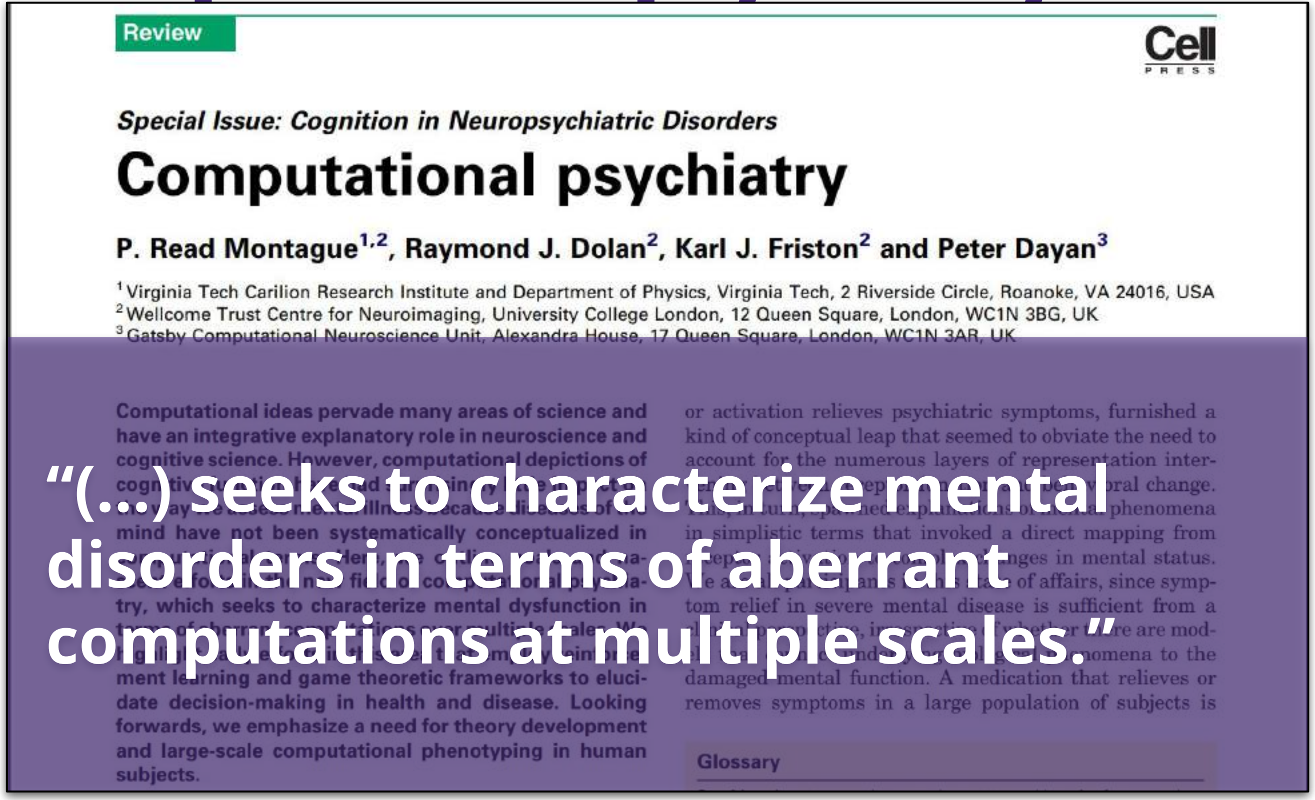
ACT-R For Computational Psychiatry: Predicting Recovery Curves for PTSD

Andrea Stocco, Briana Smith, & Lori Zoellner
University of Washington, Seattle

Computational psychiatry



"(...) seeks to characterize mental disorders in terms of aberrant computations at multiple scales."

Gold Standard: Reinforcement Learning

Redish, 2004, *Science*

Addiction as a Computational Process Gone Awry

A. David Redish

Addictive drugs have been hypothesized to access the same neurophysiological mechanisms as natural learning systems. These natural learning systems can be modeled through temporal-difference reinforcement learning (TDRL), which requires a reward-error signal that has been hypothesized to be carried by dopamine. TDRL learns to predict reward by driving that reward-error signal to zero. By adding a noncompensable drug-induced dopamine increase to a TDRL model, a computational model of addiction is constructed that over-selects actions leading to drug receipt. The model provides an explanation for important aspects of the addiction literature and provides a theoretic viewpoint with which to address other aspects.

If addiction accesses the same neurophysiological mechanisms used by normal reinforcement-learning systems (*1–3*), then it should be possible to construct a computational model based on current reinforcement-learning theories (*4–7*) that inappropriately selects an "addictive" stimulus. In this paper, I present a computational model of the behavioral consequences of one effect of drugs of abuse, which is increasing phasic dopamine levels through neuropharmacological means. Many drugs of abuse increase dopamine levels either directly [e.g., cocaine (*8*)] or indirectly [e.g., nicotine (*9*, *10*) and heroin (*11*)]. A neuropharmacologically driven increase in dopamine is not the sole effect of these drugs, nor is it likely to be the sole reason that drugs of abuse are addictive. However, this model provides an immediate explanation for several important aspects of the addiction literature, including the sensitivity of the probability of selection of drug receipt to prior drug experience, to the size of the contrasting nondrug reward, and the sensitivity but inelasticity of drugs of abuse to cost.

The proposed model has its basis in in order to accommodate the learning algorithm (*6*, *7*); however, animals (including humans) show hyperbolic discounting of future rewards (*12*, *13*). This will be addressed by including multiple discounting time scales within the model (*14*).

In temporal-difference reinforcement learning (TDRL), an agent (the subject) traverses a world consisting of a limited number of explicit states. The state of the world can change because of the action of the agent or as a process inherent in the world (i.e., external to the agent). For example, a model of delay conditioning may include an interstimulus-interval state (indicated to the agent by the observation of an ongoing tone); after a set dwell time within that state, the world transitions to a reward state and delivers a reward to the agent. This is an example of changing state because of processes external to the agent. In contrast, in a model of FR1 conditioning, an agent may be in an action-available state (indicated by the observation of a lever available to the agent), and the world will remain in the action-available state until the agent takes the action (of pushing the lever), pected and observed changes in value (*6*). This signal, termed δ, can be used to learn sequences that maximize the amount of reward received over time (*6*). δ is not equivalent to pleasure; instead, it is an internal signal indicative of the discrepancy between expectations and observations (*5*, *7*, *15*). Essentially, if the change in value or the achieved reward was better than expected ($\delta > 0$), then one should increase the value of the state that led to it. If it was no different from expected ($\delta = 0$), than the situation is well learned and nothing needs to be changed. Because δ transfers backward from reward states to anticipatory states with learning, actions can be chained together to learn sequences (*6*). This is the heart of the TDRL algorithm (*4–7*).

TDRL learns the value function by calculating two equations as the agent takes each action. If the agent leaves state S_t and enters state S_t at time t, at which time it receives reward $R(S_t)$, then

$$\delta(t) = \gamma^d[R(S_t) + V(S_t)] - V(S_k) \quad (2)$$

where γ^d indicates raising the discounting factor γ by the delay d spent by the animal in state S_k (*14*). $V(S_k)$ is then updated as

$$V(S_k) \leftarrow V(S_k) + \eta_v\delta \quad (3)$$

where η_v is a learning rate parameter.

Phasic increases in dopamine are seen after unexpected natural rewards (*16*); however, with learning, these phasic increases shift from the time of reward delivery to cuing stimuli (*16*). Transient increases in dopamine are now thought to signal changes in the expected future reward (i.e., unexpected changes in value) (*4*, *16*). These increases can occur either with unexpected reward or with unexpected cue stimuli known to signal reward (*16*) and have been hypothesized to signal δ (*4*, *7*, *16*). Models of dopamine signaling as δ have been found to be compatible with many aspects of the data (*4*, *5*, *16*, *17*).

Maia & Frank, 2016, *Nat Neurosci*

nature neuroscience

From reinforcement learning models to psychiatric and neurological disorders

Tiago V Maia[1,2] & Michael J Frank[3,4]

Over the last decade and a half, reinforcement learning models have fostered an increasingly sophisticated understanding of the functions of dopamine and cortico-basal ganglia-thalamo-cortical (CBGTC) circuits. More recently, these models, and the insights that they afford, have started to be used to understand important aspects of several psychiatric and neurological disorders that involve disturbances of the dopaminergic system and CBGTC circuits. We review this approach and its existing and potential applications to Parkinson's disease, Tourette's syndrome, attention-deficit/hyperactivity disorder, addiction, schizophrenia and preclinical animal models used to screen new antipsychotic drugs. The approach's proven explanatory and predictive power bodes well for the continued growth of computational psychiatry and computational neurology.

The limitations of the state-of-the-art in nosology in psychiatry have been much debated in the context of the development of the new edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*). There is widespread agreement that the current symptom-based system of classification must eventually be replaced with a system based on pathophysiology[1]. However, the current understanding of the neurobiology and genetics of psychiatric disorders remains too limited to form the backbone of nosology[2]. This limited understanding is also reflected in the state-of-the-art in treatment, with most psychiatric medications having been found by serendipity, rather than through rational design. Neurology typically deals with disorders with better understood etiology (for example, loss of dopaminergic neurons in Parkinson's disease), but even then it is often unclear how these etiological processes produce complex patterns of symptoms and why treatments can alleviate some deficits while exacerbating, or even causing, others[3,4]. Part of the problem is the complexity of the brain and mind and the many levels of analysis that span the two. Computational models are a valuable tool for taming

interactions; habits, goal-directed actions and their interactions; and the inter-related issues of incentive salience, motivation and vigor[5–9].

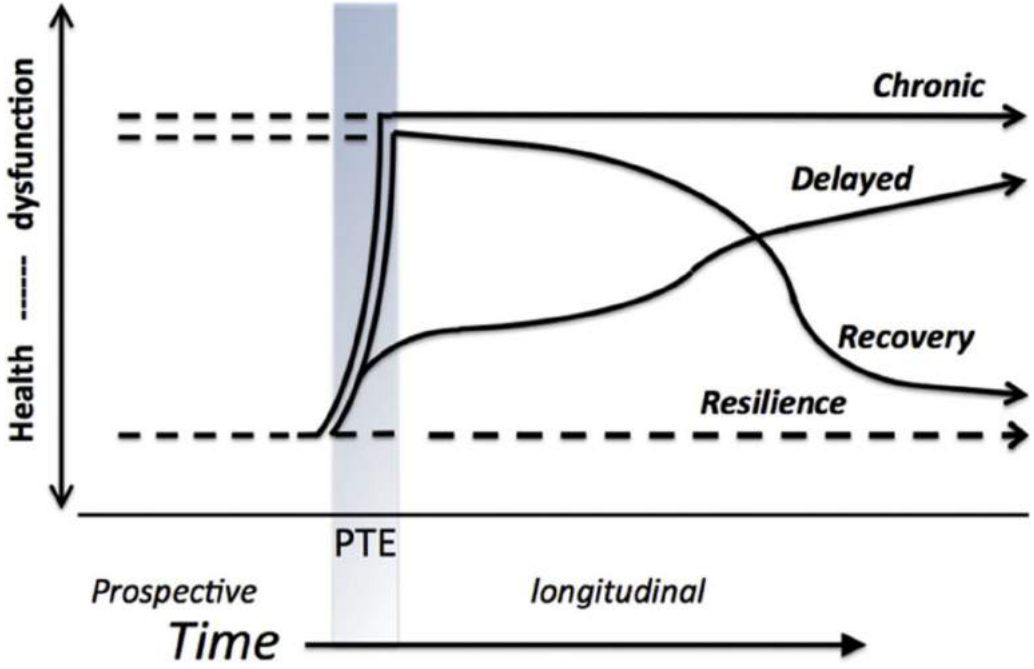
Organizing behavior in ways that obtain outcomes appropriate for the current motivational state (for example, acquiring food if hungry) and that avoid harmful outcomes is crucial for survival and is therefore a central organizing principle of the nervous system. Not surprisingly, then, disturbances of the dopaminergic system and CBGTC circuits have a key role in several psychiatric and neurological disorders. Reinforcement learning models have recently started to be applied to these disorders and have been shown to have substantial explanatory and predictive power[10–14]. The approach builds on an understanding of the computations that these circuits perform in healthy individuals and investigates how pathophysiological processes alter these computations, producing symptoms. We therefore start by reviewing the computational neurobiology of the normal functioning of these circuits. We then discuss several disorders that have benefited or are ripe to benefit from the use of reinforcement learning models. We

Can ACT-R contribute the missing piece?

Intrusive Memories

- Emotionally charged memories that intrude in everyday life and prevent normal functioning
- Staple of several disorders (depression, OCD, …)
- Particularly studied in PTSD

The problem: Not all patients are equal



Galatzer-Levy , Huangb, & Bonanno, 2018, *Clinical Psych Review*

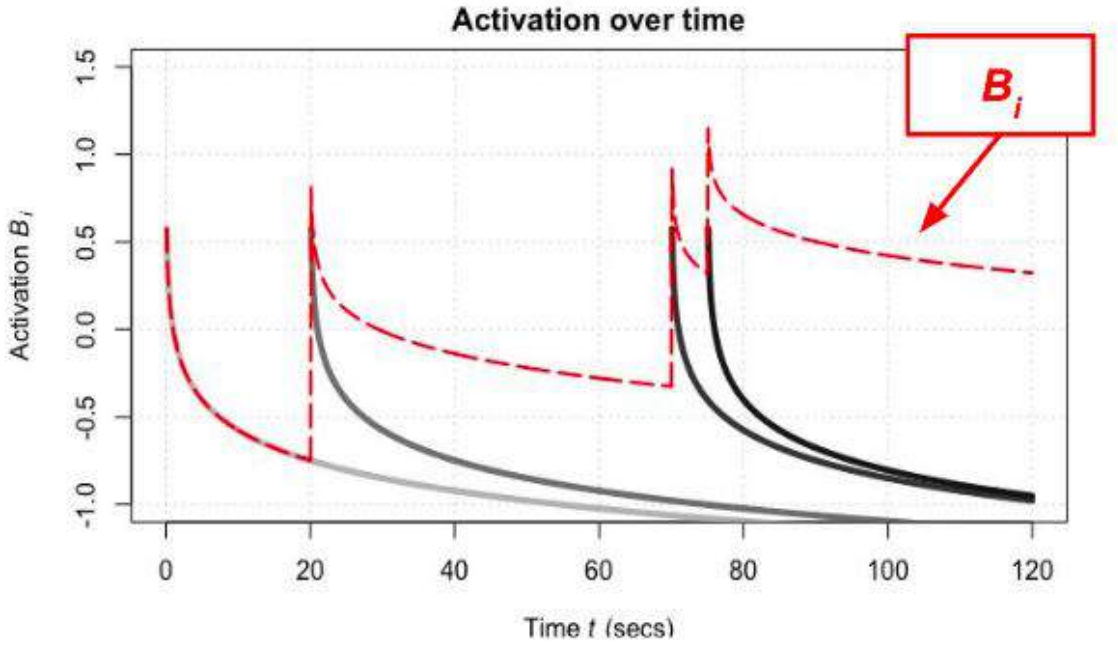
ACT-R has an excellent model of memory!

$$\frac{p(i|Q)}{p(\neg i|Q)} = \frac{p(i)}{p(\neg i)} \times \prod_{q \in Q} \frac{p(q|i)}{p(q)}$$

$$\log\left(\frac{p(i|Q)}{p(\neg i|Q)}\right) = \log\left(\frac{p(i)}{p(\neg i)}\right) + \log\left(\prod_{q \in Q} \frac{p(q|i)}{p(q)}\right)$$

Total Activation		***Base-Level Activation***		***Spreading Activation***
A_i	=	B_i $\sum_i \log(t_i^{-d})$	+	S_i $\sum_i WS_{ji}$

Memory is rational and elastic



Incorporating emotions and memory

- Activation tracks **need probability**, $A_i \sim P(i \mid Q)$
  - Assumes all memories are **created equal**.

Not all memories are equal!

(One-shot learning
at its finest)

Incorporating emotions and memory

- Activation tracks **need probability**, $A_i \approx P(i \mid Q)$
  - Assumes all memories are **created equal**.
  - Some memories are **intrinsically more important**!
- A_i should reflect a memory's **expected value,** i.e.

$$A_i \approx P(i \mid Q) * V(i)$$

- Old idea (e.g., West, Larue, and my 2004-self)

Implementation

- In addition to slots, chunks are given a scalar quantity, *V*
  - Insight: Emotions have “survival value” (Panksepp)
- *V* is actually normalized, given the rest of memories, so that *M*(*V*) = 1
- Code at **`github.com/UWCCDL/PTSD`**

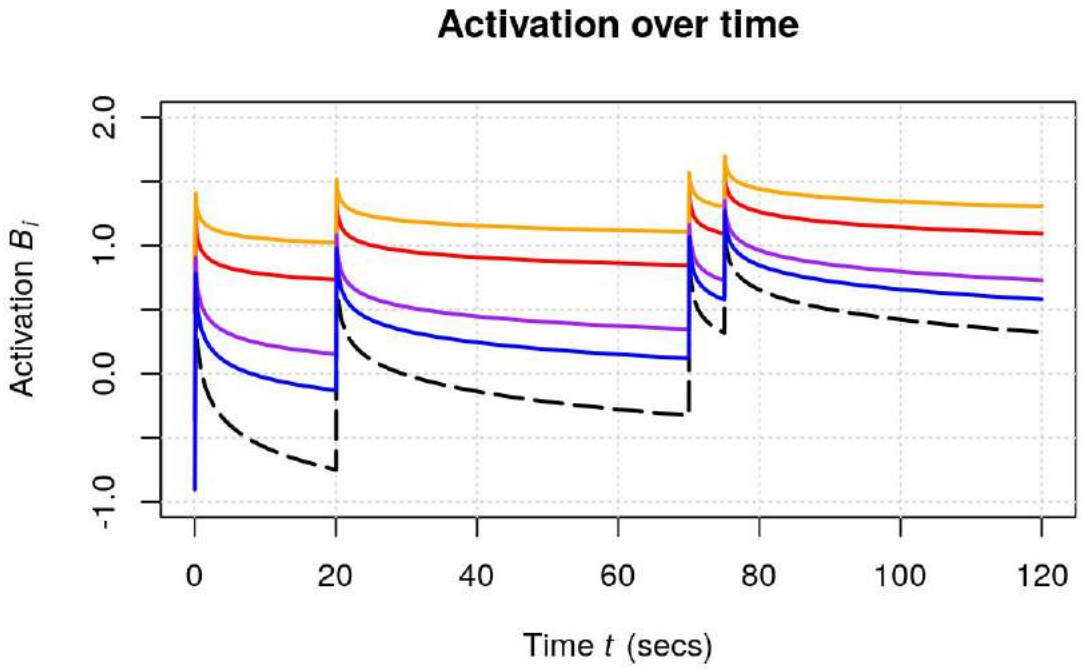
Implementation

$$\frac{p(i|Q)}{p(\neg i|Q)} = \frac{p(i)}{p(\neg i)} \times \prod_{q \in Q} \frac{p(q|i)}{p(q)} \times V(i)$$

$$\log\left(\frac{p(i|Q)}{p(\neg i|Q)}\right) = \log\left(\frac{p(i)}{p(\neg i)}\right) + \log\left(\prod_{q \in Q} \frac{p(q|i)}{p(q)}\right) + \log\left(V(i)\right)$$

Total Activation		***Base-Level Activation***		***Spreading Activation***		***Survival Value***
A_i	=	B_i $\sum_i t_i^{-d}$	+	S_i $\sum_i WS_{ji}$	+	V_i $log(V_i)$

This model is still rational, but memories are inelastic



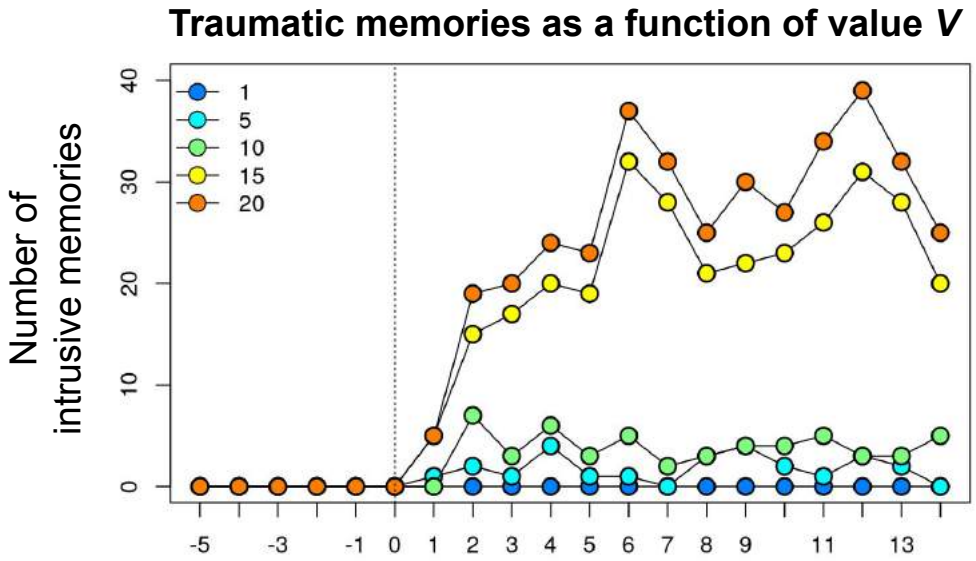
Testing the model

- Abstract agent that simulates a few months of life
- Accrues random memories at fixed intervals
- Productions implement a **perceive/retrieve/act** loop
  - "Retrieve" replaces "Decide"
  - Not too different from Christian's IBL models

Testing the model

- At predefined time T_{PTE}, a traumatic event *PTE* is introduced
  - Random memory, but $V > 1$
- We can now track down how often PTE is retrieved out of context

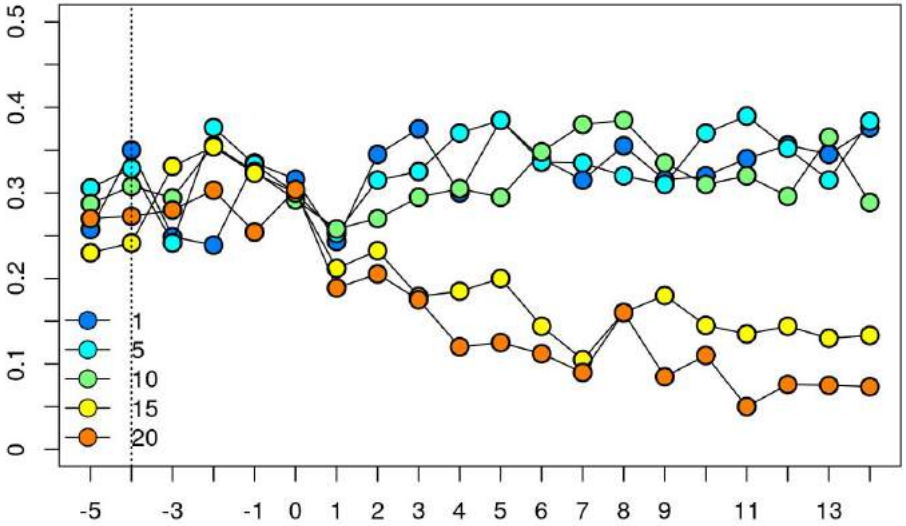
Effects of PTE Value



Non-linear effect of trauma (chronic vs. resilience)

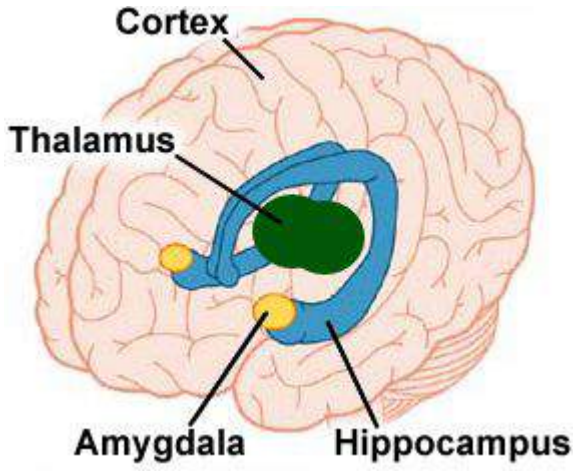
Disruption of cognitive functions

- Function: relevance of a retrieved memory *R* to the current situation *Q*
- Measured as cos(*R*, *Q*)
- Traumatic memories more disruptive

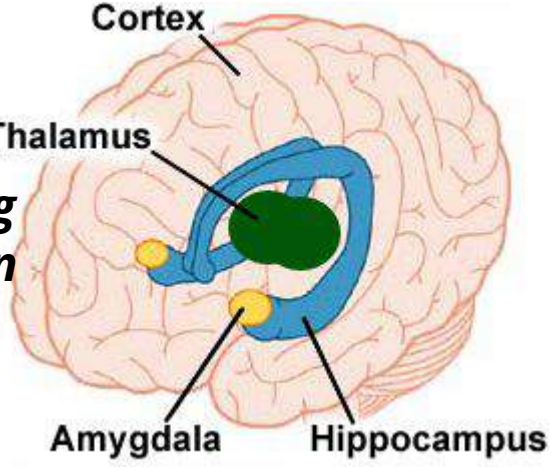


Biological interpretation

- The PTSD field is looking for **biomarkers**
- Two targets:
  - Amygdala/hippocampus circuit
  - Prefrontal/hippocampus circuit



Biological interpretation and findings



Spreading Activation

S_i

$\sum_j W S_{ji}$

Survival Value

V_i

$log(V_i)$

Base-Level Activation

B_i

$\sum_j t_j^{-d}$

- **Greater amygdala** = greater effect of trauma (greater value *V*)
- **Greater hippocampus** = smaller effect of trauma (smaller decay *d*)
- **Greater prefrontal volume** = smaller effect of trauma (greater spreading activation *W*)
- **Greater working memory** = smaller effect of trauma (again, greater *W*)

What's next

- More simulations are running!
- Predicting PTSD from brain activity (should provide *V* and *d* through neurometrics)
  - ACT-R as the missing model between biomarkers and outcomes
  - With Katie McLaughlin
- Application: Optimizing presentation of traumatic images to human raters

Thank you!

Code at: `http://github.com/UWCCDL/PTSD`