Acetylcholine, Norepinephrine, and Spatial Attention

Angela Yu Peter Dayan

Gatsby Computational Neuroscience Unit, UCL, 17 Queen Sq, London WC1N 3AR, UK

Abstract

Despite strong implication of the neuromodulators acetylcholine (ACh) and norepinephrine (NE) in cognitive tasks, there is little consensus on their computational functions. We propose that they report uncertainties: ACh signals expected uncertainty, coming from known variability within a context; NE signals unexpected uncertainty, as when the environment is dramatically changed. We apply these ideas to a novel attention task, which allows explicit manipulation of expected and unexpected uncertainties. In our model, ACh and NE interact in a precisely specified and part-opponent, part-synergistic manner. Simulation results replicate existent data, and make experimentally tractable predictions regarding neuronal responses and certain psychophysical measures.

Keywords: neuromodulation, cortical inference, uncertainty, attention, top-down

Biosketch of Angela Yu: Angela received the following degrees from MIT in 2000: BS Computer Science, BS Mathematics, BS Brain and Cognitive Sciences. She is currently wrapping up her PhD in computational neuroscience at the Gatsby Computational Neuroscience Unit, University College London.

Biosketch of Peter Dayan: Peter Dayan did his PhD at University of Edinburgh, postdocs at University of California, San Diego and University of Toronto, and taught at MIT in the Department of Brain and Cognitive Sciences, before coming to the Gatsby Computational Neuroscience in University College London.

1 Introduction

Neuromodulators such as norepinephrine (NE), acetylcholine (ACh), and dopamine (DA) are activated quickly and selectively by relevant sensory stimuli, are modulated by internal states, and exert strong effects on a wide variety of cognitive functions. These results have motivated theoretical interpretations in which the neuromodulators play critical and specific roles in neural computations[15]. ACh and NE are of special importance because of their cerebral ubiquity and their evident involvement in controlling the integration of bottom-up and top-down information in the cortex and the course of learning in a variety of tasks. However, although it is known that ACh and NE interact, at least during developmental plasticity[1], neither experiments nor theories have provided a clear picture of the nature or goals of this interaction. For instance, both neuromodulators are associated with aspects of attention[9], but there is no apparent rhyme nor reason to the tasks in which pharmacological manipulations of one or the other or, less commonly, both, have an effect[11, 10, 5].

We have recently developed a theory in which cortical ACh and NE report different aspects of uncertainty, with ACh indicating *expected* uncertainty, coming from known variability within a behavioral context, and NE signalling *unexpected* uncertainty, as when significant aspects of the task are unpredictably changed[16]. This theory is motivated by a wealth of pharmacological, behavioral, and physiological data, and is quantitative enough to make specific predictions about the interaction between the neuromodulators. In this work, we apply these ideas to a certain class of attentional tasks that are known to interact with ACh and NE and can be interpreted in the top-down, bottom-up inference framework

The classical Posner Task involves assessing the costs and benefits associated with invalid or valid cueing to a sensory location, after which detection or discrimination must be performed[8] on a sensory stimulus. Of particular interest for us is that the effects of manipulating ACh and NE on performance in Posner's task have been investigated, and indeed are qualitatively consistent with our theory. The standard versions of Posner's task engages expected uncertainty but not unexpected uncertainty. Paul Bentley (personal communications) has suggested a novel variant of a Posner task which allows systematic manipulation of both expected and unexpected uncertainty.

In section 2, we describe the conventional and extended versions of Posner's task. In section 3, we present a generative model that captures the (stochastic) experimental contingencies, and a biologically feasible ap-

proximation to the statistical inverse of this generative model, in which the interactions of ACh and NE play a critical role. In section 4, we examine the quality of the approximation, compare the model's performance in the basic Posner's task to experimental data, and show the expected effects of separate and conjoint manipulations of the neuromodulators. In section 5, we discuss the implications of the model, and relate it more precisely to our previous models.

2 Behavioral Task

In the standard Posner Task, a cue indicates the likely location of a subsequent target stimulus. Figure 1(a) shows a particular rendition of the task adapted for rats[13]. Typically, there are cue-induced benefits in reaction time and/or accuracy on correctly cued (valid) trials, and costs on incorrectly cued (invalid) trials. The resulting difference in reaction time (RT(invalid)-RT(valid)) is termed Validity Effect (VE), and has been observed to be a function of cue validity (probability of stimulus appearing where cued), which is an easily experimentally manipulable form of expected *certainty* in the top-down information. ACh, which we propose reports something like cue *invalidity*, should have effects on VE opposite to those of cue validity, a phenomenon supported by experimental evidence [7, 2, 12, 6]. However, the task contingencies, including cue validity, are kept constant throughout a session, thus allowing unexpected uncertainty no significant role. In fact, NE manipulations have not been observed to affect VE in the standard version of Posner's task[14].

We consider a richer spatial attention task, whose relationship to Posner's task will become clear shortly, that manipulates both expected and unexpected uncertainty. On each trial, the subject must first locate the target stimulus in a visually cluttered scene, and then perform a discrimination task on the target. The location of the target from trial to trial follows a particular statistical pattern; behavioral measures indicate human subjects can learn such patterns, sometimes in the absence of conscious knowledge[3]. The context acts as an implicit cue about target locations, so there is still predictive validity of contextual cues and expected uncertainty, much as in the standard Posner task. However, if the underlying patterns are allowed to change occasionally on a much slower time scale, then that induces a form of *unexpected* uncertainty – as unexpected deviations from the learned pattern accumulate, the subject should realize the need to relearn a new pattern.

3 The Model

The particular form of location patterning we consider is illustrated in Figure 1(b). It employs the simple stochastic rule that stimuli usually appear in the primary location (valid trials), but can occasionally deviate (invalid trials). When the pattern changes, a new location becomes the primary location, associated with a new probability of deviations. More formally, the model for generating stimulus locations can be described as follows: on each trial t, there is a primary location μ_t, taking on the values 1, 2, ... n, where n is the total number of discrete locations where the target stimuli may appear. The probability of the target location l_t being j given $\mu_t = i$ is $\beta_{ij}(\gamma_t) \equiv P(l_t = j|\mu_t = i; \gamma_t)$, where γ_t is the deviation parameter associated with the current primary location. β_{ij} takes on the value $1 - \gamma_t$ for $j = i$, and $\gamma_t/(n-1)$ otherwise. Another part of the model is the transition probability between primary locations, written formally as $\alpha_{ij} \equiv P(\mu_t = j|\mu_{t-1} = i)$. This takes the value α for $j = i$ and $(1-\alpha)/(n-1)$ otherwise. Typically, $\alpha \approx 1$, reflecting contextual persistence. Finally, γ_t is constant throughout all the trials associated with a particular pattern, and is redrawn from a uniform distribution when the primary location changes, *ie* $\gamma_t|\mu_t \neq \mu_{t-1} \sim U[0, \gamma_M]$, where $\gamma_M < 1$ sets the cap value for probability of deviation. Readers familiar with the statistical modeling literature would recognize this generative model as a case of a hidden Markov model with non-stationary output distribution specified by the state-dependent parameter γ.

"Learning the pattern" in our task requires the subject to maintain an on-line estimate of posterior distribution over the state variable μ_t and the associated parameter γ_t, from only the observations $\mathbf{l}_t \equiv \{l_1, ..., l_t\}$ (we ignore the fixed parameter α and assume the equilibrium condition when the subject has already learned its value). This is a standard example of Bayesian state inference and parameter estimation. We first derive the exact algorithm and then suggest a biologically plausible approximate model, in which ACh and NE control inference and learning. In the exact model, we compute a recursive form for the joint posterior:

$$\begin{aligned} p(\mu_t = i, \gamma_t|\mathbf{l}_{t-1}, l_t = j) &= p(\mu_t = i, \gamma_t, l_t = j|\mathbf{l}_{t-1})/Z_t \quad \text{where } Z_t \text{ is a normalizing term unique for each } t \\ &\propto \sum_k \int d\gamma_{t-1}\ p(\mu_{t-1} = k, \gamma_{t-1}|\mathbf{l}_{t-1}) P(\mu_t = i|\mu_{t-1} = k) \cdot \\ &\qquad p(\gamma_t|\mu_t = i, \mu_{t-1} = k, \gamma_{t-1}) P(l_t = j|\mu_t = i, \gamma_t) \\ &= \beta_{ij}(\gamma_t)(\alpha p(\mu_{t-1} = i, \gamma_{t-1} = \gamma_t|\mathbf{l}_{t-1}) + (1 - P(\mu_{t-1} = i|\mathbf{l}_{t-1}))\frac{1-\alpha}{\gamma_M(n-1)}) \quad (1) \end{aligned}$$

This last expression would seem to provide a convenient iterative algorithm for computing the posterior

distribution, except the exact integration of the posterior over γ_t, a large mixture of beta-distributions whose number of sufficient statistics grow as $O(nt)$, is representationally costly and computationally intractable

In order to use the underlying patterns in the task effectively, the brain must implement a tractable *approximate* inference algorithm with reasonable performance. We propose one such algorithm, which also utilizes ACh and NE signaling with the appropriate semantics. The idea is to approximate the posterior distribution $P(\mu_t|\mathbf{l}_t)$ with a simpler distribution P^* that requires the computation and representation of only a few quantities: the most likely primary location $\mu_t^* = j$, the estimated posterior $\lambda_t^* \equiv P^*(\mu_t^*|\mathbf{l}_t)$, the estimated most likely deviation parameter γ_t^* corresponding to the current μ_t^*, and h_t^*, the estimated number of trials associated with the current pattern. To reconstruct the full *approximate* posterior, we assume $P^*(\mu_t = i \neq j|\mathbf{l}_t) = (1 - \lambda_t^*)/(n-1)$, and the deviation parameter associated with all states $i \neq j$ as $\gamma_t^i = \gamma_0 = \gamma_M/2$, where γ_0 is the expected value of the parameter γ under the uniform prior. We suggest that ACh and NE report γ_t^* and $1 - \lambda_t^*$, respectively. γ_t^* is the expected deviation in a particular context, and therefore appropriate for ACh's role as expected uncertainty. $1 - \lambda_t^*$ is the "doubt" that the current model of location pattern is correct at all. It is as a form of unexpected uncertainty since $1 - \lambda_t^*$ is small if many trials of a particular pattern have been observed, and only increases to a significant magnitude if many more deviations have been observed than expected.

Iterative computation of this *approximate* joint posterior is tractable and efficient, and breaks down into the two major scenarios of a new observation appearing as predicted by the internal model, or not. For the former case, $l_{t+1} = k = j$, it is easy to see μ_{t+1}^* should be set to be equal to μ_t^*, and

$$\lambda_{t+1}^* \approx \frac{(1-\gamma_t^*)\lambda_t^*}{(1-\gamma_t^*)\lambda_t^* + \gamma_0(1-\lambda_t^*)/(n-1)} \tag{2}$$

where the numerator gives $P(\mu_{t+1} = j, l_{t+1} = j|\mathbf{l}_t)$, the denominator gives $P(l_{t+1} = j|\mathbf{l}_t)$, and the approximation results from assuming $\alpha \approx 1$. It makes intuitive sense that confidence in current context should increase whenever observation agrees with model expectation; in fact, a bit of algebraic manipulation reveals that λ_{t+1}^* always increases by a nonzero amount from λ_t^* in equation 2. h_{t+1}^* is updated by incrementing h_t^* by 1, and γ_{t+1}^* can be approximated as a simple running average: $\gamma_{t+1}^* = \gamma_t^* + ((1-\delta_{jk}) - \gamma_t^*)/h_{t+1}^* = \gamma_t^* - \gamma_t^*/h_{t+1}^*$, where δ_{jk} is the Kronecker delta. If $l_{t+1} = k \neq j$, the situation is more complicated, as there is a need to differentiate between the possibility of μ having changed from j to k, and the possibility of l_{t+1} simply being an outlier (we ignore the relatively minuscule possibility of μ_{t+1} being neither j nor k). The

first scenario corresponds to $A \equiv P(\mu_{t+1} = k, l_{t+1} = k|\mathbf{l}_t) \approx (\alpha(1-\lambda_t^*) + (1-\alpha)\lambda_t^*)(1-\gamma_0)/(n-1)$ and the second corresponds to $B \equiv P(\mu_{t+1} = j, l_{t+1} = k|\mathbf{l}_t) \approx \gamma_t^*\lambda_t^*/(n-1)$. It is clear that μ_{t+1}^* should be assigned k if A exceeds B, or equivalently, after some algebraic manipulations:

$$(1-\lambda_t^*) > \frac{\gamma_{t+1}^\circ}{1+\gamma_{t+1}^\circ - \gamma_0} \tag{3}$$

where $\gamma_{t+1}^\circ = \gamma_t^* + (1-\gamma_t^*)/(h_t^*+1)$ would be the new estimate of γ associated with the μ_t^*, if the context did not change. In addition, if γ_t^* ever exceeds γ_M, the system should also be alerted to a contextual change. In these cases, $\mu_{t+1}^* = k$ and the other quantities are reset as follows: $\lambda_{t+1}^* = A/(A+B)$, $\gamma_{t+1}^* = \gamma_0$, $h_{t+1}^* = 1$. Otherwise, if the probability of outlier dominates, then $\mu_{t+1}^* = j$ and $\lambda_{t+1}^* = B/(A+B)$. $h_{t+1}^* = h_t^* + 1$ as before, but now expected deviation increases: $\gamma_{t+1}^* = \gamma_t^* + ((1-\delta_{jk}) - \gamma_t^*)/h_{t+1}^* = \gamma_{t+1}^\circ$.

This approximate learning algorithm posits that ACh and NE will interact in interesting opponent and synergistic ways. 3 says that ACh level, or expected uncertainty, essentially sets the *threshold* that gates NE signaling: the more expected deviations there are in a particular context, the less likely that an outlier would alert the NE system into reporting a contextual change. On the other hand, a strong NE signal is capable of resetting the ACh level to a characteristic default level. These properties correspond to intuitive notions of how expected and unexpected uncertainties should interact.

4 Results

Simulation I: Posner's Task

As a sanity check, we simulate a special case of our spatial attention task that closely corresponds to the standard Posner task, and compare simulation results with experimental data. To link the model with behavior, we consider the behavioral measure validity effect (VE) mentioned in section 2. One natural formalization of VE is to make it proportional to the confidence in the prediction of the next location given all previous observations, or $\text{VE}_{t+1} \propto P(l_{t+1} = \mu_t^*|\mathbf{l}_t) \propto (1-\gamma_t^*)\lambda_t^*$. Thus, VE depends on two factors, the estimated probability of deviation, and the confidence in the current internal model.

Numerous studies have demonstrated that cholinergic pharmacological manipulations interact robustly with VE[7, 2, 12, 6]. In figure 3, we compare simulated cholinergic psychopharmacology to one such study[7]

(same behavioral paradigm as[13]), which administered cholinergic agonist nicotine and antagonist scopolamine at various dosages. The correspondence between experiment and theory is good: the monotonic relationships between VE and ACh are clear in both (though the graphs should not be compared literally).

As was discussed in section 2, because the task contingencies are kept constant throughout an experimental session, unexpected uncertainty does not play a significant role, and therefore it comes as no surprise that NE manipulations have been shown not to affect VE in Posner's task[14]. Likewise, as expected from our model assumptions, NE manipulation does not affect VE in our simulations (data not shown).

Simulation II: Approximation in Action

Simulations of the extended version of Posner's task make evident the critical and distinct roles played by ACh and NE. Figure 3(a) shows a sample run for a particular setting of the task with 4 stimulus locations. The approximate algorithm does a good job of learning the underlying, varying location patterns from the noisy data: both the estimate of the primary location (panel 1) and that of the deviation parameter (panel 2) quickly converge to the true values. Direct predictions about ACh and NE activities follow from the model traces in panel 2: ACh converges to the true γ, NE rises at pattern changes and also other chance accumulations of unexpected observations. Panel 3 shows the modeled VE for all the "outlier" trials over the course of the simulated session, and illustrates properties that can be verified in behavioral experiments, including in human psychophysics studies. In particular, note the strong dips when a string of unexpected outliers is observed, including when true state transitions occur and when outliers accumulate by chance. There is an inverse relationship between VE and the deviation parameter γ.

To quantify the performance of the approximate algorithm, we compare its predictive power to that of the exact algorithm and to a naive "bottom-up" algorithm, where no internal model of task is maintained except the simple assumption that the stimulus location on a given trial tends to be the same as that of the last trial. While the truly exact computation of the joint posterior in equation 1 is intractable, for reasons mentioned before, a very good approximation can be obtained by discretizing the γ space and turning the difficult integrations into tractable summations. The representational and computational complexity of this approximation means that it does not easily admit a cortical implementation, but it does provide a "gold standard" against which the performance of our ACh-NE controlled approximate algorithm can be compared.

Figure 3(b) shows the average absolute error of the three algorithms as a function of n, the number of locations. For γ drawn randomly from the interval $[0, 0.5]$, the ACh-NE algorithm performs only silghtly less well than the exact algorithm and much better than the bottom-up algorithm. The higher error rate at $n = 2$ is expected for the ACh-NE algorithm and the exact algorithm, since the 2-location problem has a particular confound: outliers more often accumulate consecutively at a single location, increasing the probability of detecting a false switch into a new state. The bottom-up algorithm, on the other hand, does not integrate information over time, and therefore is insensitive to n, and always performs at an error rate around 0.25, the average deviation rate. Figure 3(c) shows how the algorithms perform in a 4-location task as a function of γ, the deviation parameter. When the probability of deviation is less than 0.5, the ACh-NE algorithm works slightly worse than the exact algorithm and much better than the bottom-up one. When the probability of deviation dominates, the performance of the ACh-NE algorithm deteriorates to that of the bottom-up algorithm. When the deviation probability is maximal ($(n-1)/n$), all three algorithms perform at chance, as can be expected.

Simulation III: Psychopharmacology

Given the critical interactions between ACh and NE in the proposed spatial attention task, experimental manipulation of ACh and/or NE levels, either through psychopharmacology or lesions of the neuromodulatory nuclei, should result in rich patterns of impairments. Figure 4 shows the results of depleting (a) ACh and (b) NE independently and (c) jointly, at 30% of normal levels, on the same example run as in Figure 3(a). Suppressing ACh signaling leads to the underestimation of the amount of variation within a given context. Consequently, the significance of deviations from the primary location is exaggerated and causes the NE system to over-react with hyper "distractability." This results in the highly variable VE of panel 3, as the system oscillates between over-confidence and false detection of representational changes. NE depletion corresponds to excessive confidence in the internal model and results in perserverative behavior when there *has* been a pattern change. As a consequence of this reluctance to adapt to new environments, the ACh level, *ie* expected uncertainty, rises to take into account of all the accumulating evidence of deviation from the current model. Eventually, the ACh level may hit some threshold and alert the system to a representational change, but this process is protracted and inefficient compared to the case of normal NE functioning. VE levels reflect these deficits: the dramatic drop at the onset of a model switch disappears and is replaced by a

slow, gradual downward slide, corresponding to an increasingly more active ACh signal. Finally, the most interesting impairments come from joint depletion of ACh and NE: it seems from Figure 4(c) that combined depletion of ACh and NE might result in *less severe* impairments than either single depletion. Figure 4(d) shows that, indeed, under some circumstances, combined depletion of ACh and NE can actually *alleviate* impairments resulting from severe depletion of ACh alone (or NE alone, not shown). This results from non-linear interactions between ACh and NE that involve aspects of both synergy and opponency. Intuitively, the underestimation of variability corresponding to ACh depletion, as well as the subsequent hyperactive tendency to prompt the NE system into signaling representational changes, can be alleviated by inhibiting the NE system directly.

5 Discussion

We have presented a model of inference in a new spatial attention task (Bentley, personal communication). In the task, the context changes slowly, and specifies a probabilistic scheme for each trial. ACh reports the expected variability associated with each context; NE reports the degree of doubt as to the correctness of the current estimate of the context. After repeated trials associated with a particular context, the ACh level converges to a value reflecting the true variability in the context, while the NE level decreases to baseline. The performance of the approximate algorithm approaches that of the exact algorithm for a range of reasonable deviation probabilities. According to our model and simulations, NE and ACh are proposed to interact in a precisely specified and partly opponent, partly synergistic manner. Interesting consequences arise from manipulating them jointly and severally: under certain parameter settings, combined depletion can actually *ameliorate* the detrimental effects of either single depletion. This observation may have significant implications for the treatment of neurological diseases involving the depletion of either ACh or NE.

The computations of the approximate algorithm we have proposed are sufficiently simple for cortical implementation. We suggest that the frontal cortex may serve as the neural substrate for integrating the neuromodulatory signals, and representing and maintaining the internal model. The frontal cortex has dense recurrent connections with both the cholinergic and noradrenergic nuclei. Thus, it is the ideal site for the integration and updating of these various sources of signals. Besides the uncertainty signals carried by the neuromodulatory inputs, the only other two quantities needed for the computations of the approximate al-

gorithm are μ_t^* and h_t^*, both of which may well be represented and computed in the prefrontal cortex, given its well-established involvement in attentional control and working memory.

One major caveat of this work, and indeed our previous suggestions, is that, although it seems as if expected and unexpected uncertainty are very distinct quantities, it is in fact hard to draw a precise line between them. Uncertainty is only unexpected if it cannot be predicted from a model; but it is always hard to be sure that there is not a more sophisticated model which, fed with the same information, would make more proficient predictions. Such models effectively capture uncertainties about the uncertainties, thereby allowing otherwise unexpected uncertainties to be predicted. Indeed, Bayesian statistical modeling[4] often uses hierarchies of uncertainties about parameters in just such a way. Since even human subjects are not always aware of the degree of sophistication of the models they might implicitly be employing, it is hard to be exactly sure what is expected and what is not. Experimental measures such as reaction times might provide independent evidence, elucidating the levels of inference employed.

This model represents an advance on our previous suggestions. Our first model[15] only treated ACh, and used a task with just a single underlying form of uncertainty. We suggested that ACh reports this uncertainty to control cortical inference (there being no requirement for plasticity in the task); however, there, the particular form of uncertainty is similar to the unexpected uncertainty, reported by NE, in the present model. This is a good illustration of the point raised above about the lack of a completely clean distinction between expected and unexpected uncertainty. The present model is closer to our second model[16], though there, ACh reports the uncertainty associated with the estimate of a parameter, whereas here, ACh reports the value of a parameter which itself conveys information about uncertainty in the task. In more complicated tasks, both aspects may be important, and so should influence the level of ACh. The degree to which different projections of ACh can report different information to different cortical and sub-cortical locations is as yet experimentally unclear. Similarly, the way that prediction error controls NE levels subtly differs between this and the previous model; again, more experimental evidence is required.

Our model makes qualitative, and even quantitatively precise, predictions about the task of figure 1(b). The task entails the construction and maintenance of hierarchical representations, and behavioral measures such as reaction time may provide revealing insights into the various components during the course of learning. Further, since the task is not substantially more difficult for monkeys or rats than standard versions, it should be possible to measure ACh and NE activity, either electrophysiologically or by dialysis, for comparison

with traces such as those in figures 3 and 4. Finally, we have discussed the predictions made by the model about the effects of suppressing (or boosting) the levels of ACh and/or NE. In particular, we expect that suppressing NE and ACh at the same time should have a complex pattern of effects on behavior, reflecting the two neuromodulators' part-synergistic, part-antagonistic nature of interactions.

References

[1] M F Bear and W Singer. Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature*, 320(6058):172–6, 1986.

[2] A A Chiba, P J Bushnell, W M Oshiro, and M Gallagher. *NeuroReport*, 10:3119–23, 1999.

[3] S A Huettel, P B Mack, and G McCarthy. *Nature Neuroscience*, 5(5):485–90, 2002.

[4] M I Jordan, Z Ghahramani, TS Jaakkola, and L K Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37:183–233.

[5] J McGaughy, M Sandstrom, S Ruland, J P Bruno, and M Sarter. *Behav Neurosci*, 111:646–52, 1997.

[6] P M Parasuraman, R Greenwood, J V Haxby, and C L Grady. Visuospatial attention in dementia of the Alzheimer type. *Brain*, 115:711–33, 1992.

[7] J M Phillips, K McAlonan, W G K Robb, and V Brown. *Psychopharmacology*, 150:112–6, 2000.

[8] M I Posner. Orienting of attention. *Q J Exp Psychol*, 32:3–25, 1980.

[9] T W Robbins and B J Everitt. *The Cognitive Neurosciences*, chapter Arousal systems and attention. The MIT Press, 1995.

[10] S J Sara and M Segal. Plasticity of sensory responses of LC neurons in the behaving rat: implications for cognition. *Progr in Brain Research*, 88:571–85, 1991.

[11] J Turchi and M Sarter. *Neuroscience*, 104:407–17, 2001.

[12] M L Voytko, D S Olton, R T Richardson, L K Gorman, J R Tobin, and Price D L. *J Neurosci*, 14:167–86, 1994.

[13] N M Ward and V J Brown. Covert orienting of attention in the rat and the role of striatial dopamine. *J Neurosci*, 16:3082–8, 1996.

[14] E A Witte and R T Marrocco. *Psychopharmacology*, 132:315–23, 1997.

[15] A J Yu and P Dayan. Acetylcholine in cortical inference. *Neural Networks*, 15:719–30, 2002.

[16] A J Yu and P Dayan. Expected and unexpected uncertainty: ACh & NE in the neocortex. *Advances in Neural Processing Systems 15*, 2003.

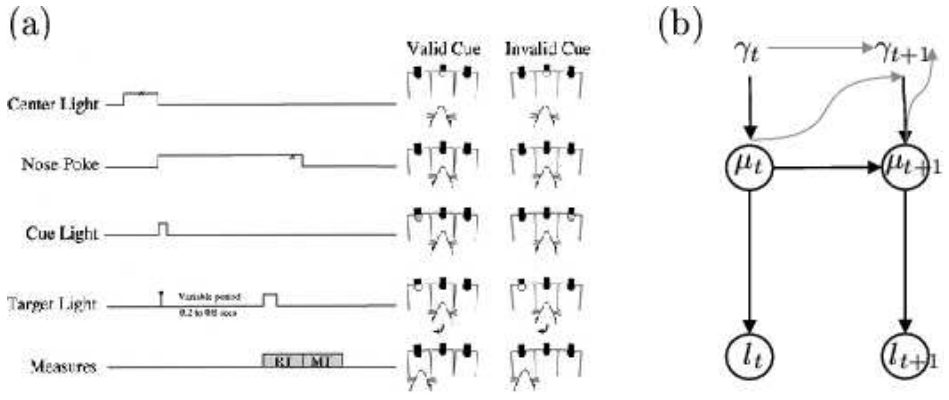

Figure 1: Experimental design. (a) An example of Posner's task adapted to rats, figure from[13]. Trial begins nose poke, terminates with reward retrieval. RT: reaction time, MT: movement time. (b) Schematic diagram of the generative model used to generate stimulus patterns in our spatial attention task.

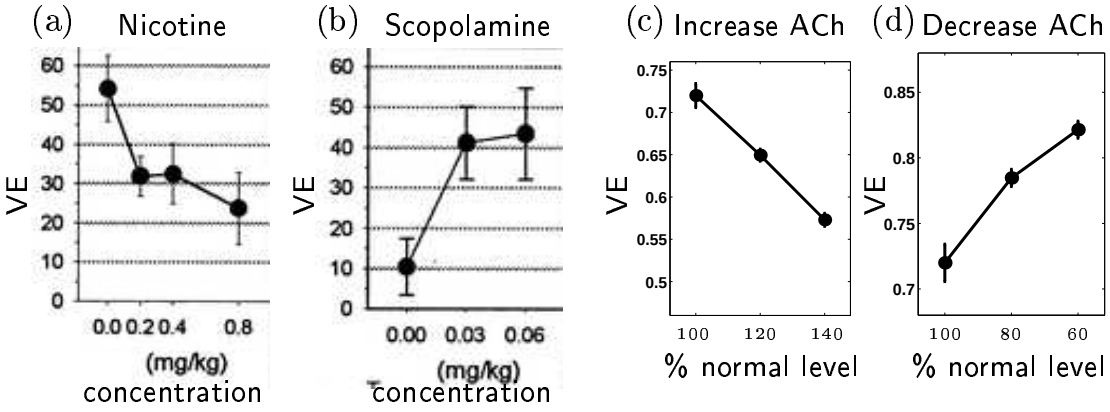

Figure 2: Posner's task. Experimental data[7] indicate (a) VE interacts with nicotine level dose-dependently, and (b) with scopolamine dose-dependently. (c) & (d) Similar interactions in simulation. 50% validity in both experiment and simulation.

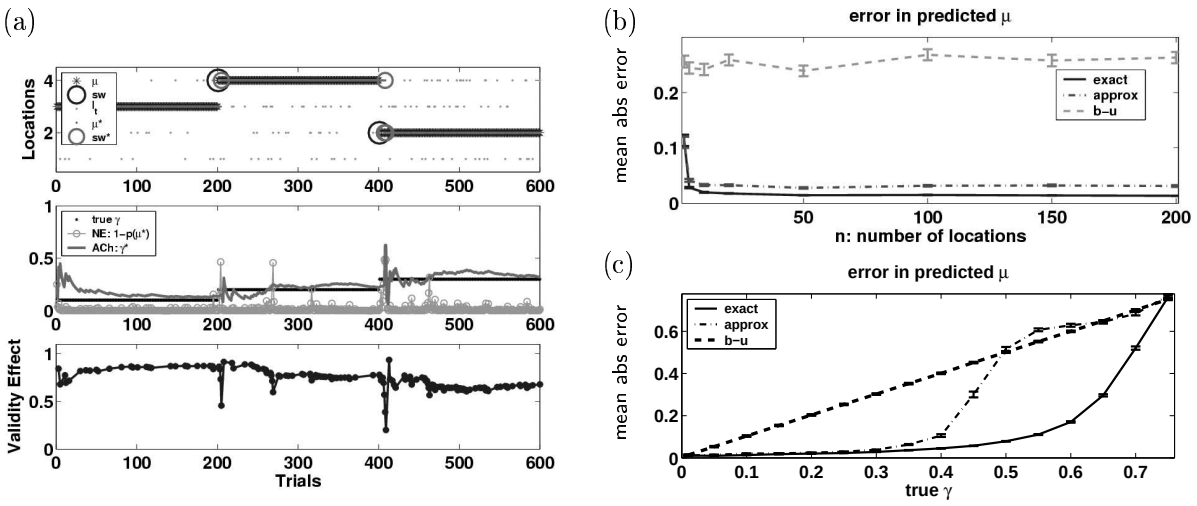

Figure 3: Quality of approximate learning, as a function of n and γ.. (a) Sample run of the approximate learning algorithm. SW: switches; *: estimated quantities. Simulation parameters: $n = 4$, $\gamma = 0.1, 0.2, 0.3$ in first, second, and third thirds of the trials, respectively. $\alpha = 0.99$ here and in all following simulations. (b) Approximate learning compared to exact and bottom-up algorithms, as a function of n. $\gamma \sim U[0, 0.5]$. (c) Approximate learning compared to exact and bottom-up algorithms, as a function of γ. $n = 4$. Error bars: standard deviation over 200 sessions of 1000 trials each.

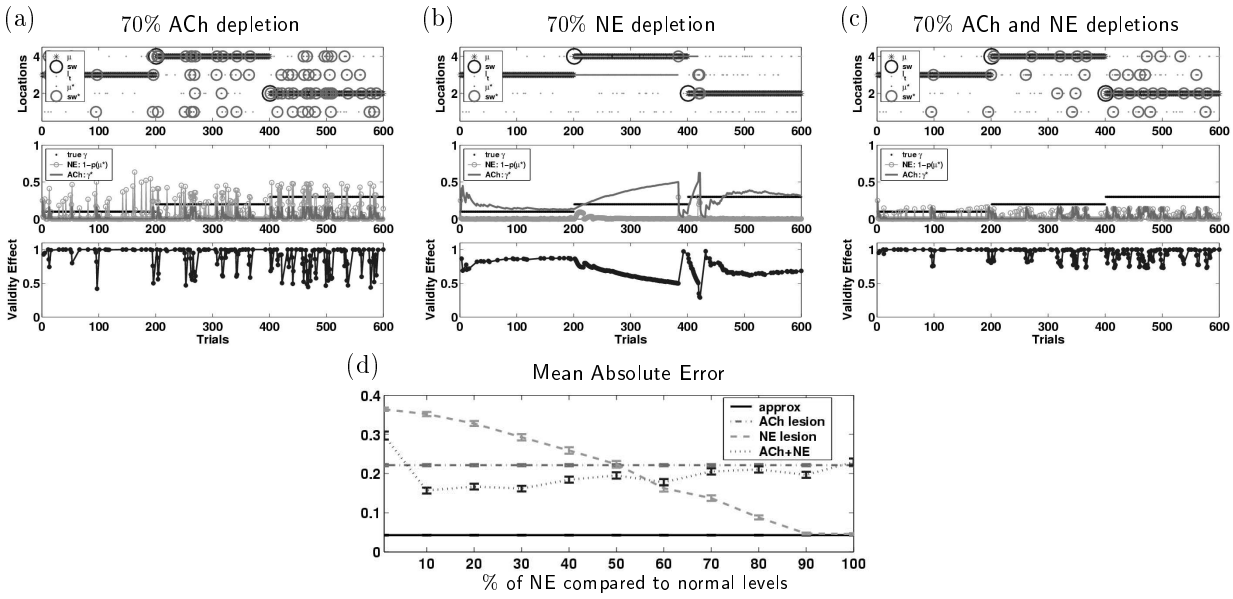

Figure 4: Simulated psychopharmacology. (a) 70% depletion of ACh (to 30% of normal levels). (b) 70% depletion of NE. (c) 70% combined depletion of ACh and NE. (d) 99% ACh depletion, varying levels of NE depletion. Absolute error in predictions of μ_{t+1} given all past observations $l_1, ..., l_t$. Combined depletion can actually alleviate severe single depletion of ACh, indicating presence of opponent interactions between ACh and NE. Error bars: standard deviation over 300 sessions of 1000 trials each.