Uncertainty and the Bayesian Brain

- sources:
  - sensory/processing noise
  - ignorance
  - change
- consequences:
  - inference
  - learning
- coding:
  - distributional/probabilistic population codes
  - neuromodulators

Multisensory Integration

- $v = s + \varepsilon_v \quad ; \qquad t = s + \varepsilon_t$
- $p[s|v,t] \propto p[v,t|s]p[s] \sim p[v|s]p[t|s]$
- $\varepsilon_v \sim N\left[0, {\rho_v}^{-2}\right] \quad ; \varepsilon_t \sim N\left[0, {\rho_t}^{-2}\right]$
- $s|v,t \;\sim\; N[\mu_s, {\rho_s}^{-2}]$
- $\mu_s = \frac{v{\rho_v}^2 + t{\rho_t}^2}{{\rho_v}^2 + {\rho_t}^2} \quad ; {\rho_s}^2 = {\rho_v}^2 + {\rho_t}^2$

$$\mathbf{v}: p[s|\mathbf{v}] \sim \mathcal{N}\left[\frac{\sum_a v_a s_a}{\sum_b v_b}, \frac{\sigma^2}{\sum_b v_b}\right] \quad \mathbf{t}: p[s|\mathbf{t}] \sim \mathcal{N}\left[\frac{\sum_a t_a s_a}{\sum_b t_b}, \frac{\sigma^2}{\sum_b t_b}\right]$$

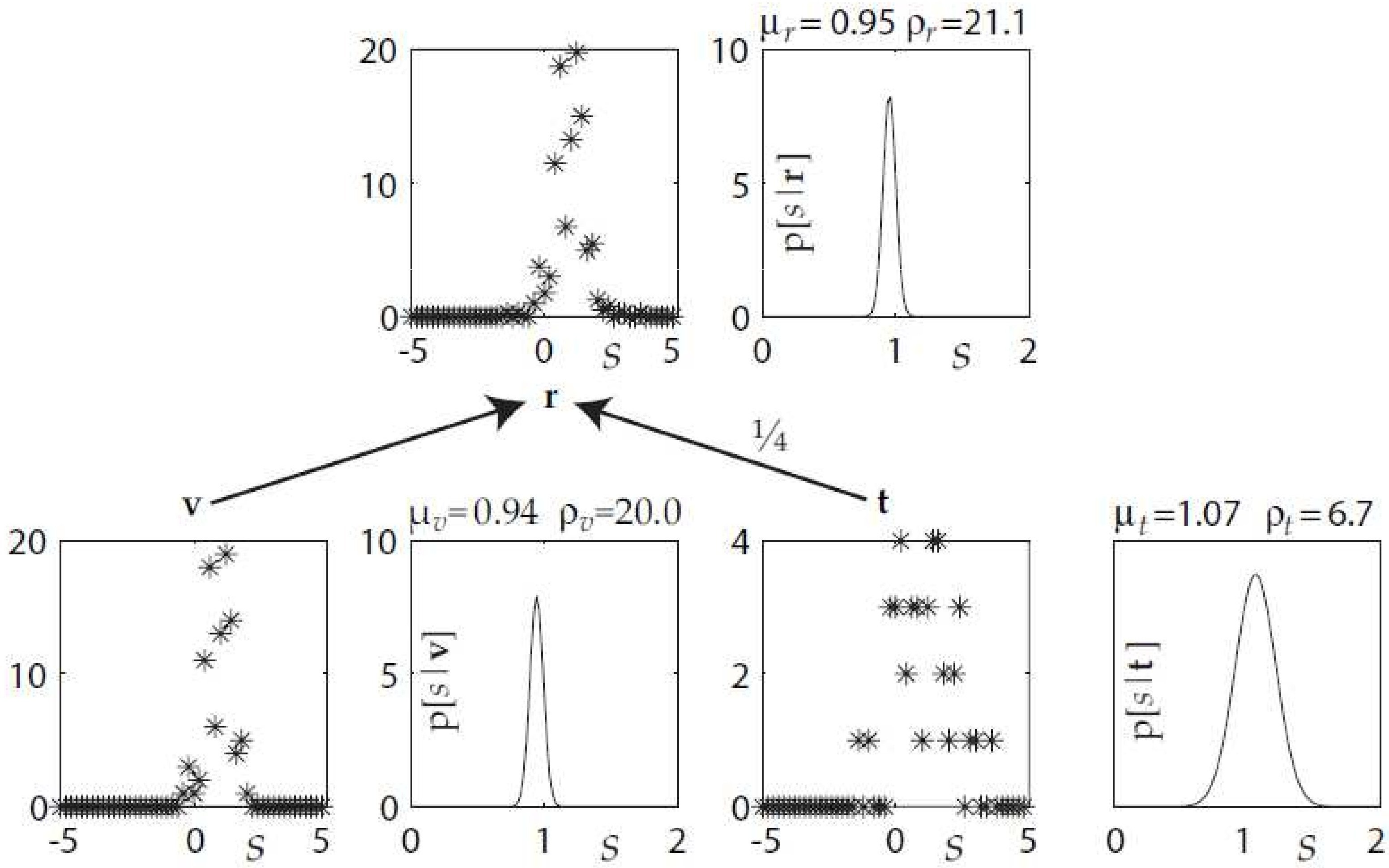
apply the previous analysis:

$$p[s|\mathbf{v},\mathbf{t}] \sim \mathcal{N}\left[\frac{\frac{\sum_a v_a s_a}{\sum_b v_b}\frac{\sum_b v_b}{\sigma^2} + \frac{\sum_a t_a s_a}{\sum_b t_b}\frac{\sum_b t_b}{\sigma^2}}{\frac{\sum_b v_b}{\sigma^2} + \frac{\sum_b t_b}{\sigma^2}}, \frac{1}{\frac{\sum_b v_b}{\sigma^2} + \frac{\sum_b t_b}{\sigma^2}}\right]$$

$$\sim \mathcal{N}\left[\frac{\sum_a (v_a + t_a) s_a}{\sum_b v_b + t_b}, \frac{\sigma^2}{\sum_b v_b + t_b}\right]$$

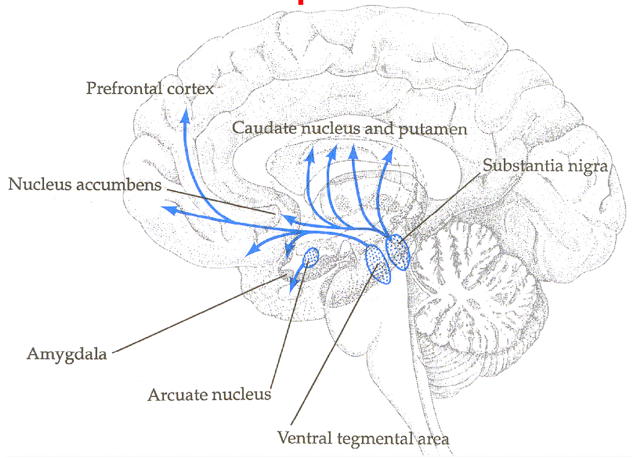
so if: $r_a = v_a + t_a$ everything will work out

Explicit and Implicit Spaces

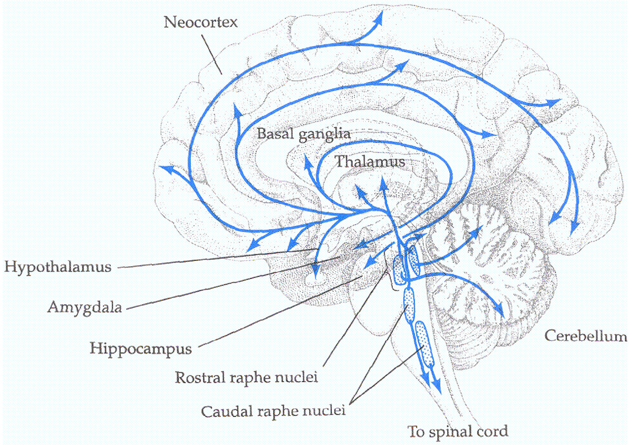


Computational Neuromodulation

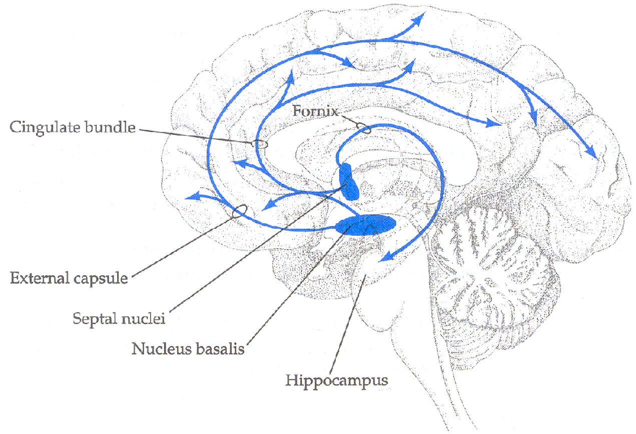
dopamine



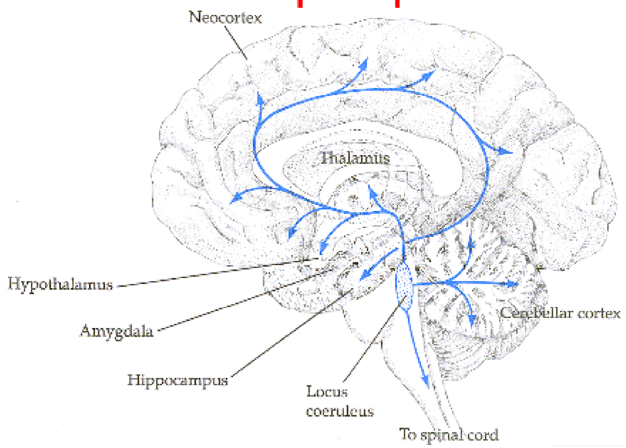
5HT



acetylcholine



norepinephrine



- general: excitability, signal/noise ratios
- specific: prediction errors, uncertainty signals

Uncertainty

Computational **functions** of neuromodulatory uncertainty:

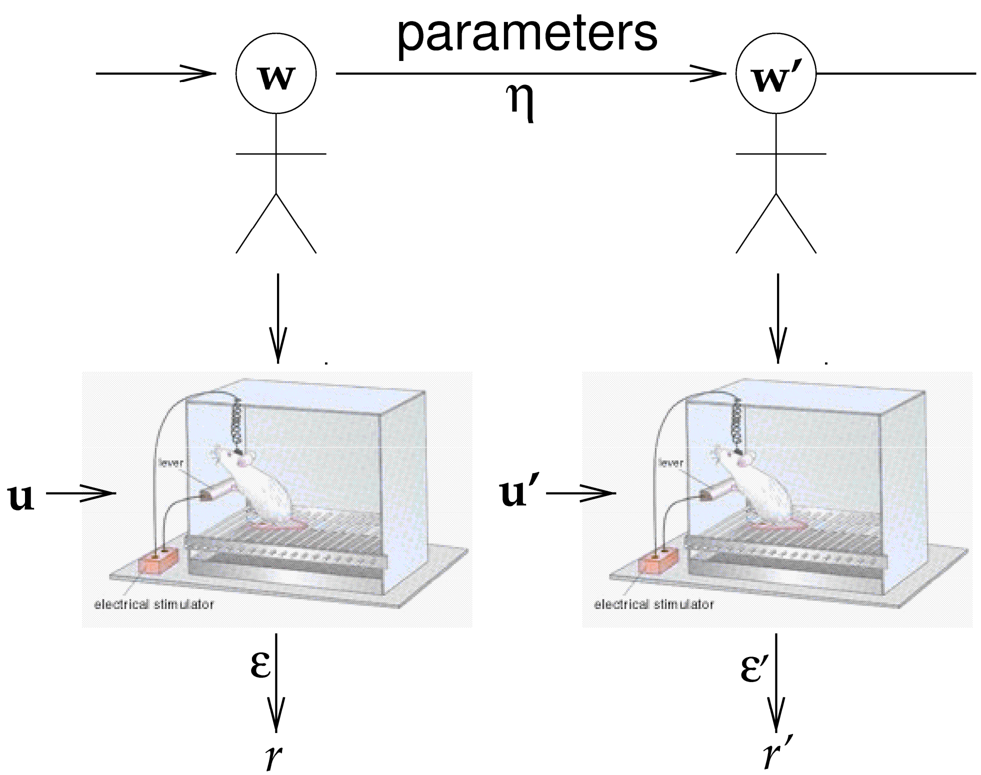
- weaken ***top-down*** influence over sensory processing
- promote ***learning*** about the relevant representations

We focus on **two** different kinds of uncertainties:

ACh - ***expected uncertainty*** from known variability or ignorance

NE - ***unexpected uncertainty*** due to gross mismatch between prediction and observation

Kalman Filter



expt $\mathbf{w}' = \mathbf{w} + \boldsymbol{\eta}$

reward given $r = \mathbf{w} \cdot \mathbf{u} + \epsilon$

allowable drift $\boldsymbol{\eta} \sim N[0, \sigma^2 \mathbb{I}]$

output noise $\epsilon \sim N[0, \rho^2]$

- Markov random walk (or OU process)
- no punctate changes
- additive model of combination
- forward inference

Kalman Posterior

The Kalman filter maintains uncertainty:

$$P(V) = \mathcal{N}[\widehat{\mathbf{w}} \cdot \mathbf{u}, \mathbf{u} \cdot \boldsymbol{\Sigma} \cdot \mathbf{u}]$$

where

Assumed Density KF

Diagonal approx to $\Sigma = \text{diag}(\sigma_i^2)$

If $\mathbf{w} \sim \mathcal{N}\left[\hat{\mathbf{w}}, \text{diag}(\sigma_i^2)\right]$, then

$$\Delta \hat{w}_i = \frac{\sigma_i^2}{\sum_j \sigma_j^2 + \rho^2}(r - \mathbf{u} \cdot \hat{\mathbf{w}})u_i$$

- Rescorla-Wagner error correction
- competitive allocation of learning
  - P&H, M

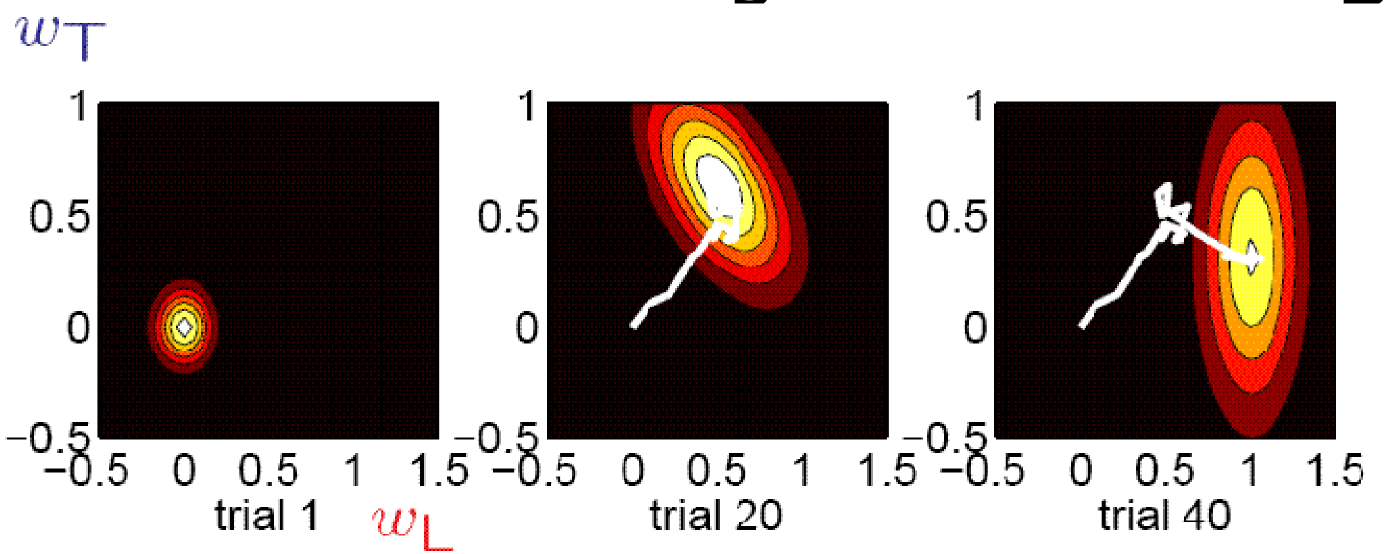
Blocking

forward	L$\to$r	L+T$\to$r	T$\to\cdot$
backward	L+T$\to$r	L$\to$r	$T\to\cdot$

- forward blocking: error correction

$$(r - \mathbf{u}\cdot\hat{\mathbf{w}})$$

- backward blocking: -ve **off-diag** $\Sigma_{LT} < 0$



Mackintosh vs P&H

- under diagonal approximation:

$$\mathsf{E}(r - \mathbf{u} \cdot \hat{\mathrm{w}})^2 = \rho^2 + \sum_j \sigma_j^2 u_i^2$$

- for slow learning,

σ_j^2 changes with correlation of $(r - V)$ and u_i

  - effect like Mackintosh

Summary

- Kalman filter models many standard conditioning paradigms
- elements of RW, Mackintosh, P&H
- but:
  - downwards unblocking

$$L \to r \Delta r \qquad L + T \to r \qquad T \nrightarrow \pm r$$

predictor competition

  - negative patterning $L \to r$; $T \to r$; $L+T \to \cdot$

stimulus/correlation rerepresentation (Daw)

  - recency vs primacy (Kruschke)

Experimental Data

ACh & **NE** have similar ***physiological*** effects

- ***suppress*** recurrent & feedback processing
  (*e.g.* Kimura *et al*, 1995; Kobayashi *et al*, 2000)
- ***enhance*** thalamocortical transmission
  (*e.g.* Gil *et al*, 1997)
- ***boost*** experience-dependent plasticity
  (*e.g.* Bear & Singer, 1986; Kilgard & Merzenich, 1998)

ACh & **NE** have distinct ***behavioral*** effects:

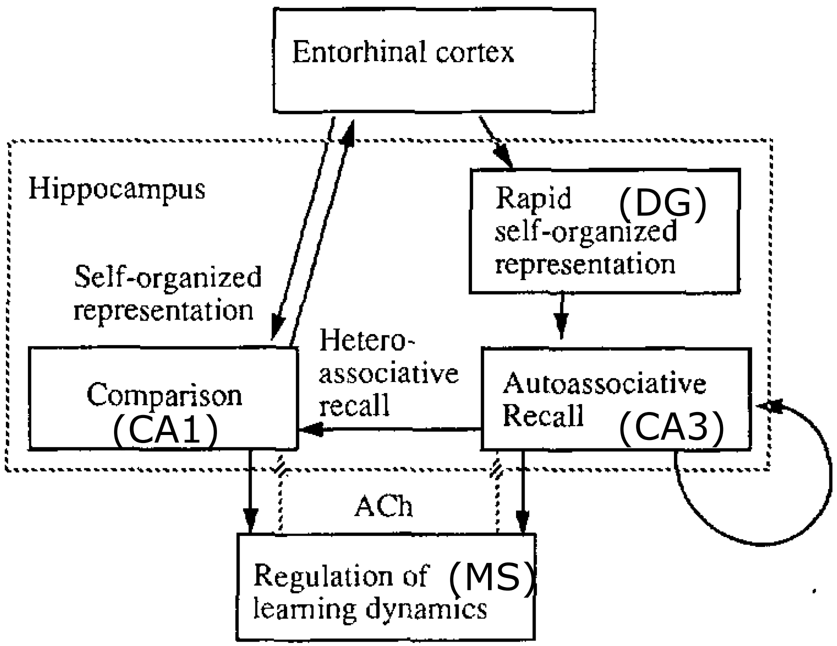
- ACh ***boosts*** learning to stimuli with uncertain consequences (*e.g.* Bucci, Holland, & Gallagher, 1998)

- NE ***boosts*** learning upon encountering global changes in the environment (*e.g.* Devauges & Sara, 1990)

ACh in Hippocampus

Given ***unfamiliarity***, **ACh**:

- ***boosts*** bottom-up, ***suppresses*** recurrent processing
- ***boosts*** recurrent plasticity



(Hasselmo, 1995)

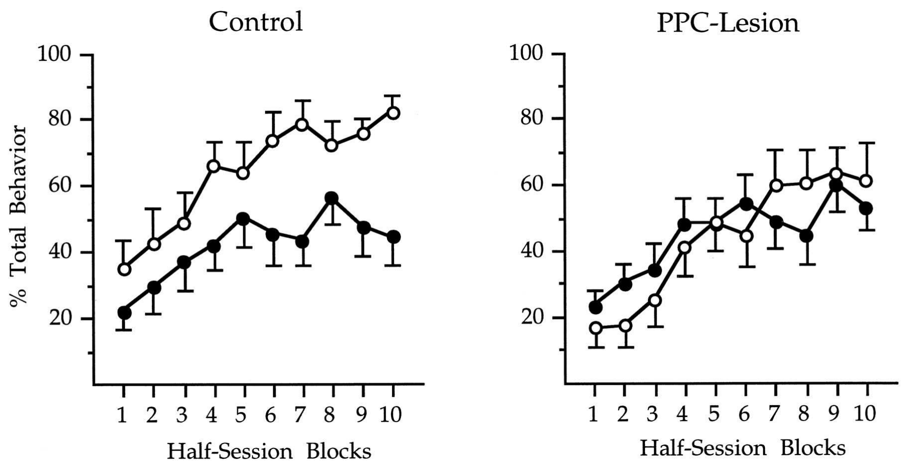
ACh in Conditioning

Given ***uncertainty***, **ACh**:

- ***boosts*** learning to stimuli of uncertain consequences

Table 1. Outline of procedures for Experiment 1

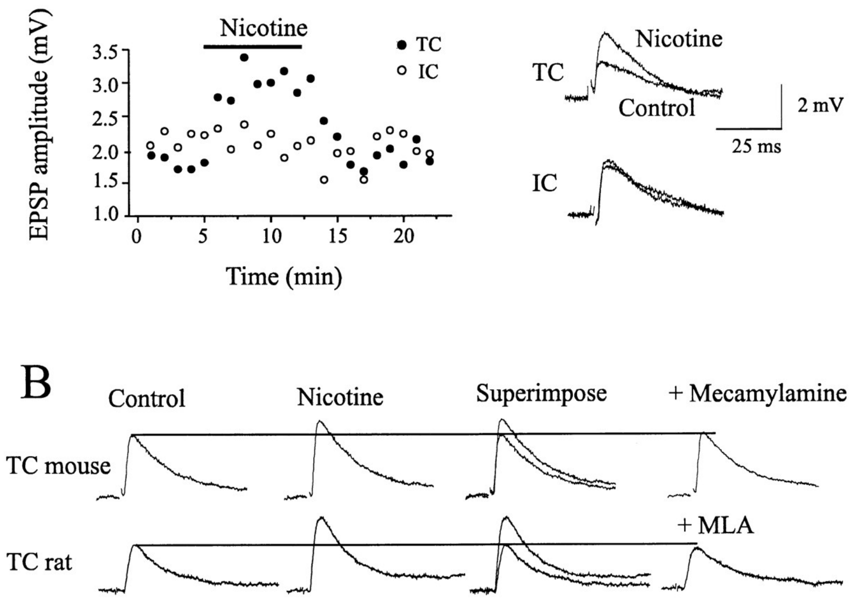
Treatment condition (groups)	Phase 1: consistent L–T relation	Phase 2: experimental change in L–T relation	Phase 3: test of conditioning to L
Consistent (CTL-C, PPC-C)	L → T → food; L → T	L → T → food; L → T	L → food
Shift (CTL-S, PPC-S)	L → T → food; L → T	L → T → food; L	L → food



(Bucci, Holland, & Galllagher, 1998)

Cholinergic Modulation in the Cortex

Electrophysiology Data



(Gil, Conners, & Amitai, 1997)

ACh agonists:

- ***facilitate*** TC transmission
- ***enhance*** stimulus-specific activity

Examples of Hallucinations Induced by Anticholinergic Chemicals

Scopolamine in normal volunteers	**Integrated, realistic hallucinations with familiar objects and faces**	**Ketchum et al. (1973)**
Intravenous atropine in bradycardia	**Intense visual hallucinations on eye closure**	**Fisher (1991)**
Local application of scopolamine or atropine eyedrops	**Prolonged anticholinergic delirium in normal adults**	**Tune et al. (1992)**
Side effects of motion-sickness drugs (scopolamine)	**Adolescents hallucinating and unable to recognize relatives**	**Wilkinson (1987)** **Holland (1992)**

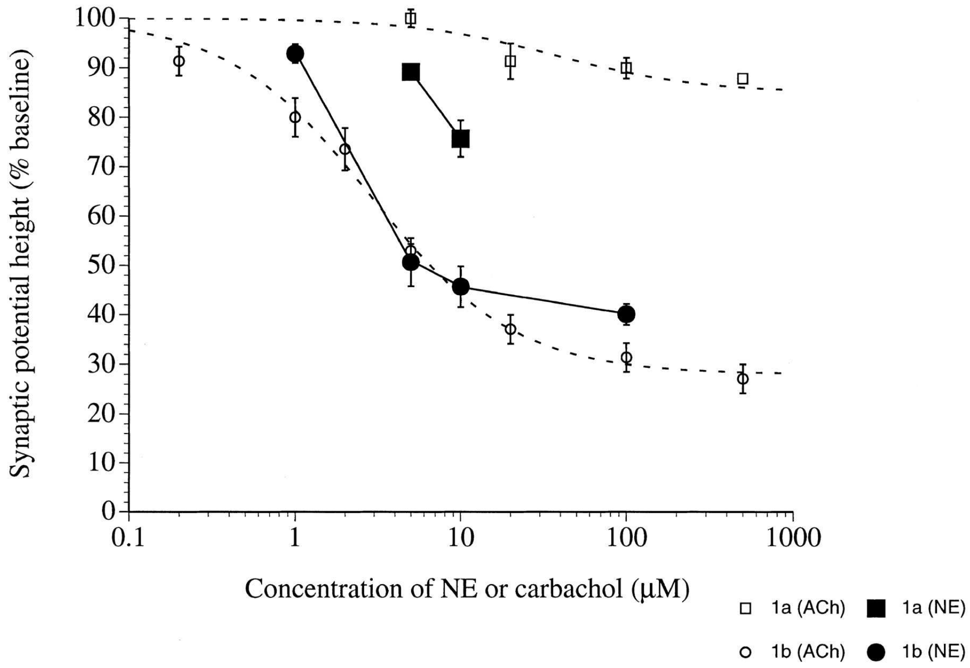
(Perry & Perry, 1995)

ACh antagonists:

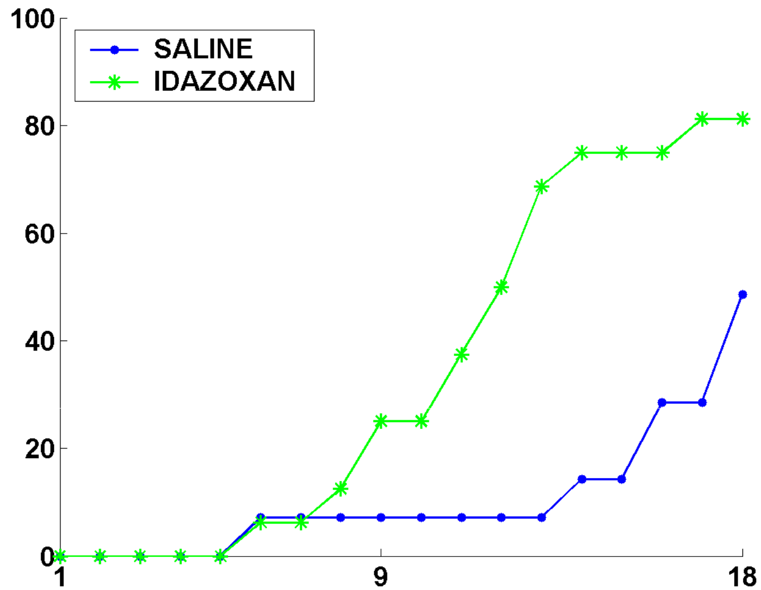
- ***induce*** hallucinations
- ***interfere*** with stimulus processing
- effects ***enhanced*** by eye closure

Norepinephrine

Something similar may be true for NE (Kasamatsu *et al*, 1981)



(Hasselmo *et al*, 1997)

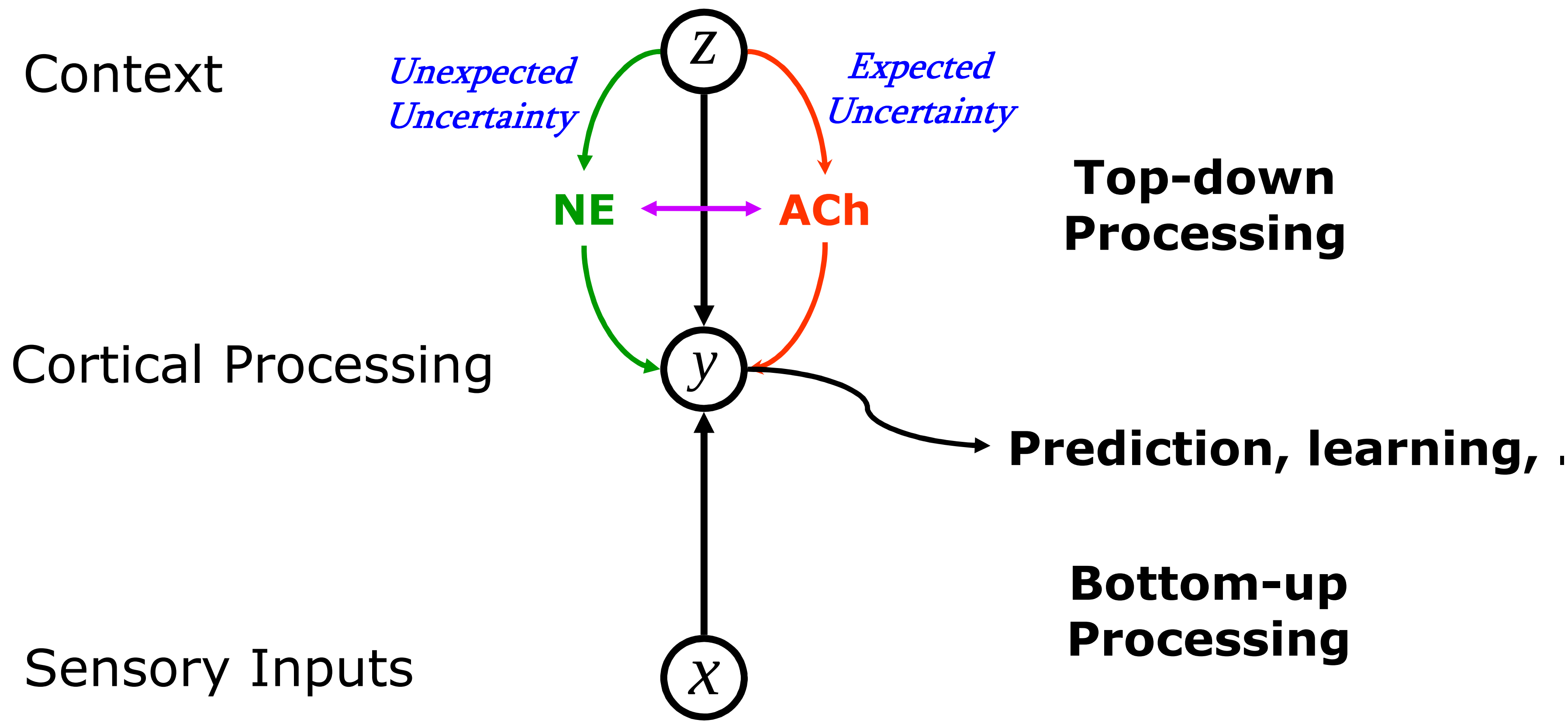


Days after task shift

(Devauges & Sara, 1990)

NE specially involved in **novelty,** confusing association with attention, vigilance

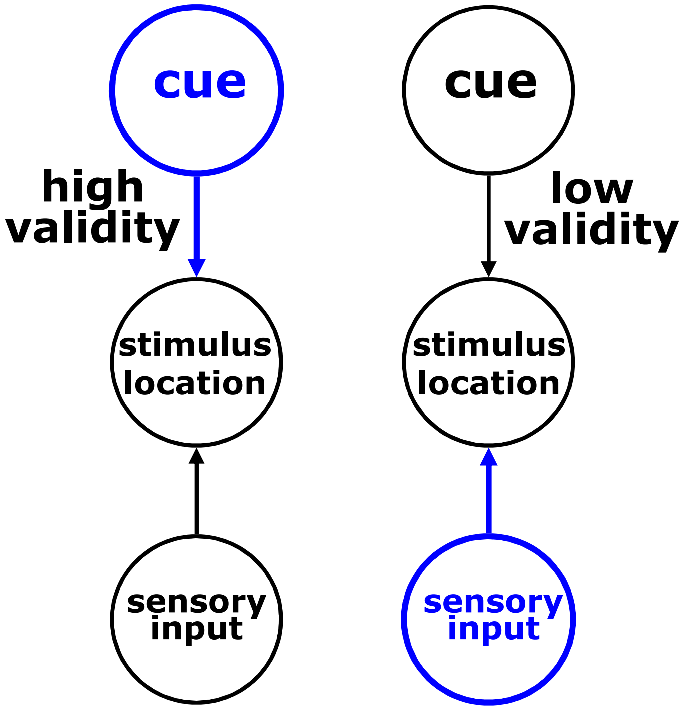
Model Schematics

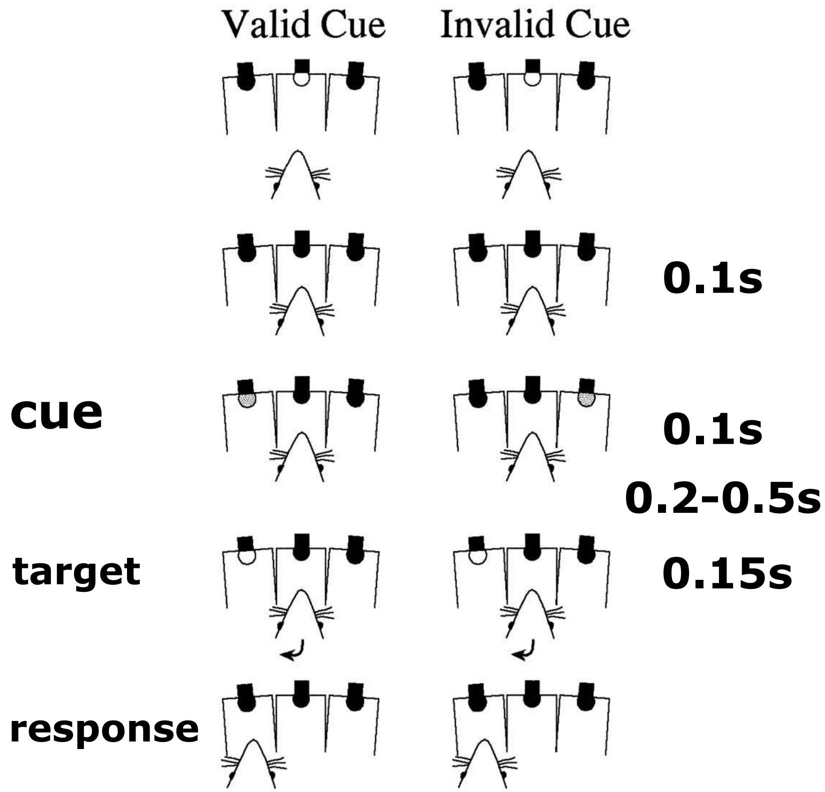


Attention

Attentional selection for (statistically) **optimal** processing, above and beyond the traditional view of resource constraint

Example 1: Posner's Task





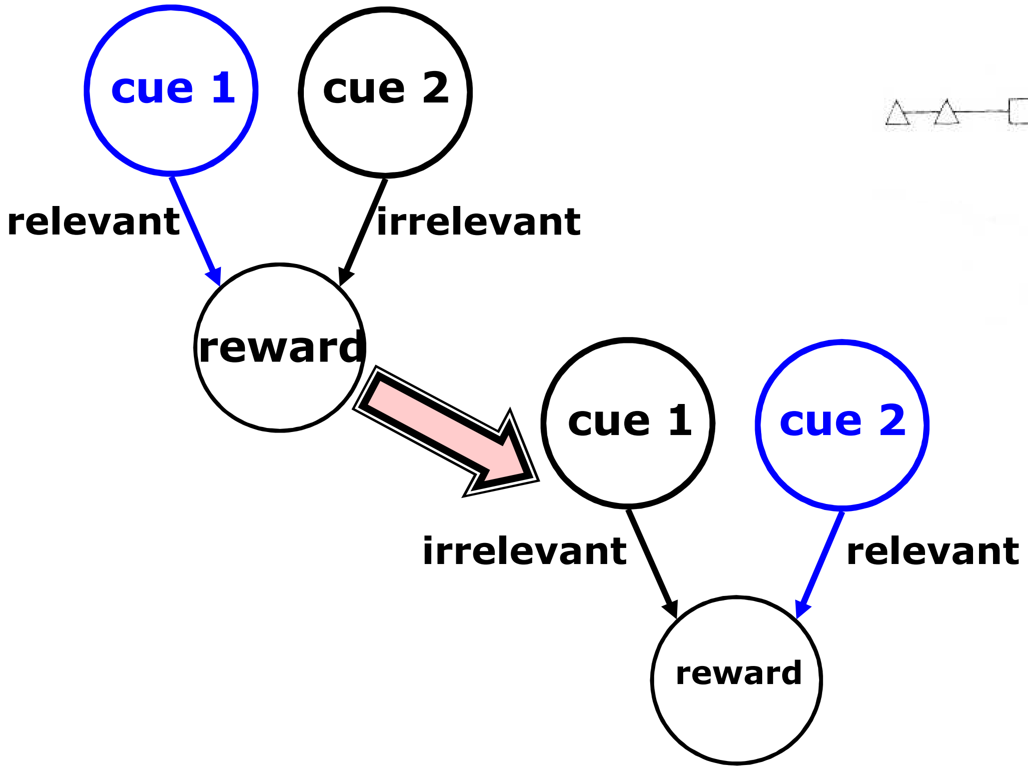
(Phillips, McAlonan, Robb, & Brown, 2000)

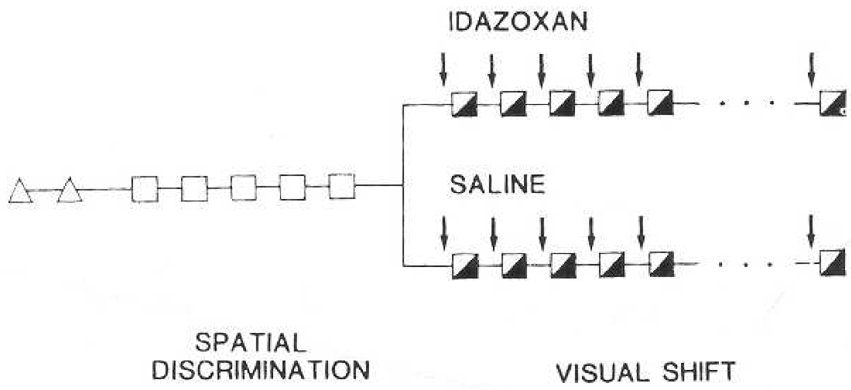
Uncertainty-driven bias in cortical processing

Attention

Attentional selection for (statistically) **optimal** processing, above and beyond the traditional view of resource constraint

Example 2: Attentional Shift





(Devauges & Sara, 1990)

Uncertainty-driven bias in cortical processing

A Common Framework

NE — Variability in **identity** of relevant cue

$$1-\lambda_t^*$$

ACh — Variability in **quality** of relevant cue

$$1-\gamma_t^*$$

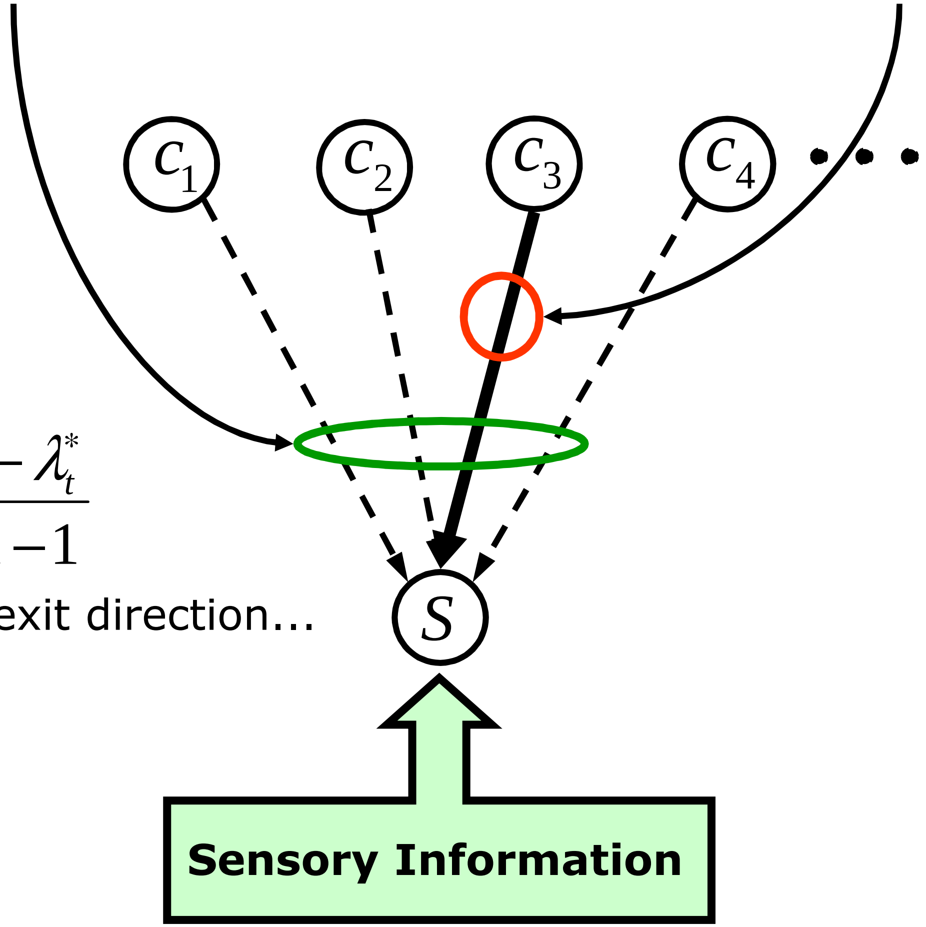
Cues: vestibular, visual, ...

$$\mu_t^* = i$$

$$P^*(\mu_t^* \mid D_t) = \lambda_t^*$$

$$P^*(\mu_t = j \neq i \mid D_t) = \frac{1-\lambda_t^*}{h-1}$$

Target: stimulus location, exit direction...

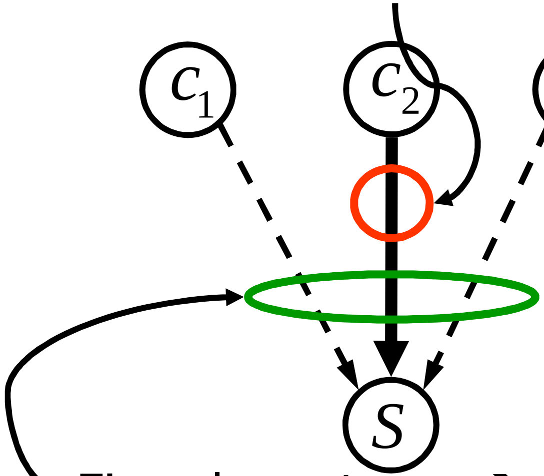


avoid representing

full uncertainty

Simulation Results: Posner's Task

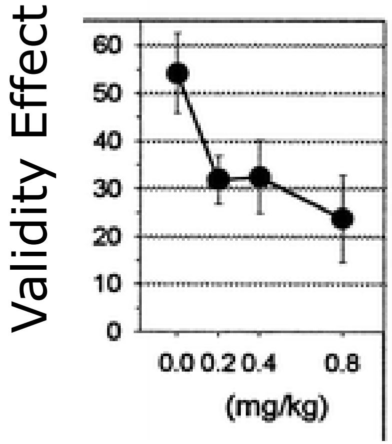
Vary cue validity ➔ Vary **ACh**



Fix relevant cue ➔ low **NE**

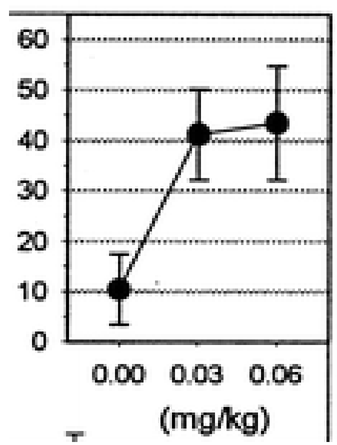
$$VE \propto (1-NE)(1-ACh)$$

Nicotine



Concentration

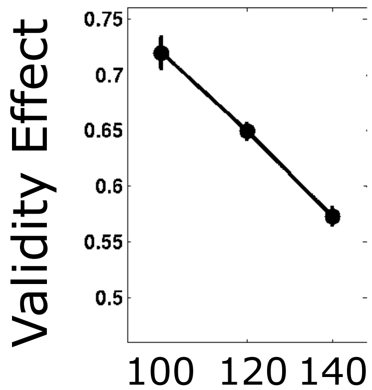
Scopolamine



Concentration

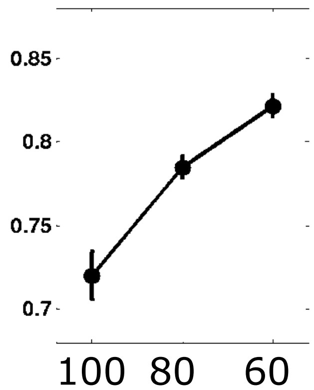
(Phillips, McAlonan, Robb, & Brown, 2000)

Increase ACh



% normal level

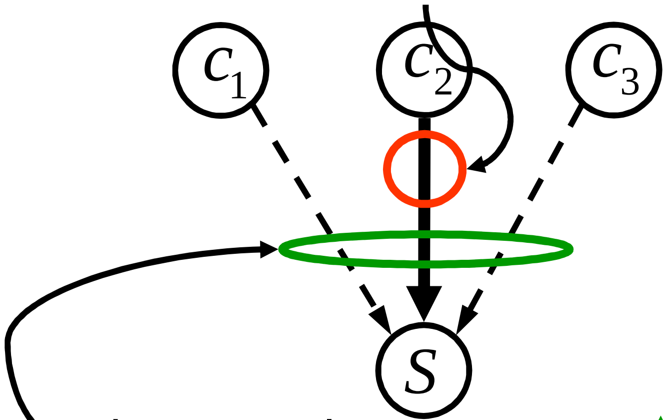
Decrease ACh



% normal level

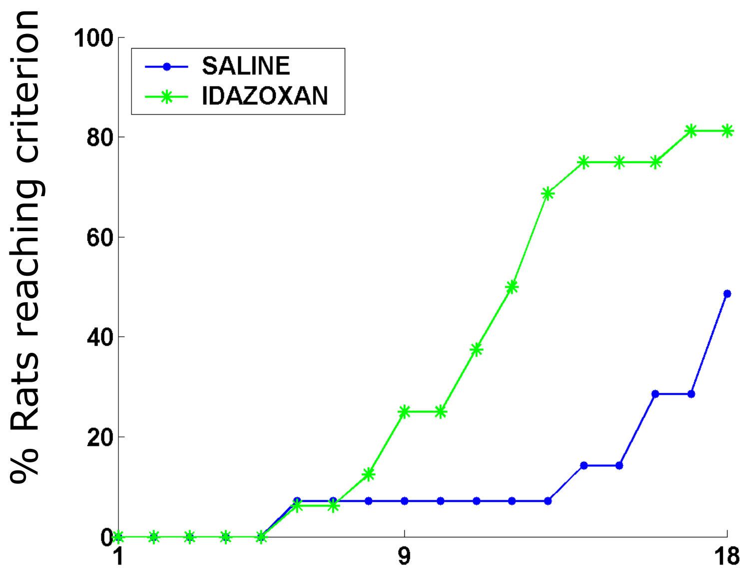
Simulation Results: Maze Navigation

Fix cue validity ➔ no explicit manipulation of **ACh**



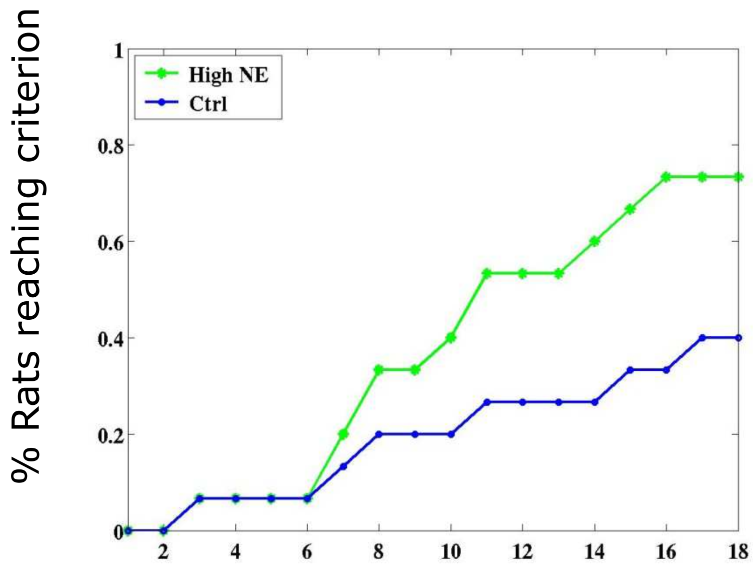
Change relevant cue ➔ ↑ **NE**

Experimental Data



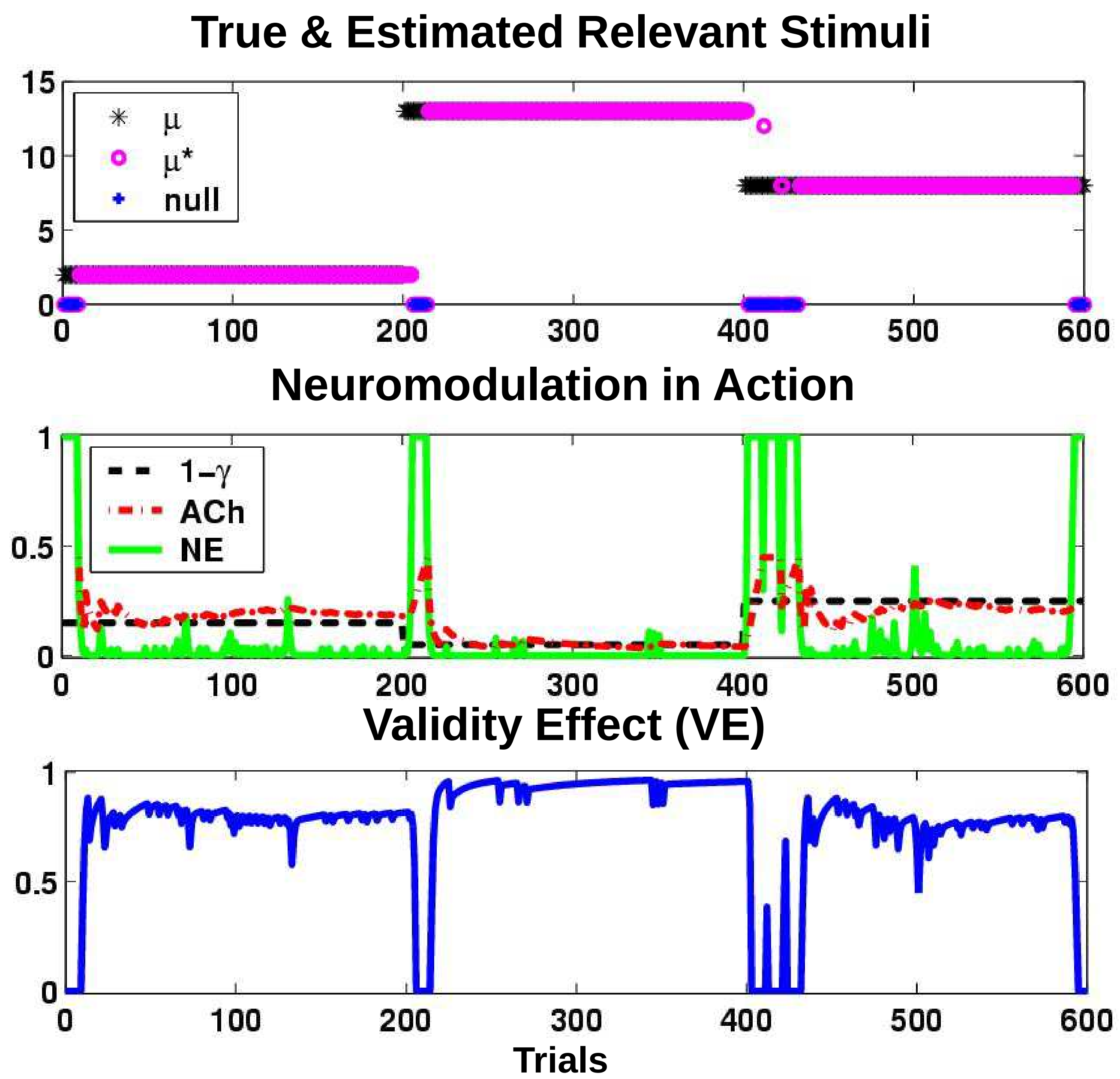
No. days after shift from spatial to visual task
(Devauges & Sara, 1990)

Model Data



No. days after shift from spatial to visual task

Simulation Results: Full Model



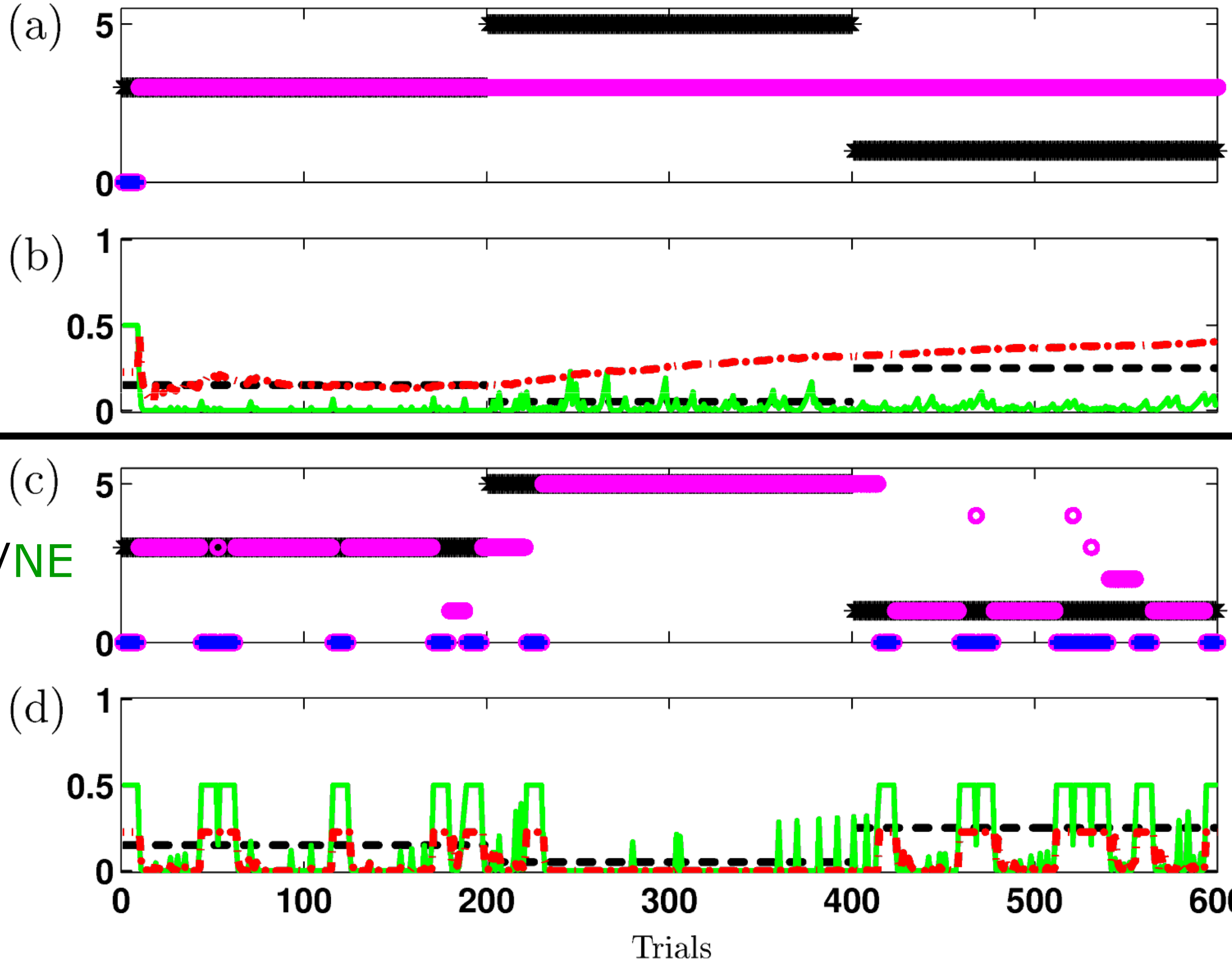
Simulated Psychopharmacology

50% NE

ACh compensation

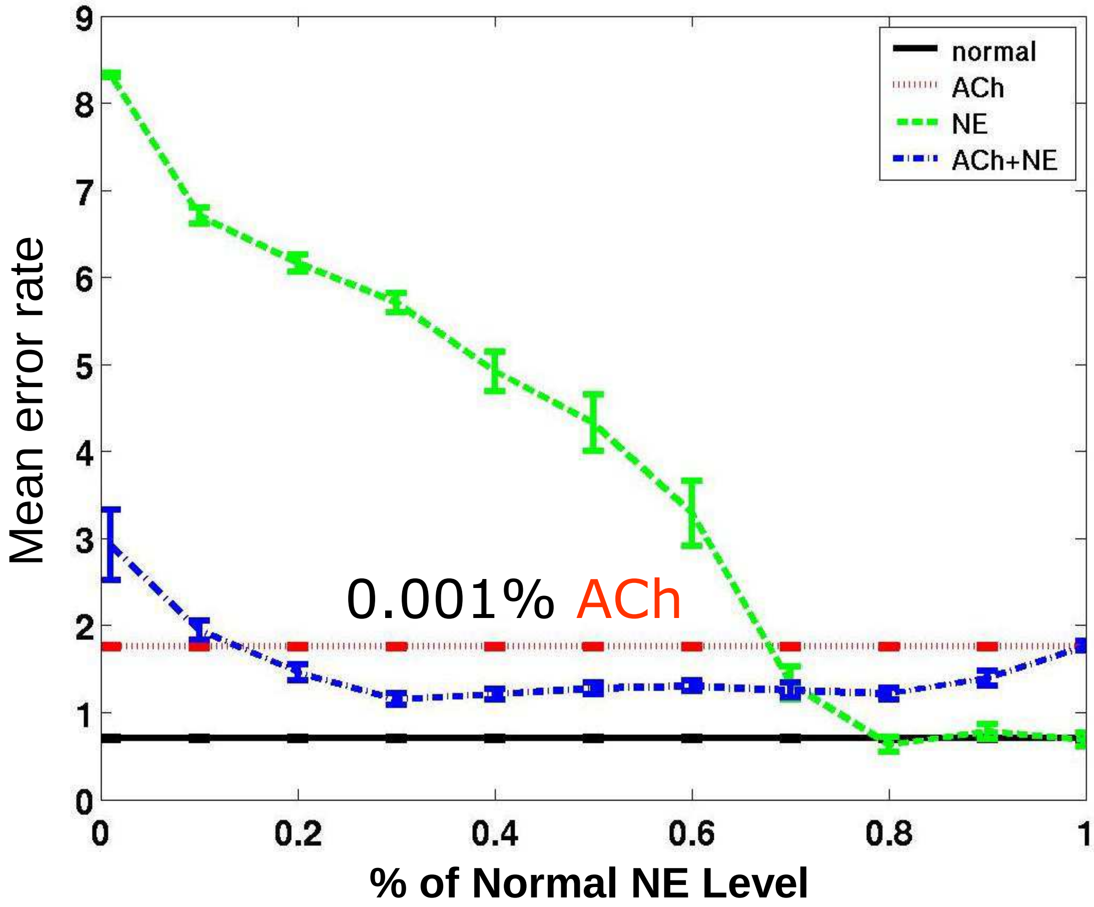
50% ACh/NE

NE can nearly catch up



Simulation Results: Psychopharmacology

NE depletion can ***alleviate*** ACh depletion revealing underlying ***opponency*** (implication for neurological diseases such as Alzheimers)

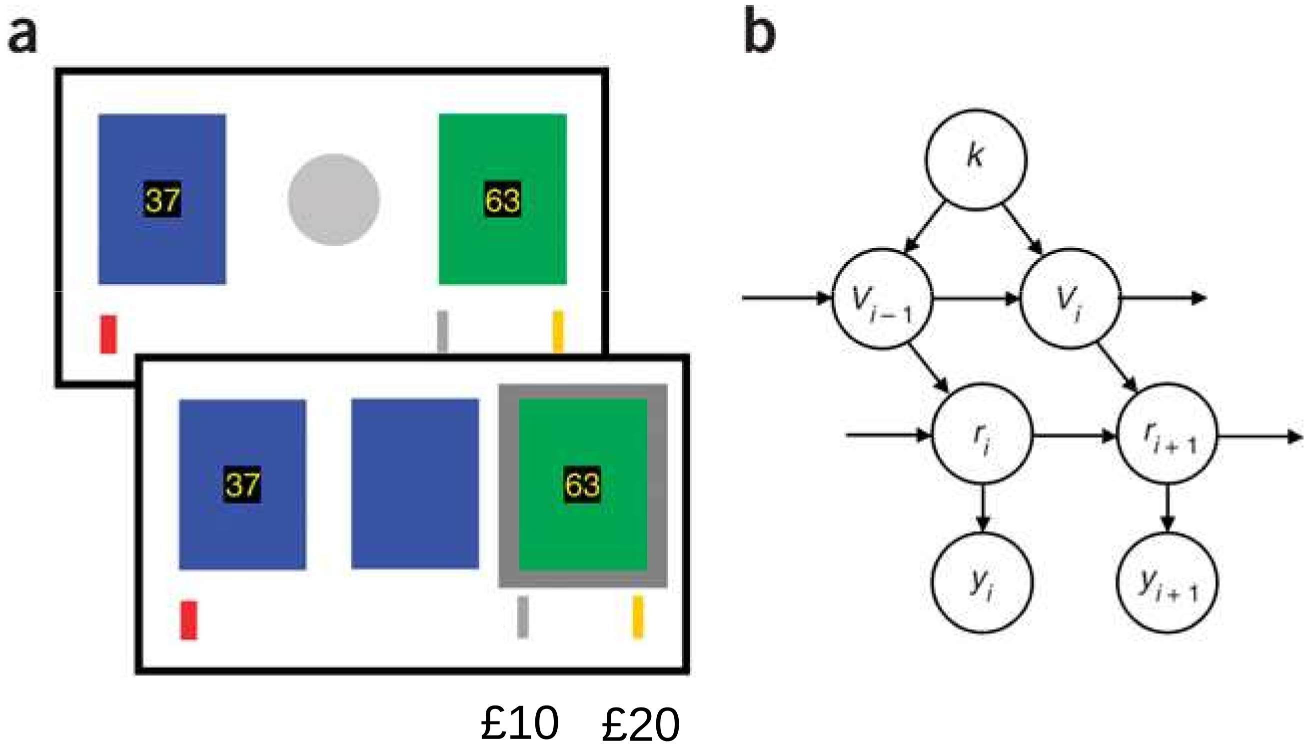


ACh level determines a ***threshold*** for **NE**-mediated context change:

$$\boxed{\mathrm{NE} > \frac{\mathrm{ACh}}{.5 + \mathrm{ACh}}}$$

high expected uncertainty makes a high bar for unexpected uncertainty

Behrens et al

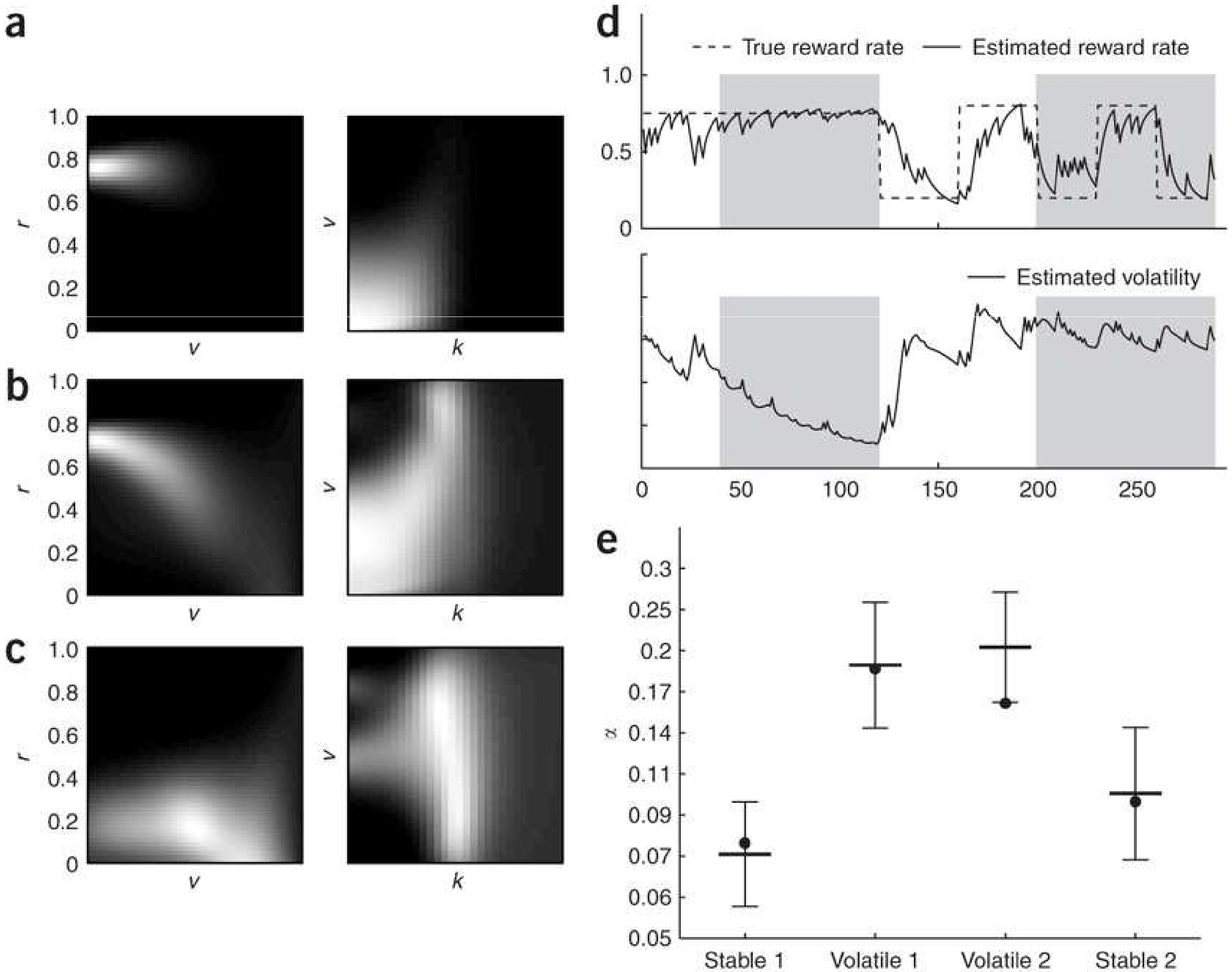


Behrens et al

stable 120

change 15

stable 25

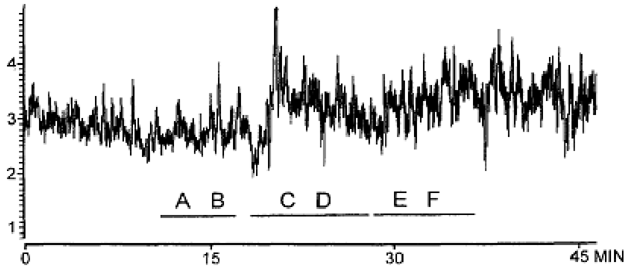


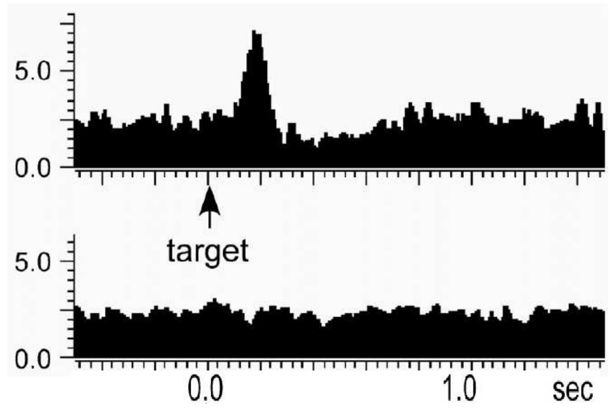
Summary

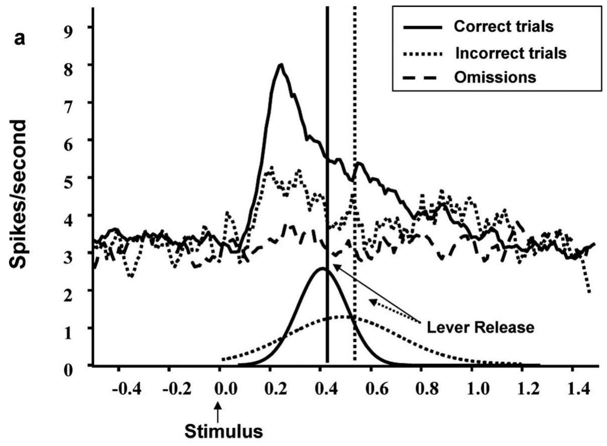
- Single framework for understanding ACh, NE and some aspects of attention
- ACh/NE as expected/unexpected uncertainty signals
- Experimental psychopharmacological data replicated by model simulations
- Implications from complex interactions between ACh & NE
- Predictions at the cellular, systems, and behavioral levels
- Consider loss functions
- Activity vs weight vs neuromodulatory vs population representations of uncertainty (ACC in Behrens)

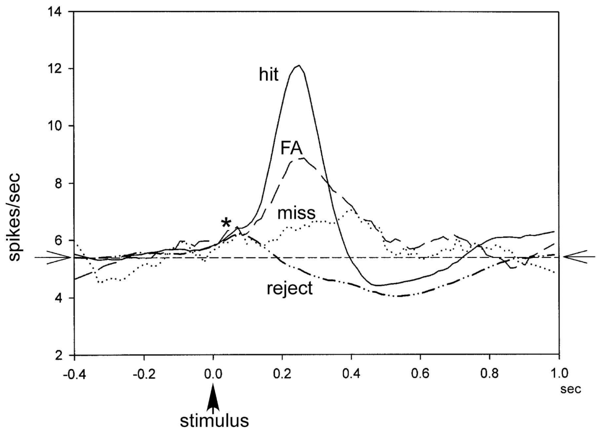
Aston-Jones: Target Detection

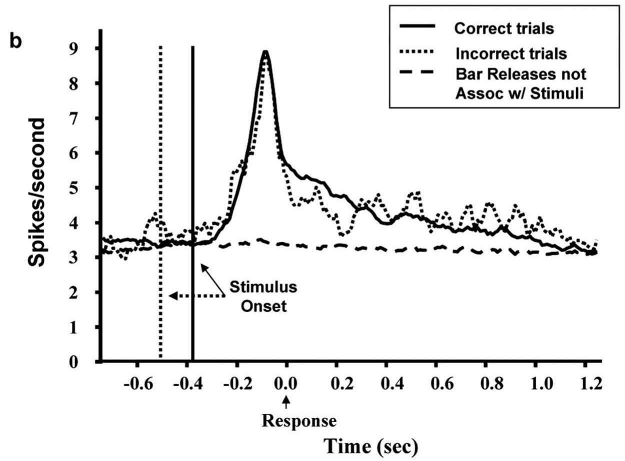
detect and react to a rare target amongst common distractors









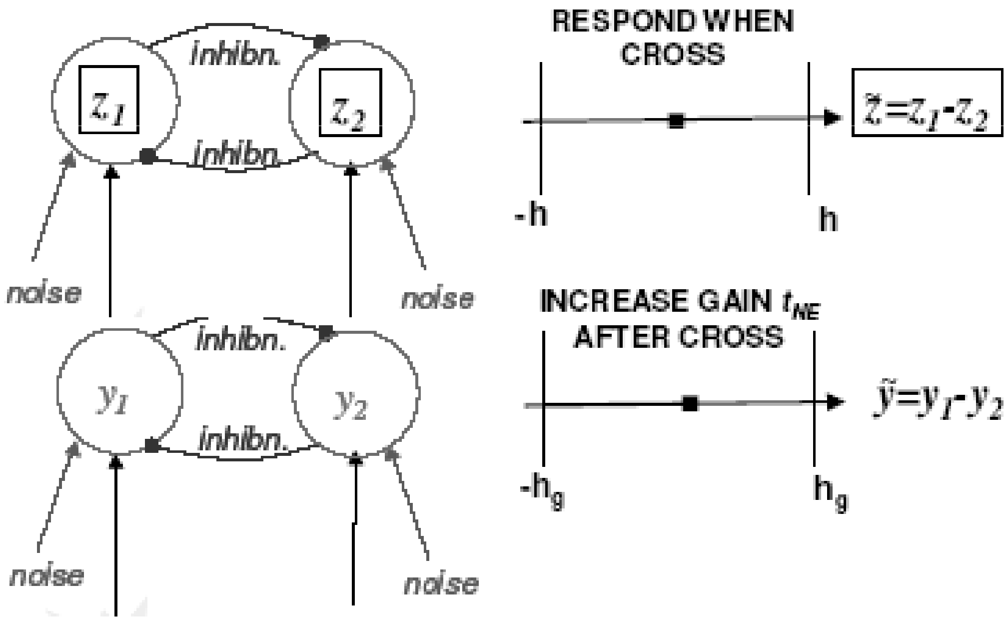


- elevated tonic activity for reversal
- activated by rare target (and reverses)
- not reward/stimulus related? more response related?
- no reason to persist as hardly unexpected

Clayton, *et al*

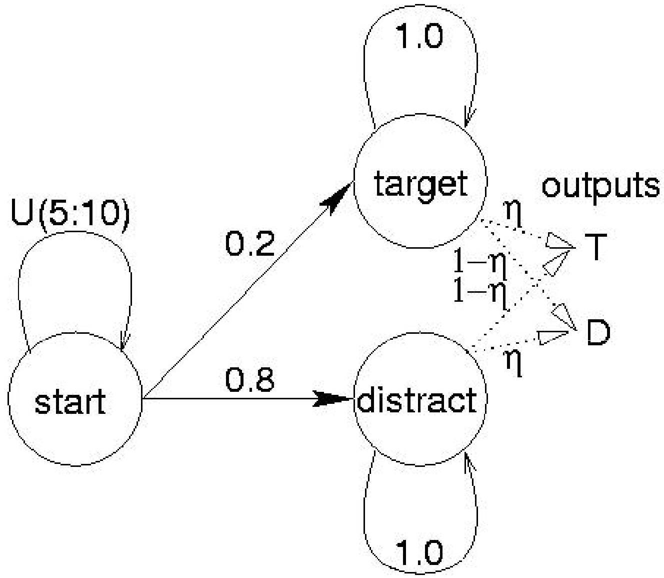
Phasic NE activity

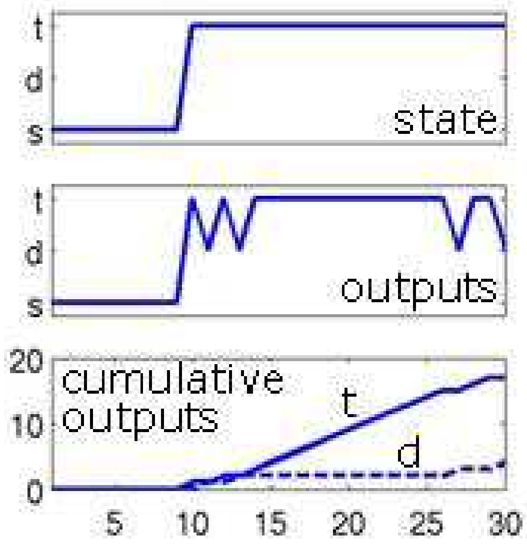
- no reason to persist under our tonic model
- quantitative phasic theory (Brown, Cohen, Aston-Jones): gain change
  - NE controls balance of recurrence/bottom-up
  - implements changed S/N ratio with target
  - or perhaps decision (through instability)
  - detect to detect
  - why only for targets?
  - already detected (early bump)

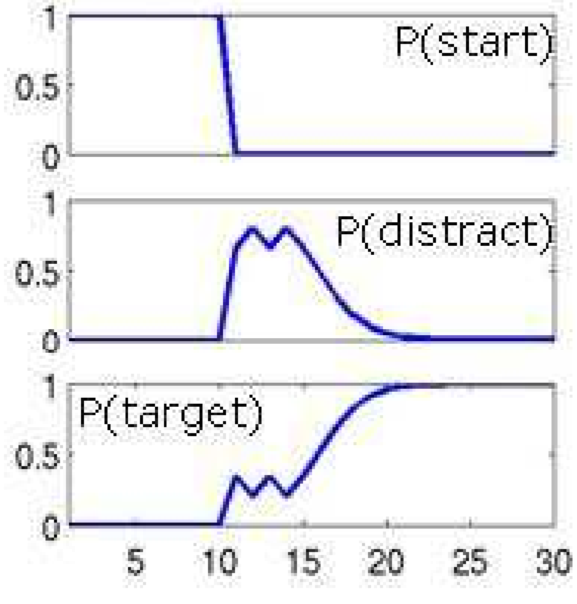


- NE reports unexpected state changes within the task

Vigilance Model







- variable time in start
- η controls confusability

- one single run
- cumulative is clearer

- exact inference
- effect of 80% prior

Phasic NE

- NE reports uncertainty about current state
  - state in the model, not state of the model
  - divisively related to prior probability of that state

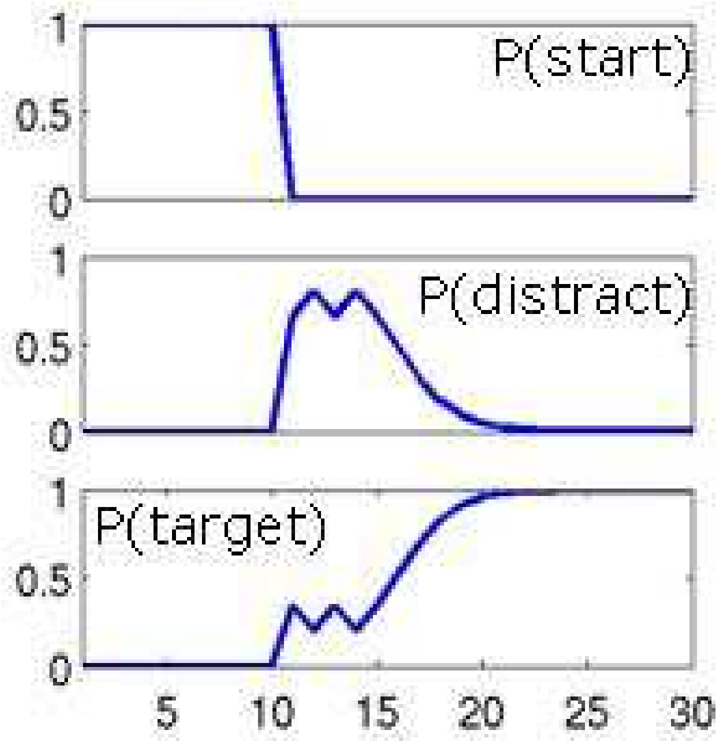
- NE measured relative to default state sequence

  start → distractor

- temporal aspect - start → distractor

- structural aspect target *versus* distractor

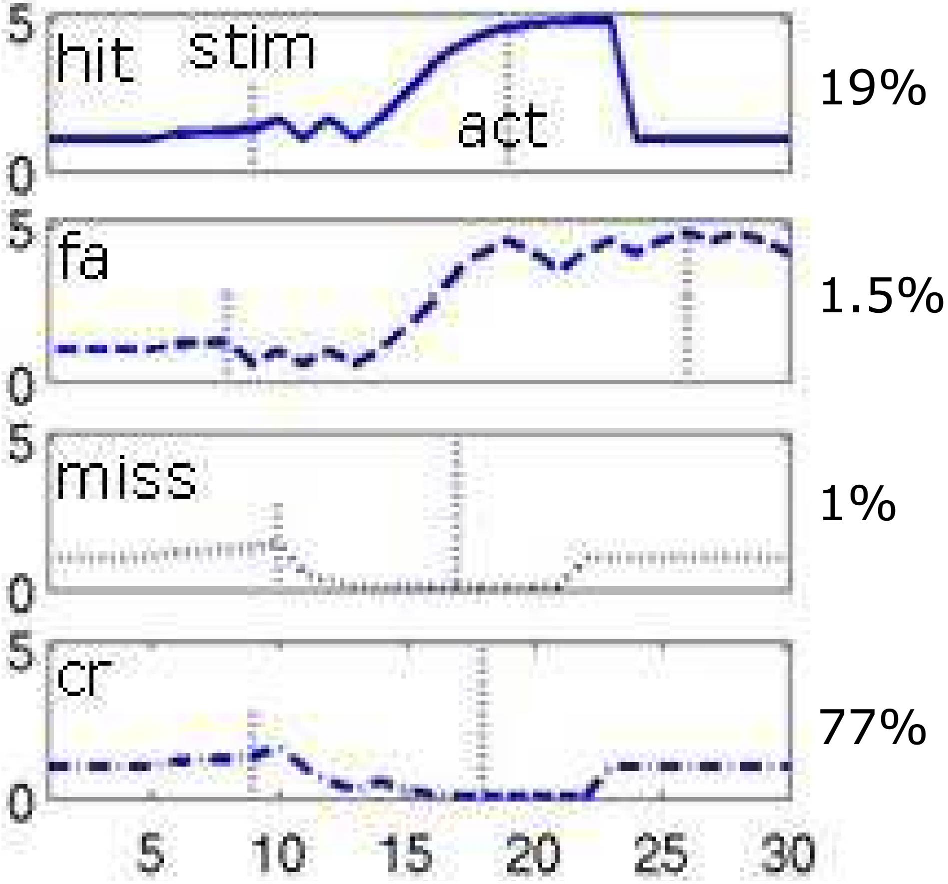
Phasic NE

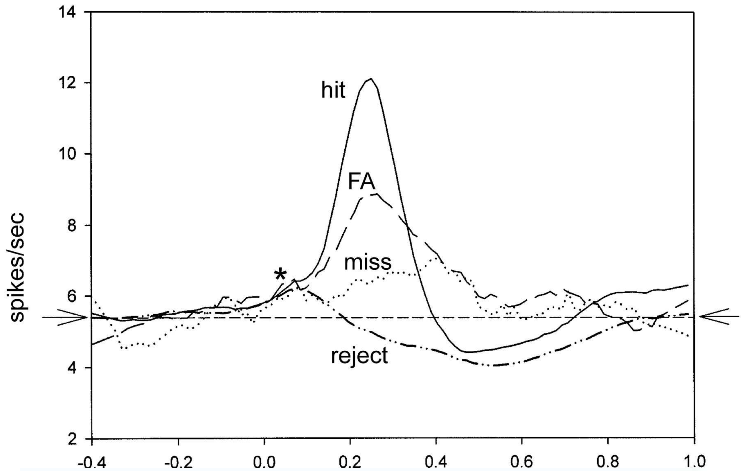


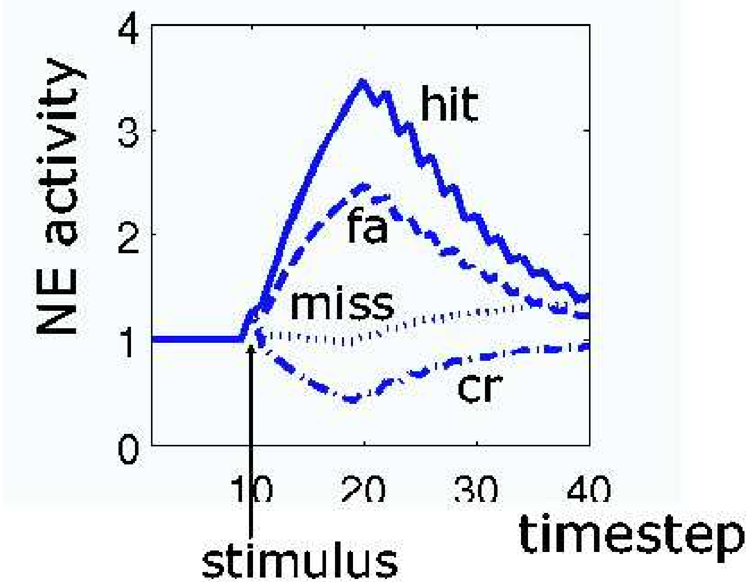
- onset response from timing uncertainty (SET)
- growth as P(target)/0.2 rises
- act when P(target)=0.95
- stop if P(target)=0.01
- arbitrarily set NE=0 after 5 timesteps

(small prob of reflexive action)

Four Types of Trial

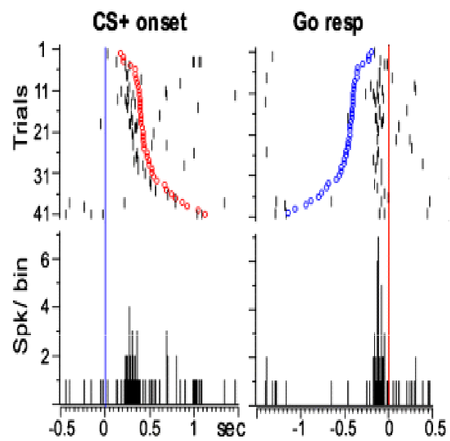


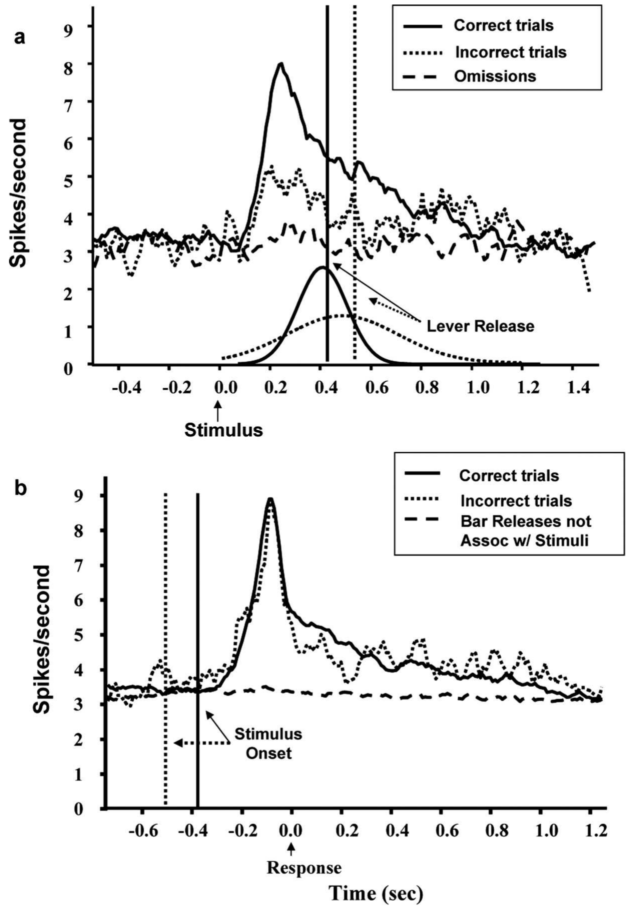


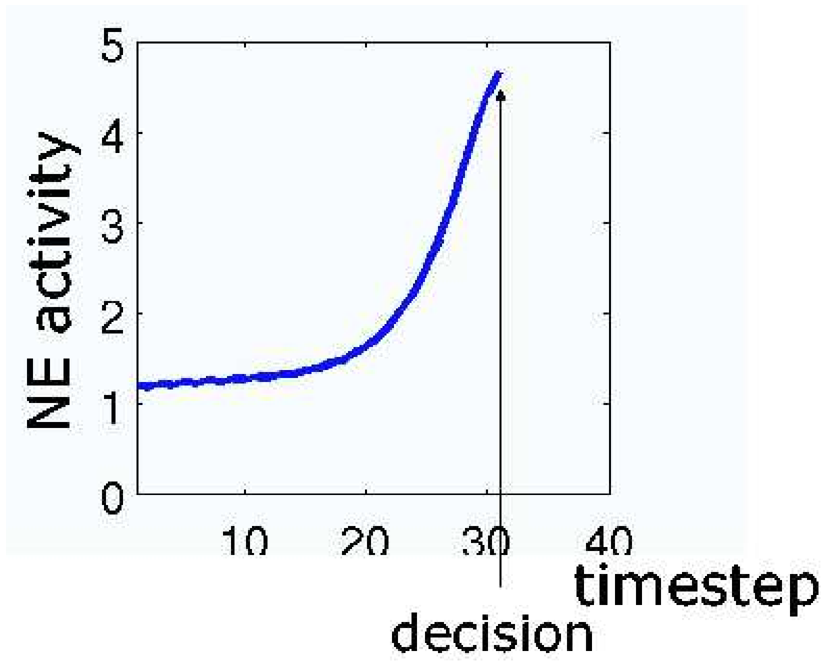


fall is rather arbitrary

Response Locking

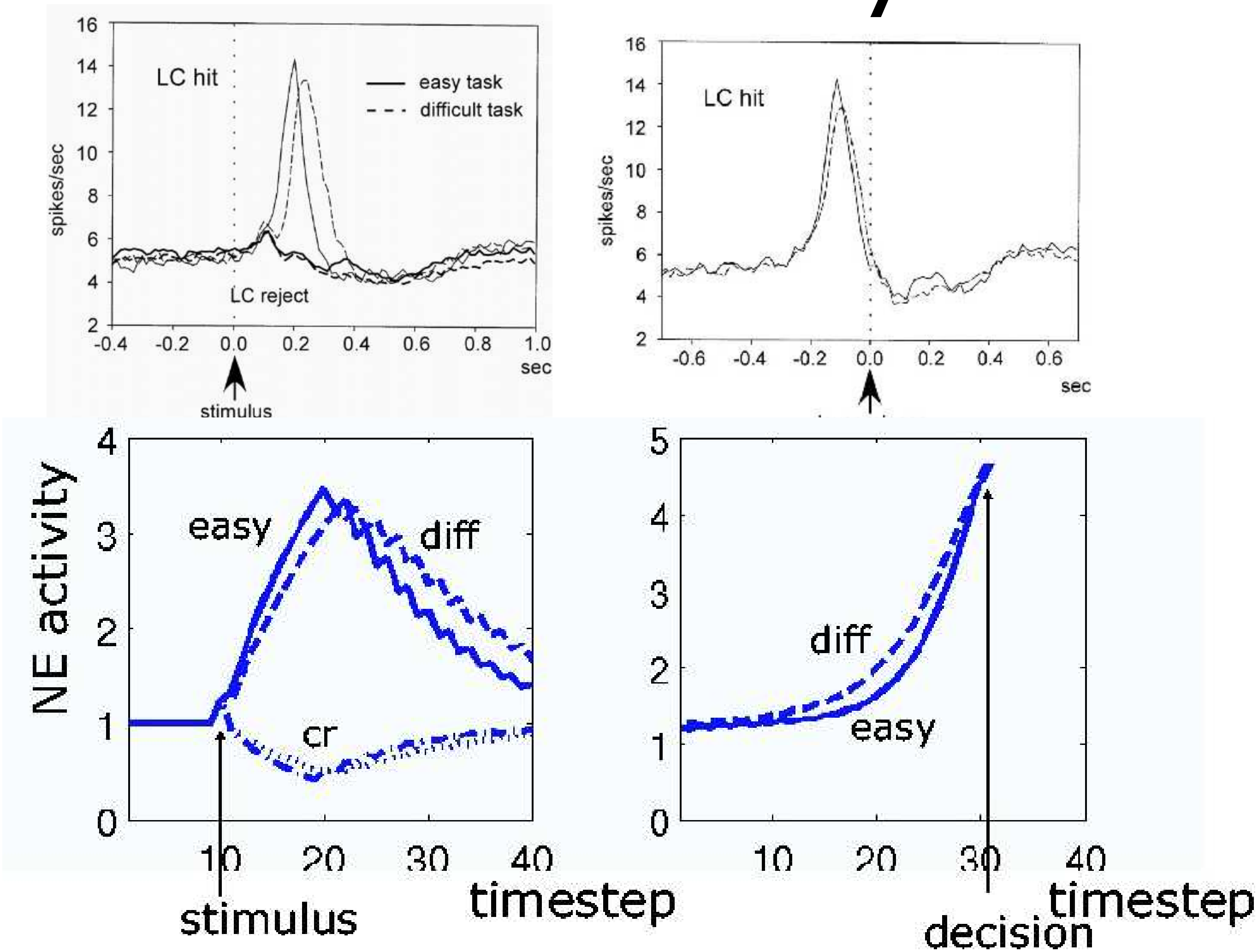






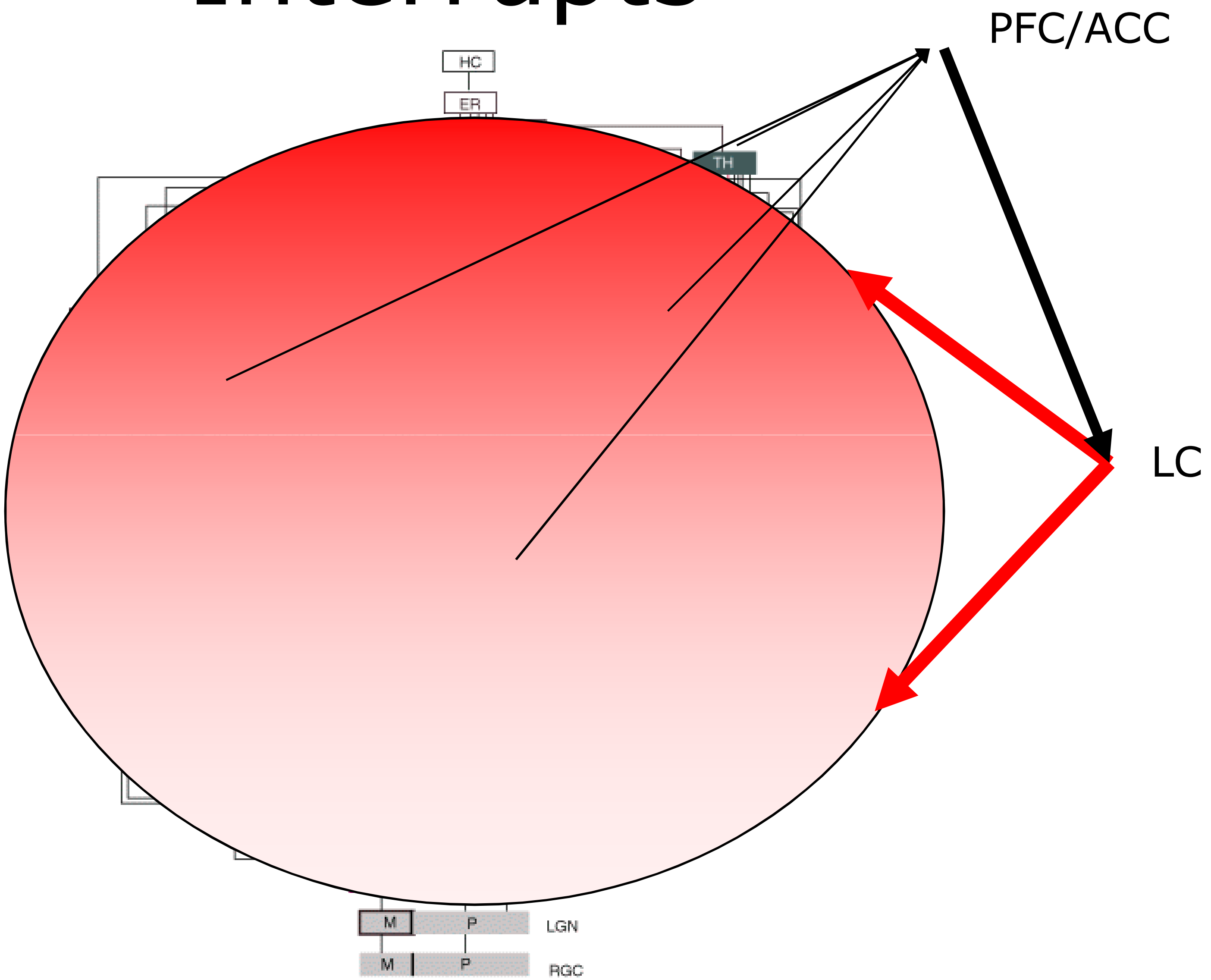
slightly flatters the model – since no further response variability

Task Difficulty



- set η=0.65 rather than 0.675
- information accumulates over a longer period
- hits more affected than cr's
- timing not quite right

Interrupts

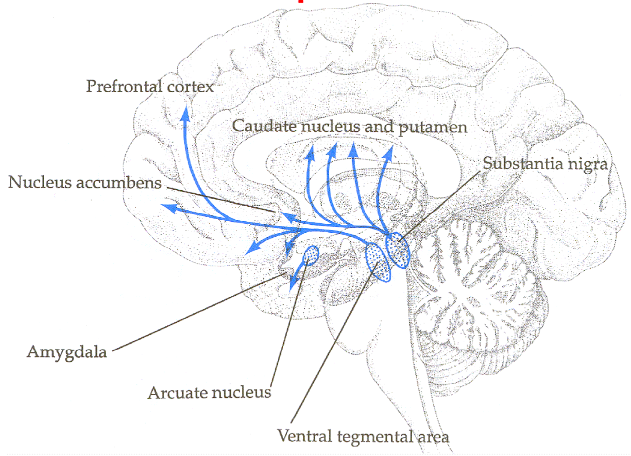


Discusssion

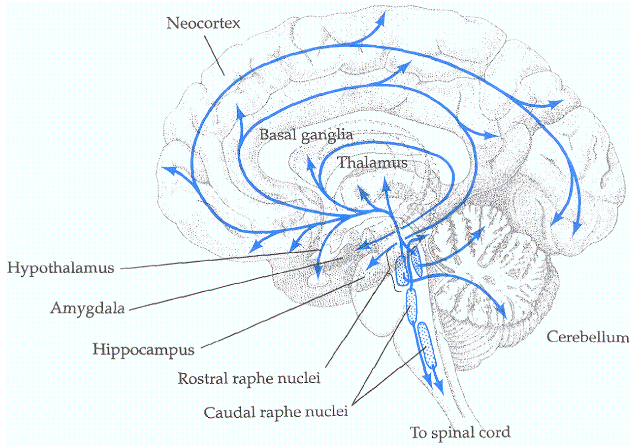
- phasic NE as unexpected state change *within* a model
- relative to prior probability; against default
- interrupts ongoing processing
- tie to ADHD?
- close to alerting – but not necessarily tied to behavioral output (onset rise)
- close to behavioural switching – but not DA
- phasic ACh: aspects of known variability within a state?

Computational Neuromodulation

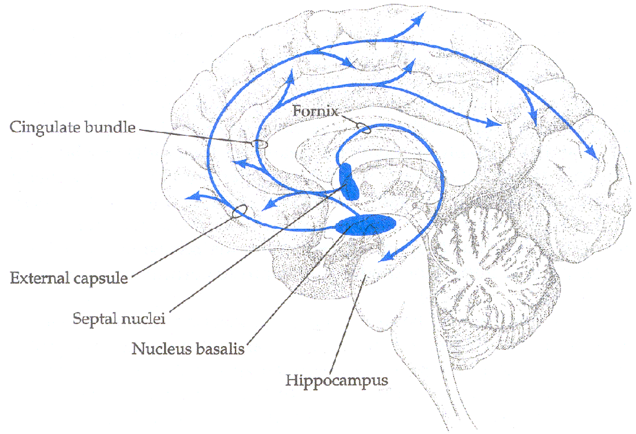
dopamine



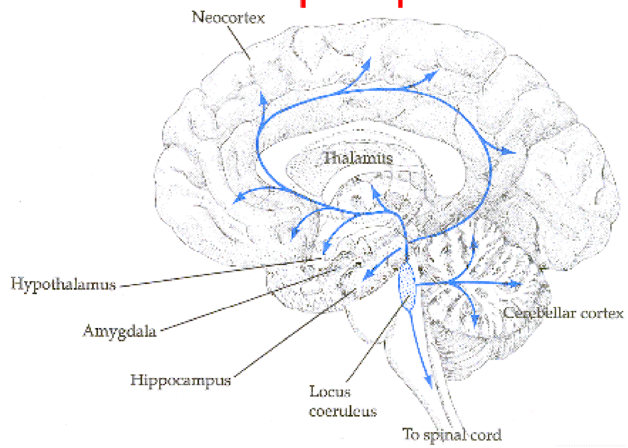
5HT



acetylcholine



norepinephrine



- general: excitability, signal/noise ratios
- specific: prediction errors, uncertainty signals

Computational Neuromodulation

Δ weight α (learning rate) x (error) x (stimulus)

- precise, falsifiable, roles for DA/5HT; NE/ACh
- only part of the story:
  - 5HT: median raphe
  - ACh: TANs, septum, *etc*
  - huge diversity of receptors; regional specificity
- psychological disagreement about many facets:
  - attention: over-extended
  - reward: reinforcement, liking, wanting, *etc*
- interesting role for imaging:
  - it didn't have to be that simple!