 4]. The result of this change in the relative weighting of information is to sharpen responses to environmental input, increase signal to noise ratio, and drive decisive responses in neural networks. Moreover, neuromodulation gates in learning such that an animal can predict relationships between sensory information and action outcomes [10, 11].

A control system for a robot, which is designed according to principles of the neuromodulatory system, could offer major advantages over conventional systems in carrying out tasks in the face of environmental challenges. Such a system could learn to take appropriate actions depending on context, environmental change, and experience. Neuromodulatory systems drive many of the fundamental behaviors crucial for an organism's survival. Cognitive functions such as attention, emotion, goal-directed behavior, and decision-making all arise from the interaction between neocortical "executive" areas and the neuromodulatory systems. Therefore a controller based on the action of neuromodulation could have much to offer the design of autonomous robots.

In this paper, we show in a neural model how bursts of cholinergic, dopaminergic and serotonergic activity can sharpen attention, and lead toward appropriate action selection in a cognitive robot. The robot's behavior is guided by a simulation, which has groups of neurons and synaptic connections between these neurons, based on known dynamical and anatomical properties of the neuromodulatory system and its interaction with surrounding brain regions. Although this neurorobot will be used to investigate how neuromodulation can lead to adaptive behavior, principles of this cognitive system may be relevant for the control of robots in general.

Methods

Robot and Experimental Apparatus

The robot used for the experiments, CARL-1, was constructed in the Cognitive Anteater Robotics Laboratory at University of California, Irvine (see Figure 1A). It consisted of a two wheeled mobile base equipped with a CCD video camera having a RF transmitter for vision, IR sensors for obstacle avoidance, and a WiFi device server (http://www.sena.com) for communication between the robot and a computer workstation. The pan and tilt position of the camera was controlled by commands to a pair of servomotors. The base of the robot was 10 inches in diameter and 8.5 inches high. A distributed network of PIC-18F2680 microcontrollers, which communicated over a CAN interface, read from CARL-1's sensors, controlled CARL-1's actuators, and communicated wirelessly with a computer workstation that contained the neural simulation. Camera video frames were transmitted wirelessly to the Firewire port of the workstation.

CARL-1's environment consisted of a 10-foot by 10-foot enclosure that contained eight light panels built into the flooring (see Figure 1B). The color of the panels at the four corners were set to Cyan, Green, Magenta, and Red at a given frequency and duration through RS-232 communication from the workstation to electronics controlling the panels. All eight panels had IR transceivers that could communicate position information to CARL-1 when it was on top of the panel.

Neural Architecture

The neural simulation that controlled CARL-1 consisted of a visuomotor area, neuromodulatory systems, action areas and behavior drivers (Figure 2). The visuomotor area consisted of sub-areas, each with 15x20 (height x width) neurons that mapped on CARL-1's field of view. These retinotopically mapped neurons responded preferentially to cyan, green, magenta, and red. The simulated neuromodulatory systems consisted of a cholinergic basal forebrain (BF) area, a serotonergic raphe nucleus (Raphe), and a dopaminergic ventral tegmental area (VTA). Each of these neuromodulatory areas contained 100 neurons. The action areas consisted of a Find and Flee area that each contained 100 neurons. The behavior driver areas consisted of a Good and Bad area that each contained 100 neurons.

Neural areas were connected through synaptic projections consisting of probability distributions of connectivity between individual neurons. Neurons in the visuomotor areas had a 10% chance of being connected to neuromodulatory neurons, and an initially weak weight (uniformly distributed between 0.05 and 0.10) that could change through experiential plasticity. Within a visuomotor subarea (e.g. Red→Red), neurons connected to neighboring neurons with a 2-dimensional Gaussian distribution having a standard deviation of 5 neurons, and an initial weight uniformly distributed between 0.8 and 1.0. Between visuomotor areas (e.g. Red→Green), neurons had a 25% chance of being connected with initial weights uniformly distributed from 0.8 to 1.0 for excitatory connections and from -0.8 to -1.0 for inhibitory connections. The behavior driver neurons had strong "all-to-all" connections to the neuromodulatory systems and the action areas. Specifically, the Good neurons had excitatory connections to VTA and Find neurons with weights set to 200, and inhibitory connections to Raphe and Flee neurons with weights set to -200. Conversely, the Bad neurons had excitatory connections to Raphe and Flee neurons with weights set to 200, and inhibitory connections to VTA and Find neurons with weights set to -

200. VTA neurons projected "all-to-all" to Find neurons, Raphe neurons projected "all-to-all" to Flee neurons, and BF neurons projected "all-to-all" to Raphe and Flee neurons with initial weights uniformly distributed from 0.1 to 0.2.

Visual neurons were set based on input from CARL-1's camera. The OpenCV library (http://sourceforge.net/projects/opencv/) was used to sub-sample the image to 30x40 pixels and run color histogram filters across the image to create separate subareas that responded preferentially to cyan, green, magenta, and red. These responses, which were normalized between 0 and 1, were used to activate visuomotor neurons of corresponding colors. These visual responses were connected topographically to visuomotor neurons with a 2-dimensional Gaussian distribution having a standard deviation of 5 neurons, and a weight uniformly distributed from 1.0 to 1.5.

Neuronal Dynamics and Synaptic Plasticity

Neural activity in CARL-1 was simulated by a mean firing rate neuron model where the firing rate of each neuron ranged continuously from 0 (quiescent) to 1 (maximal firing). The activity level of a neuron represented its average firing rate over 100ms. This model demonstrated the necessary neural dynamics, and was efficient enough to run in real-time on a robotic platform with sensors and actuators. The equation for the mean firing rate neuron model was:

$$s_i(t) = \rho_i s_i(t-1) + (1-\rho_i)\left(\frac{1}{1+\exp(-0.1I_i(t))}\right) \quad (1)$$

where t was the current time step, s_i was the activation level of neuron i, ρ_i was the persistence of the neuron, and I_i is the synaptic input. Visuomotor neurons had a persistence of 0.5 and all other neurons had a persistence of 0.1.

The synaptic input of the neuron was based on pre-synaptic neural activity, the connection strength of the synapse, and the amount of neuromodulator activity:

$$I_i(t) = \sum_j nm(t-1)w_{ij}(t-1)s_j(t-1) \quad (2)$$

where w_{ij} is the synaptic weight from neuron j to neuron i, and nm is the level of neuromodulator at synapse ij.

To simulate the effect of phasic neuromodulation, inhibitory inputs and extrinsic inputs were amplified relative to the overall neuromodulatory activity (i.e. nm was set to be ten times

the combined average activity of the simulated BF, Raphe, and VTA neural areas). Connections from the visual input neurons to visuomotor neurons, from neuromodulatory neurons to action neurons, and within the neuromodulatory systems were considered extrinsic. All other excitatory connections were considered intrinsic, and for those connections, *nm* was always equal to 1.

Connections from the visuomotor areas (Cyan, Green, Magenta, and Red) to the neuromodulatory areas (BF, Raphe, VTA), and from the visuomotor areas to the action areas (Find and Flee) were subject to synaptic plasticity that depended on the current activity of the pre-synaptic neuron, the post-synaptic neuron and the overall activity of the neuromodulatory systems.

$$\Delta w_{ij}(t) = \varepsilon(w_{ij}(0) - w_{ij}(t-1)) + \delta\Theta_{NM}(nm)s_j(t-1)(s_i(t-1) - \Theta_{BCM}) \quad (3)$$

where ε was the decay rate, which was set to 0.00001 that decayed weights back to their original value ($w_{ij(0)}$). This decay acted as a slow forgetting function and prevented over learning. δ was a learning rate set to 0.001, Θ_{NM} was a gating function in which learning only occurred when the level of neuromodulator activity (*nm* from equation 2) was greater than a threshold value, which was set to 2, and Θ_{BCM} was a sliding threshold dictating the amount of synaptic potentiation and depression. The BCM threshold changed as a function of post-synaptic neural activity [12].

$$\Delta\Theta_{BCM} = 0.001(s_i(t)^2 - \Theta_{BCM}) \quad (4)$$

Action Selection and Behavior

CARL-1's behavior switched between three states: random exploration, orienting and approaching objects of interest (*Find*), and moving away from noxious objects (*Flee*). By default, CARL-1 explored unless the difference between the average activity of the Find and Flee neural areas was greater than a threshold of 0.75, in which case the more active area would elicit the corresponding behavior.

During exploration behavior, CARL-1 would move at a constant speed while panning its camera to the left and right. CARL-1's turning rate was proportional to the camera pan position. That is, the further the camera was panned from the midline, the higher the turning rate in that direction.

During Find and Flee behavior, CARL-1 would saccade its camera to the centroid of the most salient object within its field of vision. The most salient object was chosen by applying a Softmax function to the activity of the four visuomotor areas (Cyan, Green, Magenta, Red):

$$p_c = \frac{\exp(5a_c)}{\sum_{i=1}^{4}\exp(5a_i)} \tag{5}$$

where p_c is the probability of choosing color c, a_c is the average activity of visuomotor area c, and a_i is the average activity of visuomotor area i. The Softmax function was applied every time step when CARL-1 was in Find or Flee behavior. Because there were always visual stimuli in its field of view, CARL-1 would inevitably find some object in the environment to point its camera at during Find and Flee behaviors. The camera's pan and tilt position was set to the centroid of activity for the chosen color area.

During Find behavior, CARL-1 would saccade its camera to the centroid of the most salient object within its field of vision, and orient toward that object. CARL-1's wheel velocity was proportional to the camera tilt position, that is, the lower the tilt position the slower the forward velocity. CARL-1's wheel turning rate was proportional to the camera pan position causing it to orient towards the visual target (e.g. if the camera was panned left, the wheel commands turned CARL-1 to the left). This had the behavioral effect of first fixing CARL-1's gaze on a target of interest, followed by turning the body toward the target, approaching the target, and then slowing down when close to the target.

During Flee behavior, CARL-1 would saccade its camera to the centroid of the most salient object within its field of vision, but move away from that object. The camera's pan and tilt position, which was set to the centroid of activity for the chosen color area, was used to calculate wheel commands that were, in essence, the opposite of the Find motor commands. CARL-1's turning rate was proportional to the camera pan position but in the opposite direction of the target (e.g. if the camera was panned left, the wheel commands turned CARL-1 to the right), and its velocity was inversely proportional to the camera tilt position. That is, the lower the tilt position the faster the reverse velocity. This had the behavioral effect of first fixing the gaze on a target of interest, stopping forward progress, and then turning and backing away from the target.

Simulation Computation

The neural simulation contained 6,700 neurons and roughly 1.3 million synaptic connections. The neural simulation was run on a 2xQuad-Core 2.8 GHz Intel Xeon Mac Pro Workstation on the OS X operating system using POSIX threads, OpenCV, and the FLTK Graphical User Interface (GUI) library. The simulation cycle was fixed at 100 milliseconds. During each simulation cycle, the sensor data was read and processed, the neural activations were calculated, the change in the strength of plastic connections were calculated, behavior was selected, motor commands were sent to CARL-1, and the behavioral and neural data were logged for post-experimental analysis.

Experimental Paradigm

CARL-1 was first trained to associate the color green with Find behavior, and red with Flee behavior, and then tested under various conditions.

In the training period, CARL-1 explored the environment. Occasionally, when CARL-1 was near a color panel, the operator would press a button on the GUI that would either maximally activate the Good area (see Figure 2) and turn the light panel to green, or maximally activate the Bad area (see Figure 2) and turn the light panel to red. The button would be turned off after several seconds. Training would continue in this manner until CARL-1 had experienced 10 Good and 10 Bad events.

In the testing period, CARL-1 explored its environment for 7500 simulation cycles. The four light panels were set such that each light panel had a different color. Every 8 to 10 seconds, the location of the four colors was changed randomly.

Results

Training and testing were repeated with ten different CARL-1 “subjects”. Each subject consisted of the same physical device, but possessed a unique simulated nervous system differing at the level of synaptic connections. These differences among subjects were a consequence of random draws from probability distributions of connectivity, and the variation of initial connection strengths between those neurons. However, the overall neural architecture was similar across all subjects.

Behavioral Results

After training, all ten subjects responded to green stimuli with Find behavior and to red stimuli with Flee behavior (see Figures 3 and 4). During Find behavior, CARL-1 would begin its approach to the green light panel from several feet away (see Figure 3A, Top Left). As it neared the light panel, its camera tilted down and it stopped on the panel (see Figure 3A, Top Middle). As soon as the panel changed to a neutral value color, its camera tilted up and CARL-1 shifted to exploratory behavior (see Figure 3A, Top Right). During Flee behavior, CARL-1 would stop its approach when it saw the red light panel from several feet away (see Figure 3B, Top Left). With its camera centered on the red panel, CARL-1 turned away from the threatening stimulus (see Figure 3B, Top Middle). After turning completely away from the red panel, CARL-1's gaze moved away from the salient object and it shifted to exploratory behavior (see Figure 3B, Top Right).

These behaviors were driven by phasic bursts of activity from the neuromodulatory systems. Green stimuli caused a phasic response in the VTA neurons resulting in an amplification of the green visuomotor area, a dampening of distracter colors, and a strong increase in the Find activity. For example, in the bottom of Figure 3A, different neural activities are shown just prior to and during Find behavior. When the light panel switched from Red to Green (see Figure 3A, Bottom Left), the Raphe and Flee neural areas were still active and in competition with the VTA and Find areas. However, a burst of VTA activity amplified Green and Find activity, causing a suppression of Raphe, Red and Flee activity (see Figure 3A, Bottom Right). In the bottom of Figure 3B, neural activities are shown prior to and during Flee behavior. Just prior to Flee behavior, there was moderate activity throughout the neural simulation (see Figure 3B, Bottom Left). The red area has slightly elevated activity, but it was not much more active than other color areas. Moments later, a burst of Raphe activity amplified Red and Flee neuronal responses, and caused a suppression neural activity in other visuomotor, neuromodulatory, and action areas (see Figure 3B, Bottom Right).

To further test the necessity of phasic responses in the neuromodulatory systems to generate appropriate behavioral responses, we conducted simulated lesion experiments in all subjects. In one set of experiments, the activity of neurons in the Raphe area were set to zero, and in another set of experiments, the activity of neurons in the VTA area were set to zero. These lesion groups were compared with a control group that had a complete neural simulation.

Lesions of the VTA significantly reduced the number of Find responses ($p < 0.0005$, Wilcoxon Rank Sum test; see Figure 4, left), but not the Flee responses (see Figure 4 right). Lesions of the Raphe significantly reduced the number of Flee responses ($p < 0.0005$, Wilcoxon Rank Sum test; see Figure 4, right), but not the Find responses (see Figure 4, left).

The basal forebrain is thought to increase attentional effort in challenging conditions. Therefore, we lesioned the BF alone and in conjunction with lesions of other neuromodulatory areas to better understand its functional role. Lesions of the basal forebrain area alone did not have a significant effect on behavior (see BF in Figure 4). However, a lesion of basal forebrain and VTA completely abolished the Find behaviors (see BF+VTA in Figure 4 left), and a lesion of basal forebrain and Raphe completely abolished the Flee behaviors (see BF+Raphe in Figure 4 right).

Effect of Phasic Neuromodulatory Responses on Neuronal Activity

Phasic neuromodulatory activity is thought to increase the signal to noise ratio (SNR) in neural circuits such that the organism increases the discrimination between salient and non-salient stimuli. To test this idea, we calculated a SNR metric based on the visuomotor area's response to a target color divided by the visuomotor area's response to all other colors during Find and Flee behavior.

$$SNR = \frac{vis_{tgt}}{\sum_{i=1}^{4} vis_i}; \tag{6}$$

where vis_{tgt} is the average activity of the green area during a Find behavior and the average activity of red during a Flee behavior, and vis_i is the average activity of visuomotor area i.

Lesions of neuromodulatory responses significantly lowered the SNR in the visuomotor area during behavioral responses. The SNR was significantly lower in the group with VTA lesions than in the Control group during Find behavior ($p << 0.0001$, t-test; see Figure 5, left). A lesion of BF and VTA further reduced the SNR for Find responses ($p << 0.0001$, t-test comparing VTA lesion to BF+VTA lesion; see Figure 5, left). The SNR was significantly lower in the group with Raphe lesions than in the control group for Flee behavior ($p << 0.0001$, t-test; see Figure 5, right). Lesions of both the BF and Raphe further reduced the SNR for Find

responses ($p << 0.0001$, t-test comparing Raphe lesion to BF+Raphe lesion; see Figure 5, right). The other comparisons were not significantly different ($p > 0.01$; t-test).

The responses of the neuromodulatory systems were strongly correlated with colors that predicted value. Red predicted threatening stimuli and Raphe activity increased in the presence of red (see Figure 6A). Green predicted positive valence stimuli and the VTA activity increased in the presence of green (see Figure 6B). While neuromodulatory responses increased with colors that predicted value, their responses decreased for colors that were value-independent and that were predicted a value not associated with a particular neuromodulatory system (see Figure 6).

Discussion

In the present paper, we used a cognitive robot, CARL-1 to test the hypothesis that neuromodulatory activity can shape learning, drive attention, and select actions. CARL-1 learned to approach stimuli that were predictive of positive value and move away from stimuli that were predictive of negative value (see Figures 3 and 4). An intact neuromodulatory system was necessary for correct behavioral responses (see Figure 4) and for appropriate neuromodulatory responses to stimuli (see Figures 5 and 6). These experiments suggest a mechanism of how neuromodulatory systems influence attention and decision-making.

The neural control of the cognitive robot presented here may be a design strategy for controlling autonomous systems based on principles neuromodulation found in the mammalian brain. Such a controller would flag an important environmental stimulus, cause the autonomous system to focus its attention on the appropriate signal, ignore irrelevant distracters, and quickly respond to pressing events.

Dopamine and "Wanting" Behavior

Dopamine appears to be important for "wanting", that is, the motivation process in acquiring an object [13]. Dopamine, which is found throughout the central nervous system, is produced in the ventral tegmental area. A recent proposal ties the prediction error to wanting by suggesting that incentive salience is the expected future reward that maps actions to rewards [14]. Alternatively, it has been proposed that dopamine is involved with the discovery of new actions and it influences action-outcome contingencies [11]. From the evidence, it appears that dopamine is an important signal for the acquisition of value-laden objects.

In the present paper, we showed that dopaminergic neuromodulation arising from a simulated Ventral Tegmental Area was necessary for value-laden "wanting" responses. When CARL-1's dopaminergic system was intact, it approached stimuli that were predictive of positive value, and ignored neutral stimuli (see Figure 3A and control group in Figure 4). When CARL-1's VTA was lesioned, the number of Find responses, which signify "wanting", significantly decreased (see VTA group in Figure 4). Instead of approaching these positive-value stimuli, CARL-1 treated green objects as neutral stimuli.

Serotonin and "Risky" Behavior

Serotonin originates in the Raphe nucleus and its effect on the nervous system appears to be related to the control of stress. The structures, which receive serotonin from the Raphe, modulate behavioral response to threats, and risks [6]. For example, serotonin plays an important role in social anxiety and social threats in primates [15].

In our experiments with CARL-1, we showed that serotonergic neuromodulation arising from a simulated Raphe nucleus was needed to respond appropriately to threatening stimuli. When CARL-1's serotonergic system was intact, it moved away from threatening stimuli, and ignored neutral stimuli (see Figure 3A and control group in Figure 4). But when CARL-1's Raphe was lesioned, its behavior became "risky" in that it approached Red stimuli as if they were of neutral value (see Raphe group in Figure 4).

Acetylcholine and Attentional Effort

Acetylcholine originates from the basal forebrain and projects to the cortex, amygdala, and hippocampus. The basal forebrain appears to enhance input processing and the allocation of attentional resources for important stimuli under challenging conditions [16]. Removal of cholinergic projections to the parietal and frontal cortex impairs the ability to increase attentional effort [17].

In our experiments, the simulated basal forebrain enhanced CARL-1's ability to attend to salient objects. Removal of the basal forebrain alone through simulated lesions did not have a significant effect on CARL-1's behavior or the signal to noise response in the visuomotor area (see BF in Figures 4 and 5). However, removal of the basal forebrain and another neuromodulatory area, such as VTA or Raphe significantly reduced the appropriate behavioral responses and the signal to noise ratio well below the levels where only Raphe or VTA were

lesioned (see BF+Raphe and BF+VTA in Figures 4 and 5). This suggests a compensatory mechanism for the basal forebrain and other regions, and it is in agreement with the notion that ACh increases the allocation of attentional resources.

Role of Phasic Neuromodulation

Phasic bursts of neuromodulatory activity were necessary to shape CARL-1's behavior during training, and to drive appropriate behavioral responses during testing. The phasic response of the simulated neuromodulators caused CARL-1 to attend to appropriate stimuli, ignore distracters, and take decisive actions (see Figure 3). Neuromodulator activity was strongly correlated with stimuli that were value-laden (see Figure 6). When phasic neuromodulation was impaired, the signal to noise ratio of the system decreased (see Figure 5), and CARL-1 made poor decisions (see Figure 4). This link between phasic neuromodulation and accurate action selection is in agreement with empirical data from animal models [11]. It appears that phasic neuromodulation is important for shifting attention when environmental demands require such vigilance [5].

The Neurorobot Approach

Neurorobotics and cognitive robotics are emerging fields in computer science, neuroscience, and engineering [18]. Neurorobots not only provide a tool for studying brain function by embedding neural simulations on a robotic platform, but they also provide the groundwork to develop intelligent machines based on neurobiological principles. The present work showed how a model of neuromodulation could be used to shape a robot's behavior, such that it focused its attention on important events, and made effective decisions.

Although it could be argued that virtual environments could be used for the present work, the real environment is required for several reasons [19, 20]. First, simulating an environment can introduce unwanted and unintentional biases into the model. For example, a computer-generated object presented to a vision model has its shape and segmentation defined by the modeler and directly presented to the model. In a simulation, the color and shading of an object is typically uniform and noise free. However a device that views objects on the floor of a room has to segment the shape and figure from the ground based on its own active vision and deal with camera sensor noise, occlusions, viewing angles, and varying light conditions. Second, because real environments are rich, multimodal, and noisy; an artificial design of such an environment is

computationally intensive and difficult to simulate. However, all these interesting features of the real world come for free when the robot is allowed to freely move and actively sense in an environment. Finally, there are theoretical implications that can be characterized by the slogan "understanding through building" [20]. To truly understand the system being studied, it is essential to build the actual physical system. Real physical systems tend to yield the most insights because they include the most details in their design and are grounded in the physics of the real world.

In the work presented here, CARL-1 overcame much of the environmental and sensory noise through phasic neuromodulation. Because the environment was varied and interesting, it developed interesting, experience-dependent responses that would be difficult to replicate in a simulated environment. For example, the Find and Flee responses that emerged through CARL-1's learning were fairly complex. CARL-1 would focus its attention on a salient object, by aiming its camera at the object of interest, as it either approached or moved away from the stimulus (see Figure 3). In particular, the Flee response gave the impression of an animal warily eyeing a threatening object as it slowly backed way.

Neuromodulation as a Robot Controller

While conventional robots and autonomous systems require some level of supervision and tuning of parameters to fit a particular domain, biological organisms have the ability to respond quickly and appropriately in an ever-changing world. We have shown how a model of the neuromodulatory system and surrounding regions, can cause a robot to: (1) sharpen its sensory systems, (2) attend to behaviorally relevant objects and ignore distractions, (3) learn to predict the value and outcome of its decisions, and (4) respond decisively and appropriately to environmental events.

Other groups have taken a similar approach in modeling neuromodulation and action selection. The phasic response of dopamine has been modeled to examine reward anticipation behavior in a robot [21]. Models of the basal ganglia have been tested on robots to demonstrate action selection and switching behavior [22]. In a robotic system that has correlates with features of the noradrenergic system, "cyber rodents" explored new behaviors when their battery packs are full, but took more exploitative behavior when their battery packs were nearly empty [23]. A study of selection and learning in a simulated robot showed how modulating attentional effort

could induce learning and memory [24]. These studies, which use neural architectures to guide behavior and test models of cognition, are in a similar vein to the present work.

Our work with CARL-1 differs from the above studies in that it describes a specific neural mechanism for neuromodulation and shows how this mechanism can lead to decisive behavior under noisy conditions. That is, a neural network can quickly change from arbitrary responses to a winner take all response by amplifying connections carrying sensory information. This mechanism has been shown in the present experiments with CARL-1, in theoretical modeling [1], and in empirical data [3, 4].

Cognitive robots and neurorobots provide a synergy between empirical and simulated data, which can lead to improvements in the model and predictions in the modeled organism. An advantage of the neurorobot approach taken here is that it provides a model that can be directly tested against animal models; both in its behavioral response and in its neuronal response. Another advantage of this approach is that cognitively and neurally inspired robots can provide a framework for a new class of intelligent machines. We have presented a design strategy, based on principles of the neuromodulatory system, which controlled the behavior of autonomous robot systems. This research showed that such a system could respond appropriately to environmental changes. Although the field is at a nascent stage, researchers in cognitive robotics are following working models: biological nervous systems and human cognition. If scientists are able to find the underlying principles of these working models and engineers can construct machines based on these principles, it will result in a major advancement in the field of robotics.

Acknowledgments

This work was supported by the National Science Foundation (EMT/BSSE Award No.: 0829752) and the Office of Naval Research (Award No.: N000140910036).

[1] J. L. Krichmar, "The Neuromodulatory System – A Framework for Survival and Adaptive Behavior in a Challenging World," *Adaptive Behavior,* vol. 16, pp. 385-399, 2008.

[2] L. A. Briand, H. Gritton, W. M. Howe, D. A. Young, and M. Sarter, "Modulators in concert for cognition: Modulator interactions in the prefrontal cortex," *Prog Neurobiol,* vol. 83, pp. 69-91, Oct 2007.

[3] M. E. Hasselmo and J. McGaughy, "High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation," *Prog Brain Res,* vol. 145, pp. 207-31, 2004.

[4] M. Kobayashi, K. Imamura, T. Sugai, N. Onoda, M. Yamamoto, S. Komai, and Y. Watanabe, "Selective suppression of horizontal propagation in rat visual cortex by norepinephrine," *Eur J Neurosci,* vol. 12, pp. 264-72, Jan 2000.

[5] G. Aston-Jones and J. D. Cohen, "Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance," *J Comp Neurol,* vol. 493, pp. 99-110, Dec 5 2005.

[6] M. J. Millan, "The neurobiology and control of anxious states," *Prog Neurobiol,* vol. 70, pp. 83-244, Jun 2003.

[7] M. G. Baxter and A. A. Chiba, "Cognitive functions of the basal forebrain," *Curr Opin Neurobiol,* vol. 9, pp. 178-83, Apr 1999.

[8] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science,* vol. 275, pp. 1593-9, Mar 14 1997.

[9] A. J. Yu and P. Dayan, "Uncertainty, neuromodulation, and attention," *Neuron,* vol. 46, pp. 681-92, May 19 2005.

[10] Q. Gu, "Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity," *Neuroscience,* vol. 111, pp. 815-35, 2002.

[11] P. Redgrave and K. Gurney, "The short-latency dopamine signal: a role in discovering novel actions?," *Nat Rev Neurosci,* vol. 7, pp. 967-75, Dec 2006.

[12] E. L. Bienenstock, L. N. Cooper, and P. W. Munro, "Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex," *J Neurosci,* vol. 2, pp. 32-48, Jan 1982.

[13] K. C. Berridge, "Motivation concepts in behavioral neuroscience," *Physiol Behav,* vol. 81, pp. 179-209, Apr 2004.

[14] S. M. McClure, N. D. Daw, and P. R. Montague, "A computational substrate for incentive salience," *Trends Neurosci,* vol. 26, pp. 423-8, Aug 2003.

[15] K. K. Watson, J. H. Ghodasra, and M. L. Platt, "Serotonin transporter genotype modulates social reward and punishment in rhesus macaques," *PLoS ONE,* vol. 4, p. e4156, 2009.

[16] M. Sarter, W. J. Gehring, and R. Kozak, "More attention must be paid: the neurobiology of attentional effort," *Brain Res Rev,* vol. 51, pp. 145-60, Aug 2006.

[17] D. J. Bucci, P. C. Holland, and M. Gallagher, "Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli," *J Neurosci,* vol. 18, pp. 8038-46, Oct 1 1998.

[18] A. K. Seth, O. Sporns, and J. L. Krichmar, "Neurorobotic models in neuroscience and neuroinformatics," *Neuroinformatics,* vol. 3, pp. 167-70, 2005.

[19] J. L. Krichmar and G. M. Edelman, "Brain-Based Devices for the Study of Nervous Systems and the Development of Intelligent Machines," *Artificial Life,* vol. Vol. 11, pp. 63-78, 2005.

[20] R. Pfeifer and J. Bongard, *How the body shapes the way we think: A new view of intelligence*. Cambridge, MA: MIT Press, 2007.

[21] W. H. Alexander and O. Sporns, "An Embodied Model of Learning, Plasticity, and Reward," *Adaptive Behavior,* vol. 10, pp. 143-159, 2002.

[22] T. J. Prescott, F. M. Montes Gonzalez, K. Gurney, M. D. Humphries, and P. Redgrave, "A robot model of the basal ganglia: behavior and intrinsic processing," *Neural Netw,* vol. 19, pp. 31-61, Jan 2006.

[23] K. Doya and E. Uchibe, "The Cyber Rodent project: Exploration of adaptive mechanisms for self-preservation and self-reproduction," *Adaptive Behavior,* vol. 13, pp. 149-160, Jun 2005.

[24] J. Garforth, S. L. McHale, and A. Meehan, "Executive attention, task selection and attention-based learning in a neurally controlled simulated robot," *Neurocomputing,* vol. 69, pp. 1923-1945, 2006.

Figure Captions

Figure 1. CARL-1 and Experimental Setup. **A.** CARL-1 is a wheeled mobile robot with a RF-CCD camera for vision, IR sensors for obstacle avoidance, and wireless RS-232 for communication with a computer workstation. **B.** CARL-1's environment consisted of an enclosure with eight light panels. Each panel could communicate its position to CARL-1 if it was on top of the panel. The four corner panels could be set to one of four colors.

Figure 2. Schematic of Neural Architecture. Each ellipse denotes a neural area that contains simulated neurons. The arrows between neural areas denote synaptic projections containing many connections between neurons. Within area connections, inhibitory connections and connections from Behavior Drivers to Action areas are omitted for clarity (see text for details). The neural simulation contains 6,700 neurons and roughly 1.3 million synaptic connections.

Figure 3. CARL-1 behavior and neural activity. **A.** Find Behavior. **Top row**. Snapshots of CARL-1 during Find behavior. **Bottom row. Left.** Selected neural areas just prior to Find behavior. **Right.** Selected neural areas during Find behavior. Each pixel denotes a neuron, where the activity is color-coded from quiescent (dark blue) to maximally active (bright red). **B.** Flee Behavior. **Top row**. Snapshots of CARL-1 during Flee behavior. **Bottom row. Left.** Selected neural areas just prior to Flee behavior. **Right.** Selected neural areas during Flee behavior.

Figure 4. Behavioral responses for the 10 subjects with an intact simulated nervous system (Control), lesion of the simulated Raphe nucleus (Raphe), lesion of the simulated Ventral Tegmental Area (VTA), lesion of the Basal Forebrain (BF), and lesions of multiple areas (BF+Raphe and BF+VTA). On each box in the plot, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually with plus signs.

Figure 5. Signal to noise ratio (SNR) of target color activity to all color activities during Find and Flee behaviors. Plots show the mean SNR (see equation 6) for the 10 subjects with an

intact simulated nervous system (Control), lesion of the simulated Raphe nucleus (Raphe), lesion of the simulated Ventral Tegmental Area (VTA), lesion of the Basal Forebrain (BF), and lesions of multiple areas (BF+Raphe and BF+VTA). Error bars denote the standard deviation.

Figure 6. Scatter plots of visuomotor activity versus neuromodulatory activity for 10 subjects over all Control trials. The Pearson's correlation coefficient (r) is given at the top of each plot. **A.** Color neural activity versus Raphe activity. **B.** Color activity versus VTA activity.

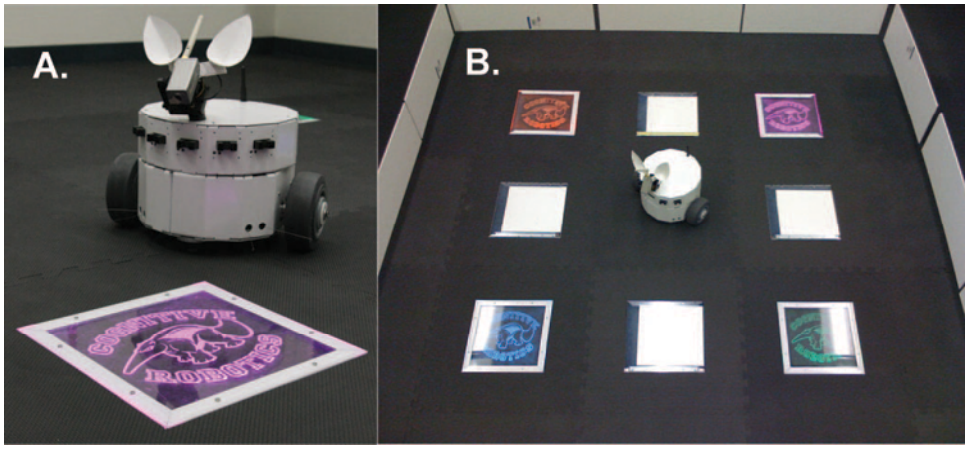

Figure 1.

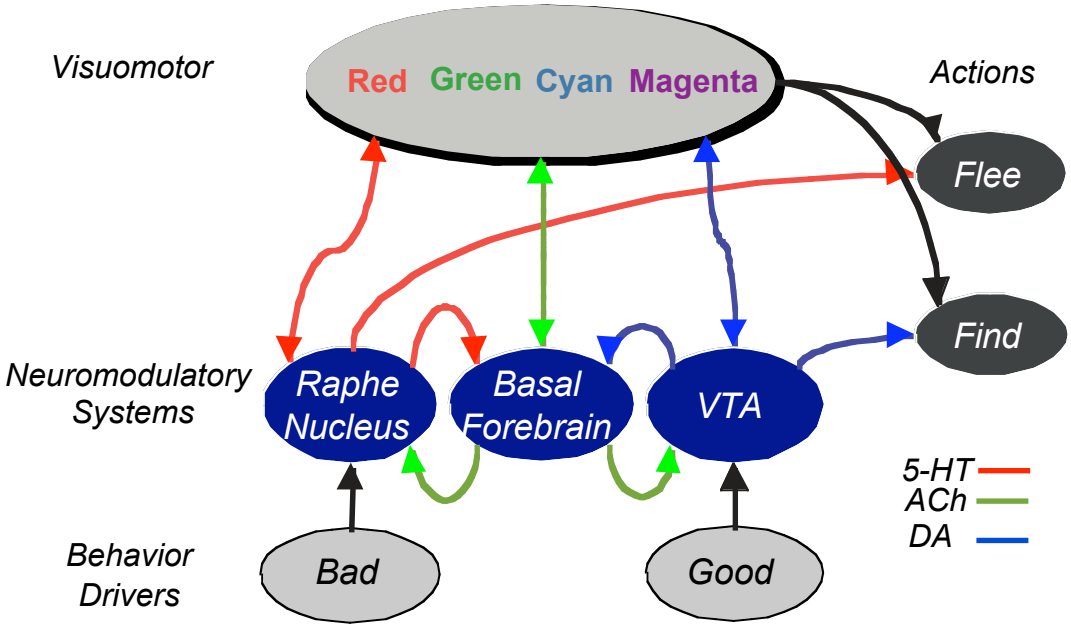
Visuomotor
Red Green Cyan Magenta
Actions
Flee
Find
Neuromodulatory Systems
Raphe Nucleus
Basal Forebrain
VTA
5-HT
ACh
DA
Behavior Drivers
Bad
Good

Figure 2.

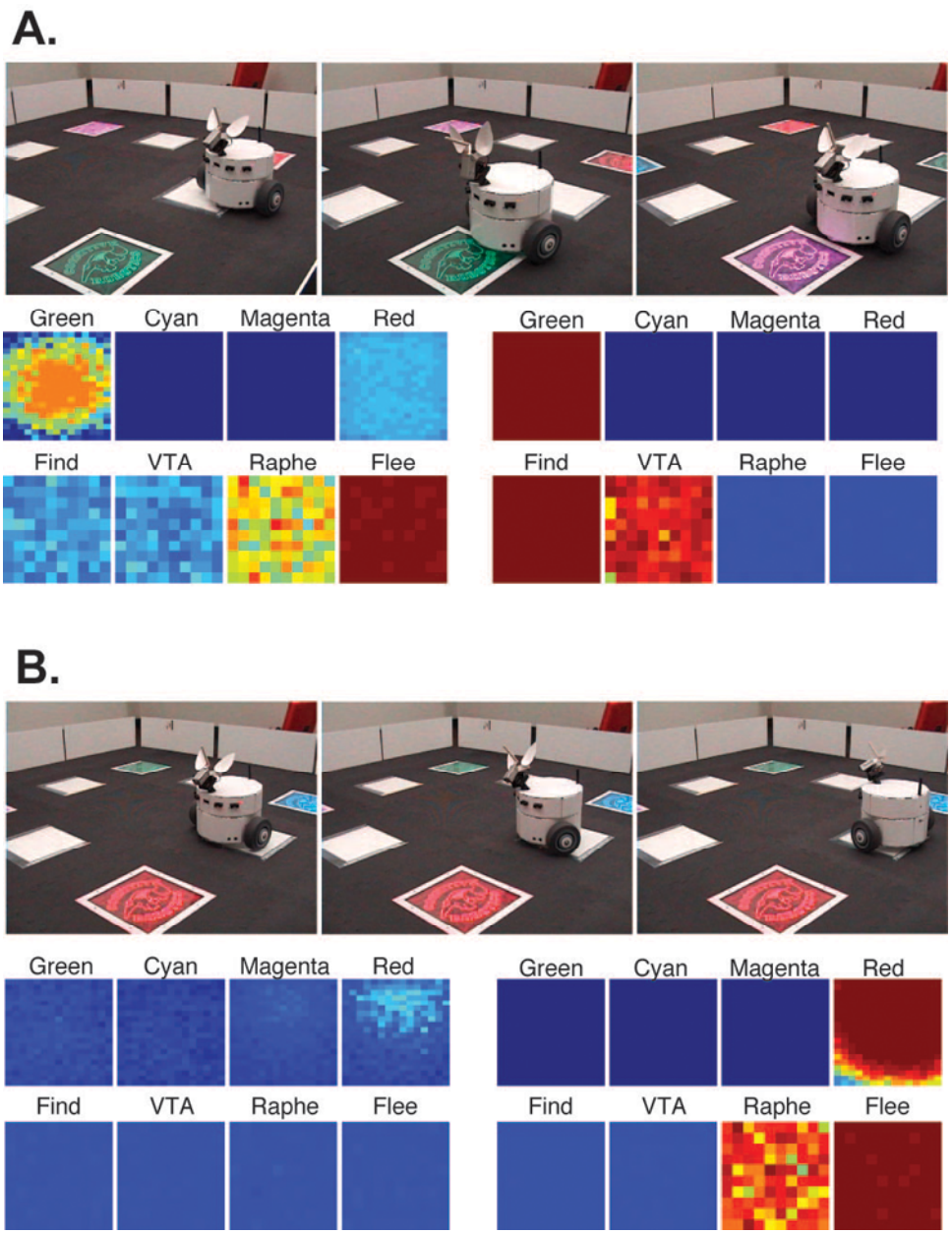
A.
Green
Cyan
Magenta
Red
Green
Cyan
Magenta
Red
Find
VTA
Raphe
Flee
Find
VTA
Raphe
Flee
B.
Green
Cyan
Magenta
Red
Green
Cyan
Magenta
Red
Find
VTA
Raphe
Flee
Find
VTA
Raphe
Flee

Figure 3.

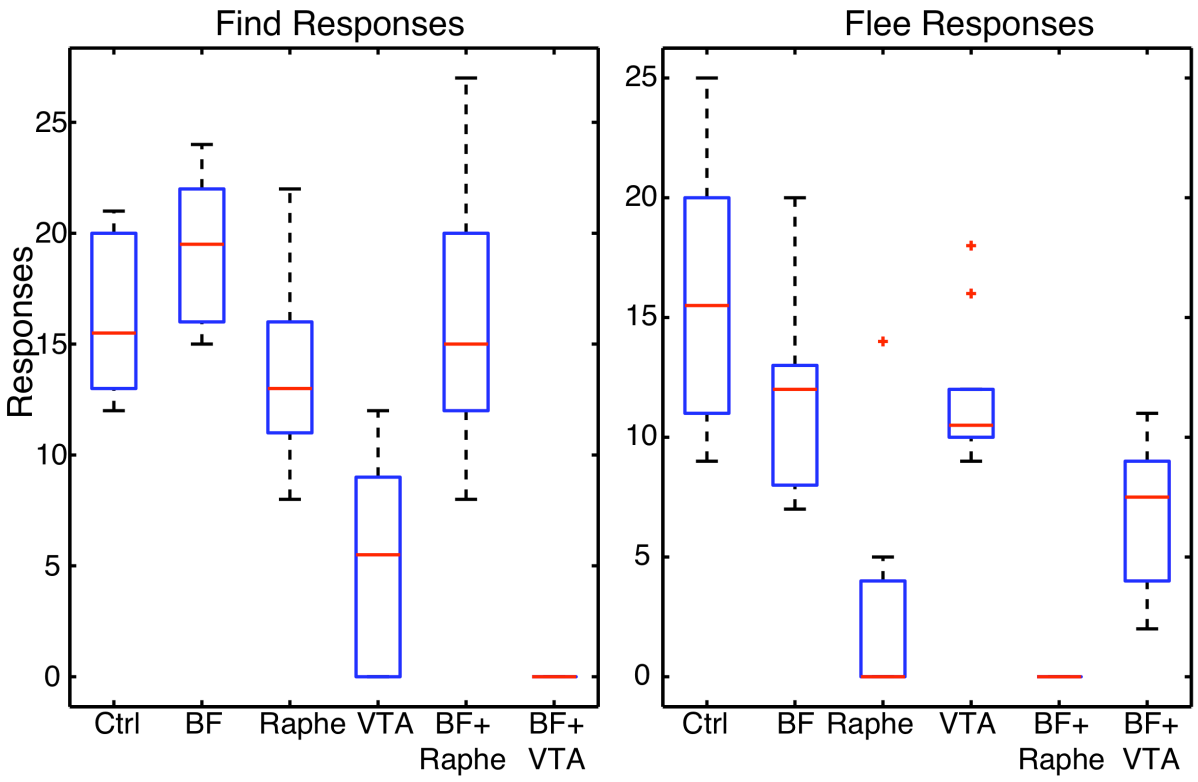
Find Responses
Flee Responses
Responses
25
20
15
10
5
0
Ctrl
BF
Raphe
VTA
BF+ Raphe
BF+ VTA

Figure 4.

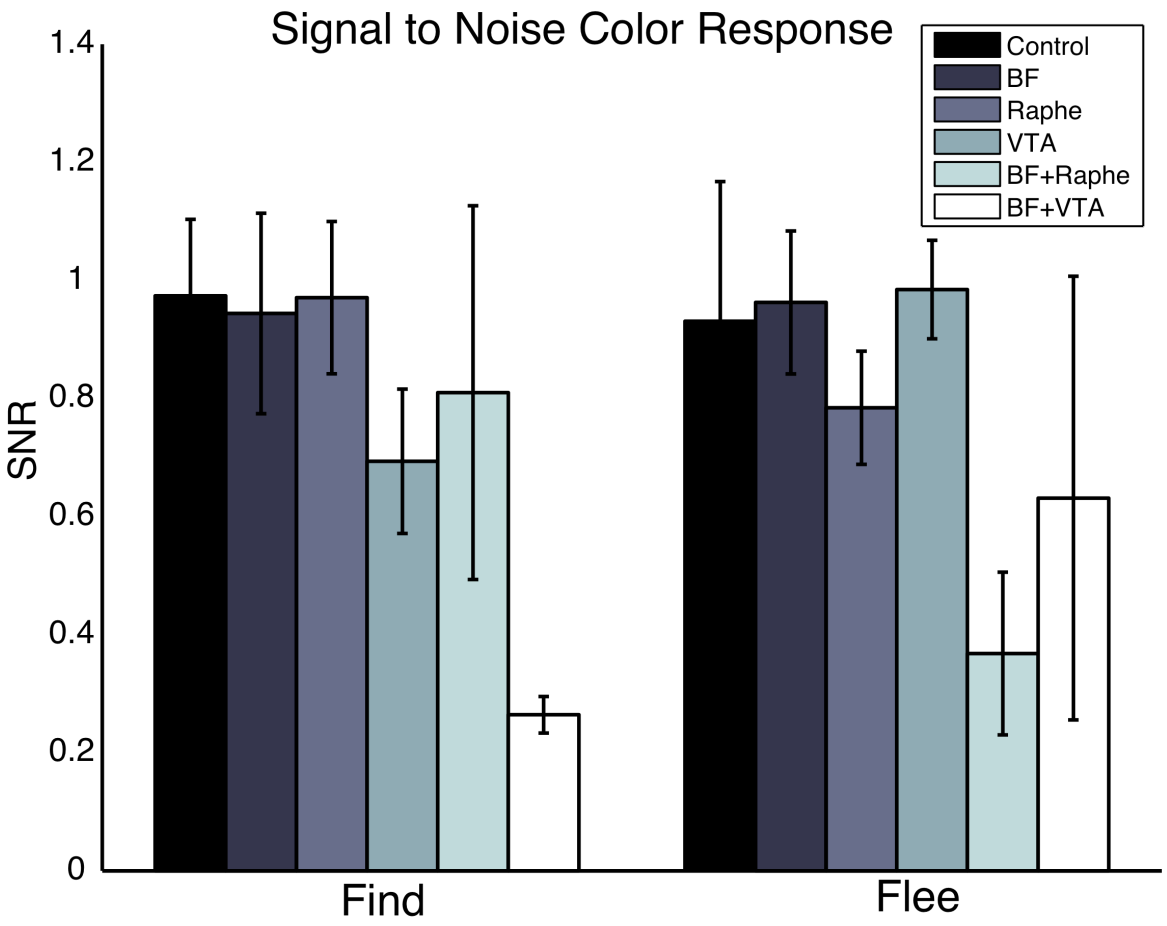
Signal to Noise Color Response
Control
BF
Raphe
VTA
BF+Raphe
BF+VTA
SNR
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Find
Flee

Figure 5.

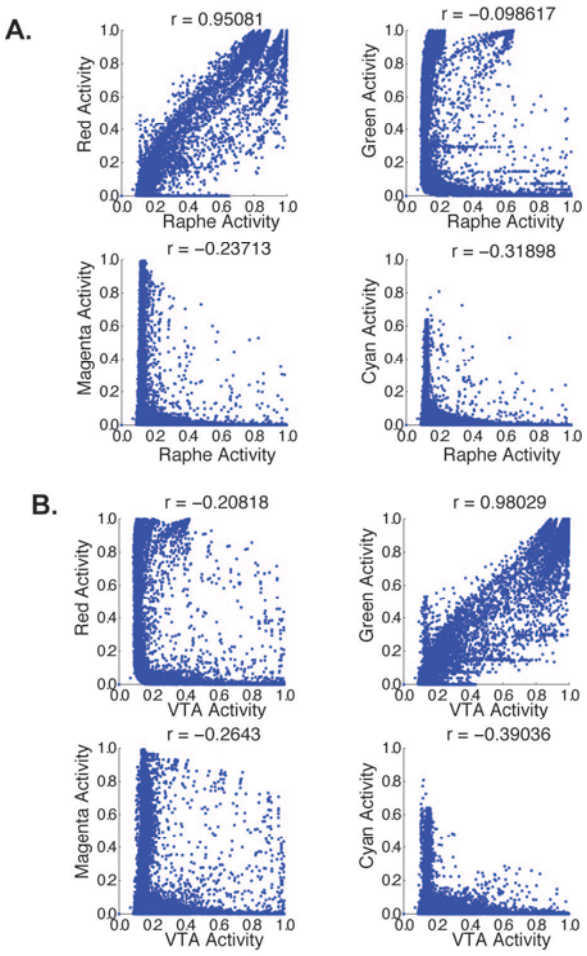
A.
r = 0.95081
Red Activity
Raphe Activity
r = −0.098617
Green Activity
Raphe Activity
r = −0.23713
Magenta Activity
Raphe Activity
r = −0.31898
Cyan Activity
Raphe Activity
B.
r = −0.20818
Red Activity
VTA Activity
r = 0.98029
Green Activity
VTA Activity
r = −0.2643
Magenta Activity
VTA Activity
r = −0.39036
Cyan Activity
VTA Activity

Figure 6.