A Network Model of Catecholamine Effects: Gain, Signal-to-Noise Ratio, and Behavior

David Servan-Schreiber, Harry Printz, Jonathan D. Cohen

At the level of individual neurons, catecholamine release increases the responsivity of cells to excitatory and inhibitory inputs. A model of catecholamine effects in a network of neural-like elements is presented, which shows that (i) changes in the responsivity of individual elements do not affect their ability to detect a signal and ignore noise but (ii) the same changes in cell responsivity in a network of such elements do improve the signal detection performance of the network as a whole. The second result is used in a computer simulation based on principles of parallel distributed processing to account for the effect of central nervous system stimulants on the signal detection performance of human subjects.

The catecholamines norepinephrine and dopamine are neuroactive substances that are presumed to modulate information processing in the brain rather than to convey discrete sensory or motor signals. Release of norepinephrine and dopamine occurs over wide areas of the central nervous system (CNS), and the postsynaptic effects of the release of these catecholamines are long-lasting (1). One important effect consists of an enhancement of the response of target cells to other afferent inputs, inhibitory as well as excitatory [(2); reviewed in (3)].

Increases or decreases in catecholaminergic tone have many behavioral consequences, including effects on motivated behaviors, attention, learning, memory, and motor behavior. At the information processing level, catecholamines appear to affect the ability to detect a signal when it is embedded in noise [reviewed in (4)]. However, there is no adequate account of how these changes at the system level relate to the effect of catecholamines on individual cells. Several investigators (5–8) have suggested that catecholamine-mediated increases in a cell’s responsibility can be interpreted as a change in the cell’s signal-to-noise ratio (SNR). By analogy, they proposed that this change at the unit level may account for changes in signal detection performance at the behavioral level. We explore here the relation between these two levels, using mathematical and computational models of individual neurons and networks of neurons.

We assume that the response of a typical neuron can be described by a strictly increasing function $f_G(x)$ from real-valued inputs to the interval $(0, 1)$. This function relates the strength of a neuron’s net afferent input $x$ to its probability of firing or activation. We do not require that $f_G$ be differentiable or even continuous. We call $f_G$ the activation function.

For instance, the family of logistics, given by

$$f_G(x) = \frac{1}{1 + e^{-(Gx + B)}}$$

has been proposed as a model of neural response functions (9). The logistics are all strictly increasing, for all values of $G > 0$ and all values of $B$.

The potentiating effect of catecholamines on responsibility can be modeled as a change in the shape of the activation function. In the case of the logistic, this is achieved by increasing the value of the gain parameter $G$, as illustrated in Fig. 1B. As $G$ increases, the value $f_G(x)$ comes arbitrarily close to 1 if $x > 0$ and arbitrarily close to 0 if $x < 0$ (10).

Consider the signal detection performance of a network in which the response of a single unit is compared with a threshold to determine the presence or absence of a signal. We assume that in the presence of the signal this unit receives a positive (excitatory) net afferent input $x_A$ and in the absence of the signal it receives a null or negative (inhibitory) input $x_A$. When zero-mean noise is added to this quantity, in the presence as well as in the absence of the signal, the unit’s net input in each case is distributed around $x_I$ or $x_A$, respectively (Fig. 1A). Therefore, its response is distributed around $f_G(x_I)$ or $f_G(x_A)$, respectively.

In other words, the input in the case where the signal is present is a random variable $X_S$, with probability density function (PDF) $p_X(x)$, and in the absence of the signal it is the random variable $X_A$, with PDF $p_X(x)$. These inputs then determine the random variables $Y_{GS} = f_G(X_S)$ and $Y_{GA} = f_G(X_A)$, with PDFs $p_{Y_{GS}}$ and $p_{Y_{GA}}$, which represent the response in the presence or absence of the signal for a given value of $G$ (Fig. 1C).

If the input PDFs $p_X(x)$ and $p_X(x)$ overlap, the output PDFs $p_{Y_{GS}}$ and $p_{Y_{GA}}$ will also overlap. Thus, for any given threshold $\theta$ on the $y$ axis

![Fig. 1. (A) Example of the probability density function (PDF) of the net input in the cases of signal absent ($p_{X_A}$) and signal present ($p_{X_A}$). (B) The logistic function at two values of gain $G$. This function has been proposed as a model of neural responsivity. The unit's activation at zero net input corresponds to a neuron's baseline firing rate. Positive net inputs correspond to excitatory stimuli, negative net inputs correspond to inhibitory stimuli. For the graphs drawn here, we set the bias $B = -1$. The negative bias renders the function asymmetric around a net input of 0. This asymmetry is often found in the response function of actual neurons (22). Increasing $G$ drives up a unit's response to a positive input and drives down its response to a negative input. (C) Examples of the PDFs of the activation value of a unit in the presence ($p_{Y_{GA}}$) and in the absence ($p_{Y_{GA}}$) of the signal. These are the PDFs of the transformed RVs, $Y_{GS} = f_G(X_S)$, and $Y_{GA} = f_G(X_A)$. Each PDF is drawn for two different values of $G$, in the case where $f_G$ is the logistic.](http://www.sciencemag.org)
used to categorize the output as "signal present" or "signal absent," there will be some misses and some false alarms. The best the system can do is to select a threshold that optimizes performance. More precisely, the expected payoff or performance of the unit is given by

\[ E(\theta) = \lambda + \alpha \cdot \Pr(Y_{GS} \geq \theta) - \omega \cdot \Pr(Y_{GA} \geq \theta) \] (2)

where \(\lambda, \alpha,\) and \(\omega\) are constants that together reflect the prior probability of the signal and the payoffs associated with correct detections (also called hits), correct ignores, false alarms, and misses. Note that \(\Pr(Y_{GS} \geq \theta)\) and \(\Pr(Y_{GA} \geq \theta)\) are the probabilities of a hit and a false alarm, respectively. By solving the equation \(\frac{dE(\theta)}{d\theta} = 0\), we can determine a threshold, \(\theta^*\), that maximizes \(E\). We call \(\theta^*\) the optimal threshold.

From examination of Fig. 1, it might be supposed that, by changing the activation function, one can improve signal detection performance. But this is not so. For any activation function \(f\) that satisfies our assumptions and any fixed input PDFs \(p_X\) and \(p_Y\), the unit’s performance at optimal threshold is the same. This is our constant optimal performance theorem (COPT); it is stated and proved in (11). In particular, for the logistic, increasing the gain \(G\) does not induce better performance. It may change the value of the threshold that yields optimal performance, but it does not change the actual performance at optimum. This is because a strictly increasing activation function gives a one-to-one mapping from input to output values. This makes it possible to express Eq. 2 exclusively in terms of the input PDFs \(p_X\) and \(p_Y\), and \(\alpha, \omega\), and \(\lambda\). Because it is the overlap between \(p_X\) and \(p_Y\) that limits performance, and because this overlap does not vary with the gain, the performance at optimal threshold is constant.

We now examine the effect of changing the gain on the SNR of the output of a single unit. In electrical engineering, the SNR measures the difficulty of extracting a continuous-time signal \(s(t)\) from a noisy background \(n(t)\). The SNR compares the average power input to the receiver in the presence of the signal, \(S = \langle |s(t)|^2\rangle\), with the average power input in the absence of the signal, \(N = \langle |n(t)|^2\rangle\) (12). If \(s(t)\) is a small perturbation added to \(n(t)\), then \(S \approx N\), and the signal will be difficult to detect. On the other hand, if the signal amplitude is high and the noise amplitude is low, then \(S >> N\). Thus, the ratio \(S/N\) measures how difficult it is to distinguish signal from noise.

In the case of a single unit, if the unit’s input is \(x\), its output is \(y = f(x)\). Because this quantity represents the firing rate of the neuron for a given stimulus presentation, if each neural spike contains the same amount of energy, the power the neuron delivers will be proportional to \(y\). Thus, over many stimulus presentations, the average power delivered in the presence of signal is proportional to \(\mu(Y_{GS})\), the mean of \(Y_{GS}\), and in the absence of signal is proportional to \(\mu(Y_{GA})\). Hence, the ratio of the average power delivered in the presence of the signal to the average power delivered in the absence of the signal, that is, the SNR, is \(\frac{\mu(Y_{GS})}{\mu(Y_{GA})}\).

In general, raising \(G\) will drive up \(\mu(Y_{GS})\) and drive down \(\mu(Y_{GA})\), increasing the SNR of a single unit. Yet by the COPT, the performance of the unit at optimal threshold remains the same, because the effect of an increase of \(G\) on \(\rho_Y\) and \(\rho_V\) is not captured by the mean alone. Increases in \(G\) will in general alter the shapes of these PDFs, possibly driving apart the main concentrations of probability mass but simultaneously extending their tails (Fig. 1C). The erroneous intuition that separating the means will improve performance arises from the assumption that the effect of an increase in \(G\) is to translate the output PDFs rigidly away from one another. For this reason, it is misleading to explain the performance effects of catecholamines solely in terms of the SNR.

Although increasing \(G\) does not affect the signal detection performance of a single element, it does improve the performance of a chain of such elements. By a chain, we mean an arrangement in which the output of the first unit provides the input to another unit. Let us call this second element the response unit. We monitor the output of this second unit to determine the presence or absence of a signal (Fig. 2).

As in the previous discussion, noise is added to the net input of each unit in the chain in the presence as well as in the absence of a signal (13). We represent noise as a random variable (RV) \(V\), with PDF \(\rho_V\) that we assume to be independent of gain. Let the RVs \(X_1, X_2, Y_{GS}, Y_{GA}\), and their PDFs all be defined as in the single-unit case. Now, because noise is added to the net input of the response unit as well, the input of the response unit is the RV \(Z_{GS} = Y_{GS} + V\) or \(Z_{GA} = Y_{GA} + V\), again depending on whether the signal is present or absent. We write \(\rho_{Z_{GS}}\) and \(\rho_{Z_{GA}}\) for the PDFs of these RVs. Then \(\rho_{Z_{GS}}\) is the convolution of \(\rho_{Y_{GS}}\) and \(\rho_V\), and \(\rho_{Z_{GA}}\) is the convolution of \(\rho_{Y_{GA}}\) and \(\rho_V\). Convoluting the output PDFs of the input unit with the noise distribution increases the overlap between the resulting distributions (\(\rho_{Z_{GS}}\) and \(\rho_{Z_{GA}}\)) and therefore decreases the discriminability of the input to the response unit.

How are these distributions affected by an increase in \(G\) on the input unit? By the COPT, we already know that the discriminability of \(Y_{GS}\) and \(Y_{GA}\) is unchanged. Furthermore, we have assumed that the noise distribution is independent of \(G\). It would therefore seem that a change in \(G\) should not affect the discriminability of \(Z_{GS}\) and \(Z_{GA}\). However, under very general conditions, the overlap between \(\rho_{Z_{GS}}\) and \(\rho_{Z_{GA}}\) decreases when the \(G\) of the input unit increases, thereby improving performance of the two-layered system. We call this the chain effect.

The chain effect arises because the noise added to the net input of the response unit is not affected by variations in \(G\). Increasing \(G\) separates the means of the output PDFs of the input unit, \(\mu(Y_{GS})\) and \(\mu(Y_{GA})\), even though this does not affect the performance of this unit. Suppose all the probability mass were concentrated at these means. Then \(\rho_{Z_{GS}}\) would be a copy of \(\rho_V\) centered at \(\mu(Y_{GS})\), and \(\rho_{Z_{GA}}\) would be a copy of \(\rho_V\) centered at \(\mu(Y_{GA})\). Thus, in this case, increasing \(G\) would rigidly translate \(\rho_{Z_{GS}}\) and \(\rho_{Z_{GA}}\) apart, thereby reducing their overlap and improving performance. A similar effect arises in more general circumstances, when \(Y_{GS}\) and \(Y_{GA}\) are not concentrated at their means. The chain performance theorem, stated and proved in (11), gives sufficient conditions for the appearance of this effect.

The above analysis has shown that increasing the \(G\) of the activation function of individual units in a very simple network can improve signal detection performance. We now present computer simulation results showing that this phenomenon can account for catecholamine-induced performance improvements in a common behavioral test of signal detection.

The continuous performance test (CPT) (14) has been used extensively to study attention and vigilance in behavioral and clinical research. In this task, individual letters are displayed tachystoscopically in a sequence on a computer monitor. In one common version of the task, a target event is to be reported when two consecutive letters are identical. Performance on this task has been shown to be sensitive to drugs or
Fig. 3. Simulation of the continuous performance task. (A) Diagram of recurrent three-layer network (12 input units, 30 intermediate units, 10 output units, and 1 response unit). Each unit projects to all units in the subsequent layer. In addition, each output unit also projects to each unit in the intermediate layer. Letters are presented to the network as patterns of activation over the input units, which act as feature detectors. During training, the network learns to activate the output unit corresponding to the letter being presented on the input. In addition, the recurrent connections provide the network with the pattern of activation evoked on the output layer by the presentation of the previous letter. The network learns to activate the response unit when two consecutive letters are identical and to turn it off in all other cases. The activation of a unit in the intermediate or output layer depends on the activations of all units in the layers providing input to it. Each input is weighted by the corresponding connection strength, which can be positive or negative. The sum of the weighted inputs is then passed through the logistic function to determine the unit's activation. The gain parameter G is the same for all intermediate and output units. In the simulation of the placebo condition, G = 1.0; in the simulation of the drug condition, G = 1.1. Bias B is -1 in both conditions. (B) Performance of human subjects and of the simulation on the CPT. Filled markers indicate the performance of human subjects with placebo and methylphenidate (16). Empty markers indicate the results of the simulation.

REFERENCES AND NOTES
2. The effect of catecholamines on target cells is generally characterized as an enhancement of stimuluselicited responses with respect to background firing.
The C7 Cluster: Structure and Infrared Frequencies

J. R. Heath, R. A. Sheeks, A. L. Cooksy, R. J. Saykally

Observation and characterization of the C7 cluster are reported. Carbon clusters are produced by laser vaporization of a graphite target followed by supersonic expansion of the vaporized material within a gas dynamically focused argon jet. Thirty-six sequential rovibrational lines of the ν4 antisymmetric stretch fundamental of C7 are probed by gated detection of diode laser absorption. The observed spectrum is characteristic of a symmetrical linear molecule. Analysis of the spectrum indicates an effective average bond length of 1.2736(4) angstroms and a vibrational frequency of 2138.1951(10) reciprocal centimeters, in excellent agreement with ab initio calculations. This work will facilitate the quantum chemical characterization of this cluster.

Small carbon clusters (less than 12 atoms) have recently attracted the attention of numerous investigators from a wide variety of disciplines. This is largely due to the ubiquitous nature of these species; they have been observed in astrophysical sources (1), in sooting flames (2), in acetylene photolysis (3, 4), and in plasmas produced by laser vaporization of graphite (5–7). This suggests that unsaturated carbon clusters play a critical, if not central, role in the high-temperature chemistry of carbon-rich environments.

Ab initio and semiempirical theory of small carbon clusters has been under constant development for several decades (8, 9). Much of this work has recently been reviewed by Welner and Van Zee (10). Odd-numbered clusters of up to 11 atoms are expected to have linear Σ ground states, with the lowest triplet states existing at much higher energy. Even-numbered clusters of up to 10 atoms are predicted to have two low-energy configurations: an open shell linear Σ state and a monocyclic singlet state. There is much debate regarding the detailed properties of these even clusters (10). Odd-numbered clusters up to C7 are predicted to be more stable than the adjacent even clusters (9).

Despite this high level of theoretical activity, experimental results have been sparse. The development of tandem mass spectroscopic technologies has enabled researchers to study carbon cluster cation photofragmentation (6) and anion photoelectron spectroscopy (7). Results from these experiments have been consistent with theory. Only recently, however, have definitive experiments capable of testing detailed theoretical predictions been possible. Over the past 2 years a number of research groups have characterized the C3 cluster with high-resolution laser techniques and have obtained sufficient information for the construction of an accurate molecular potential surface (3, 12–14). Last year we accomplished a detailed laboratory characterization of the C3 cluster (5) using infrared laser spectroscopy. That experiment was reported simultaneously with the detection of C3 in the carbon star IRC+10216 by Bernath, Hinkle, and Keedy (1). Additional bands of C2 have been detected and analyzed by Moazzen-Ahmadi, McKellar, and Amano (4, 15). Ab initio calculations are in close agreement with those experimental results.

In this paper we describe direct observation and characterization of the C7 cluster, carried out with an infrared laser spectroscopy technique similar to that used in our study of C3. Thirty-six sequential rovibrational lines have been measured and assigned to the ν4 antisymmetric stretch vibrational transition of C7. The observed spectrum is characteristic of a symmetrical linear molecule with a closed electronic shell. Analysis of the spectrum indicates a ground state rotational constant of 0.030929(21) cm⁻¹.