Unravelling the Fear Circuit: A Novel Computational Model of Fear Acquisition, Extinction & Recovery

Shreya Rajagopal, & Thad Polk, University of Michigan

Introduction

Background

•Fear responses to a conditioned stimulus typically decline if the stimulus is presented without an aversive stimulus (**fear extinction**). Fear extinction forms the basis of common treatments for anxiety disorders like **exposure therapy.**

•**Neuroscientific insights from studying fear extinction in animals** hold promise for understanding why fear sometimes returns after it is extinguished, but the invasive nature of these experiments makes it hard to generalize these insights to humans and translate them to clinical practice.

•We bridge this gap by introducing a novel computational model of fear acquisition, extinction, and recovery by integrating key findings from the fear circuit in the animal brain to simulate return-of-fear phenomena – specifically fear **Renewal** (*immediate return of fear in a non-extinction context*)[1,2] and **Spontaneous Recovery** (*Delayed return of fear in the extinction context*)[3].

•The model also generates interpretable and **testable predictions**. For example, it predicts that repeatedly pairing a fear-conditioned stimulus with a positive stimulus (**Counterconditioning**) – instead of no stimulus in extinction – will reduce the return of fear.

Incorporating Key Neuroscientific Insights

• **Distinct Pathways for Processing Cue and Context Information:** Cue information directly travels to the amygdala while context information takes a more involved route[4]. We update cue-amygdala connection weights more rapidly than context-amygdala weights.

•**Competition between Positive and Negative Stimulus Responsive Neurons in the Basolateral Amygdala (BLA) to decide the Fear Level Associated with a Cue-Context Pair:** If the net activation of the negative BLA population is larger than the net positive, the response is of negative valence (fear) and vice-versa.

•**Extinction Engram is formed in the positive BLA population**, by recruiting a random subset of neurons to neutralize the fear response[5].

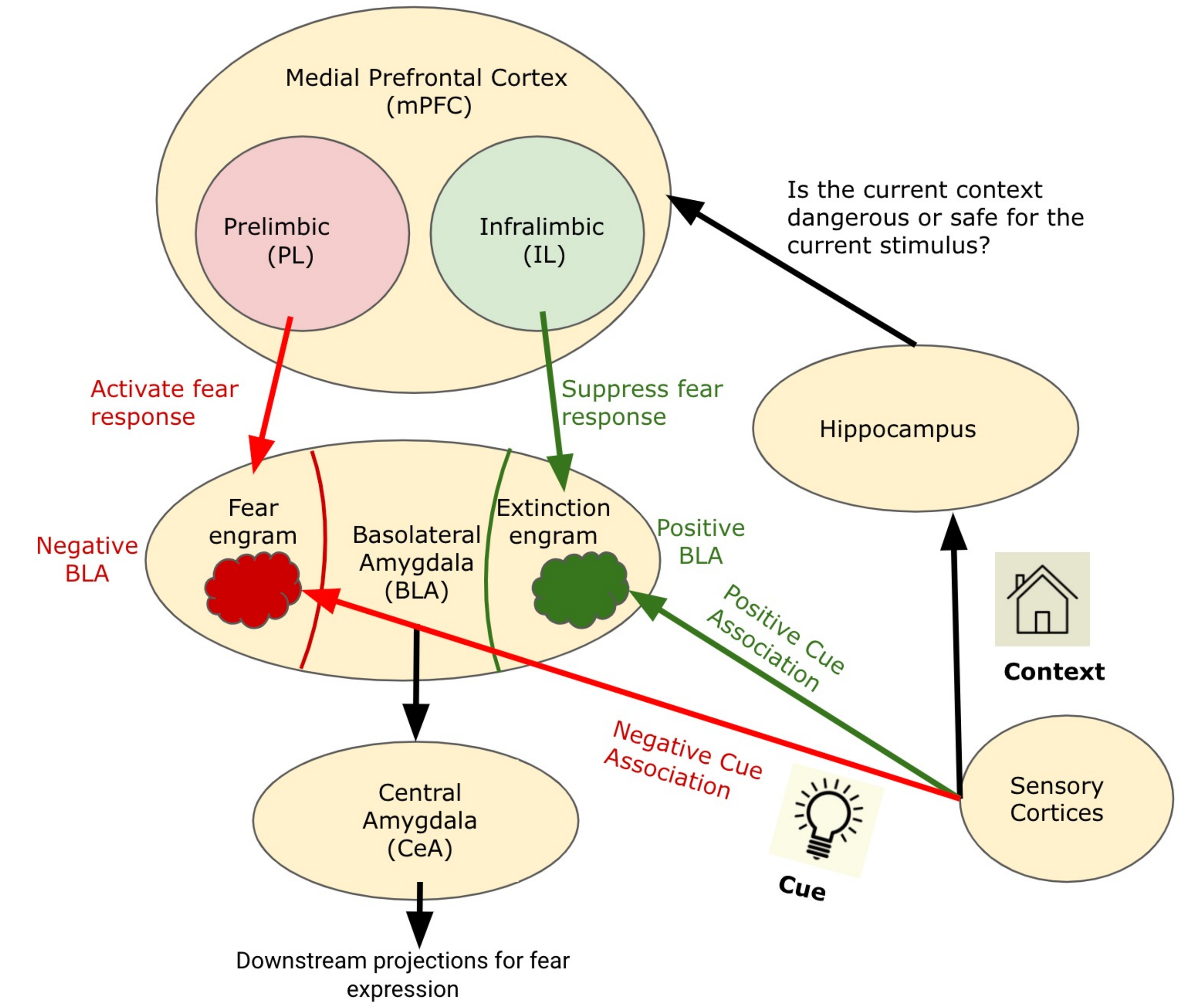
•**Decay of Context Associations over Time**: Context-Amygdala associations deteriorate with time faster than the cue-amygdala associations.

•**Prediction error incorporated in computing the final fear response**: A larger difference in net BLA activation from one trial to the next leads to a larger change in fear responses.

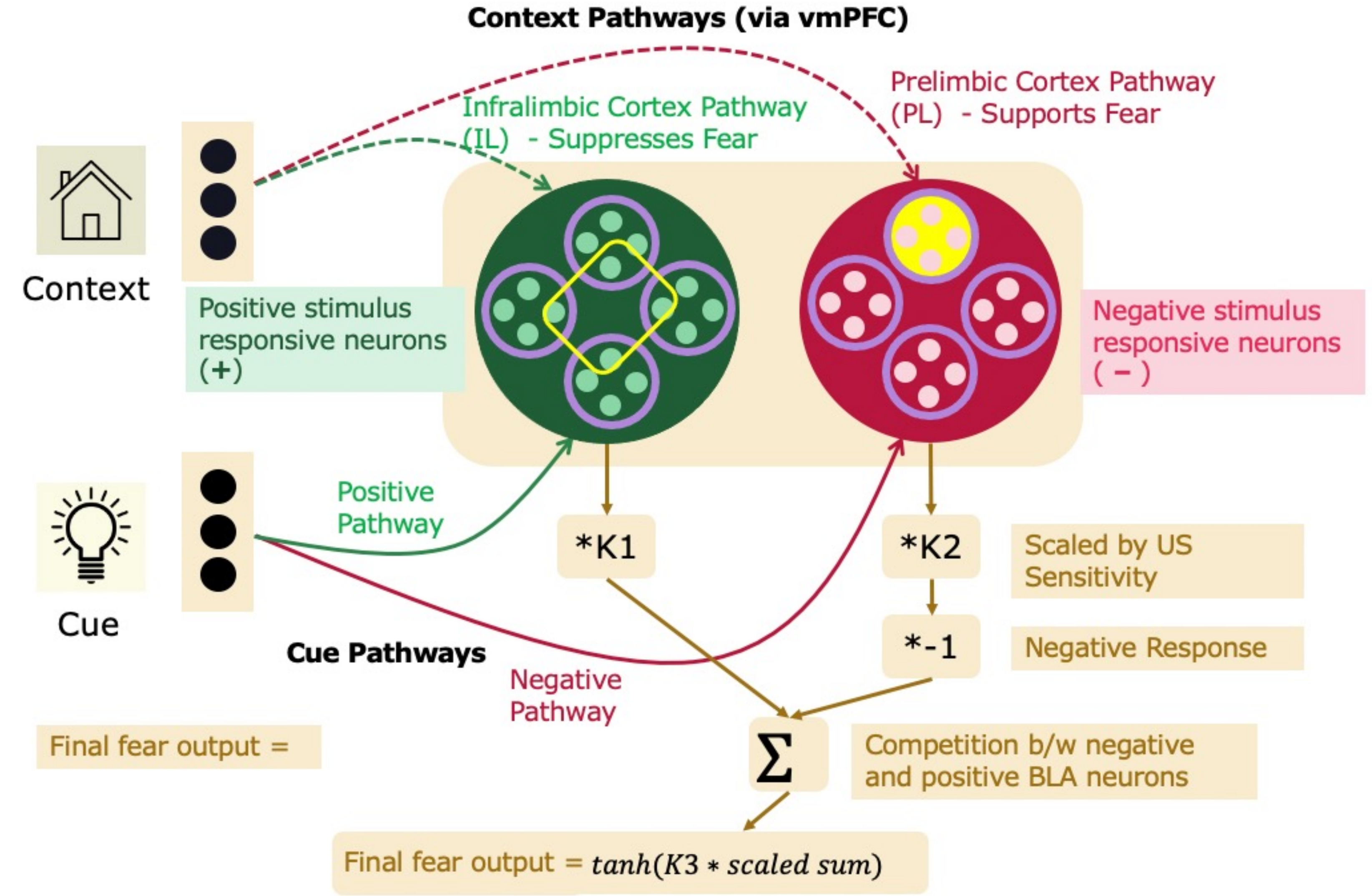
•The model also incorporates different **Reward Sensitivities** to different stimuli.

Results

I. The Circuit of Fear

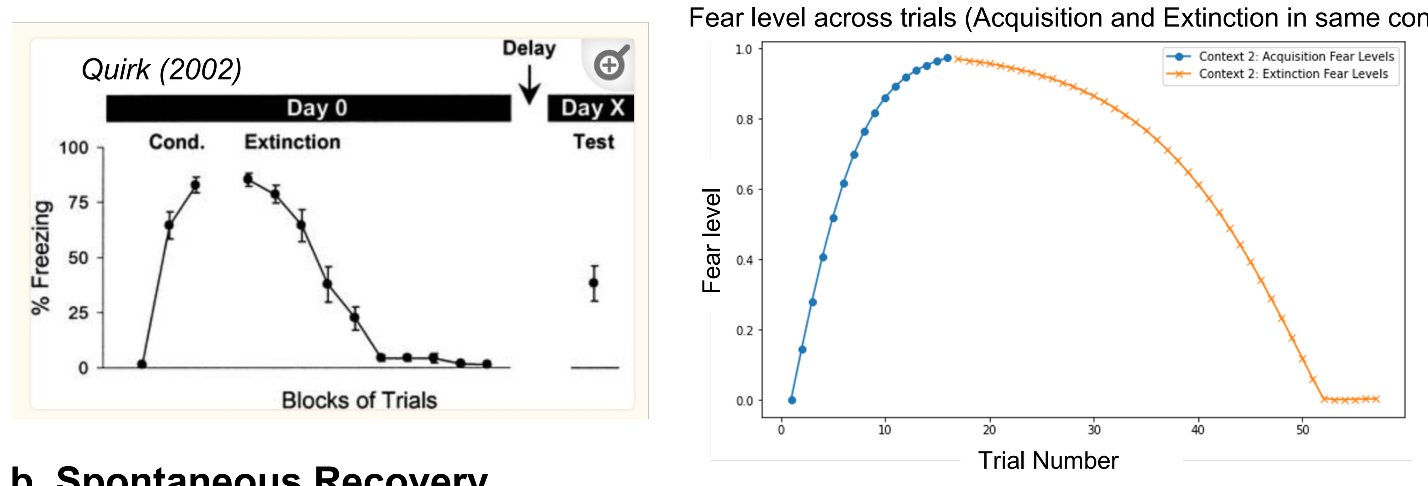


II. Model Architecture

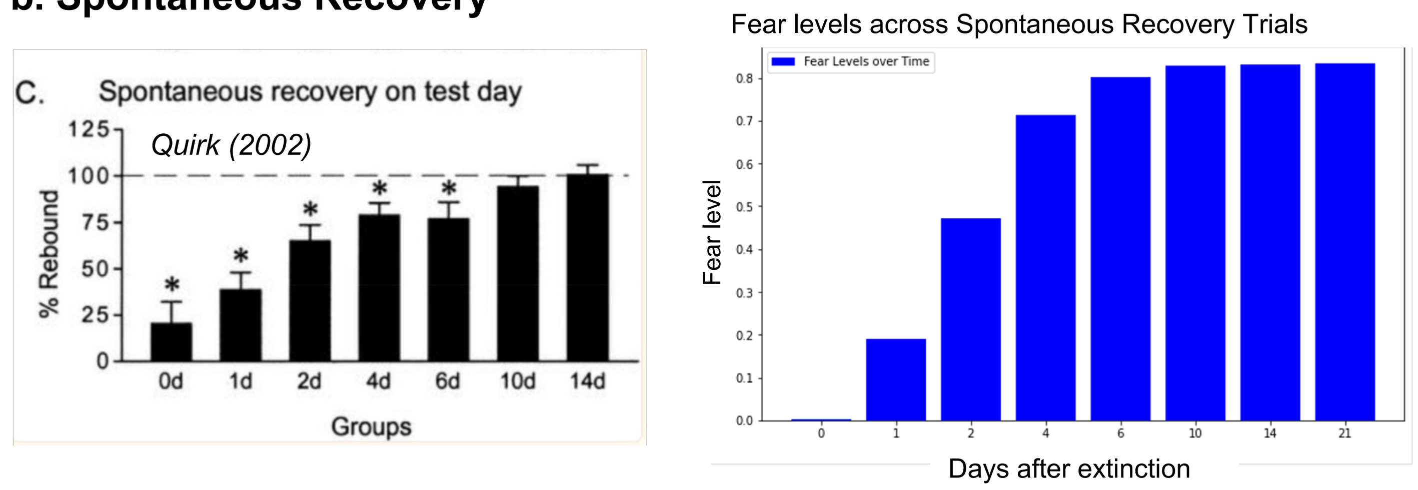


III. Model Results from Replicating Empirical Phenomena

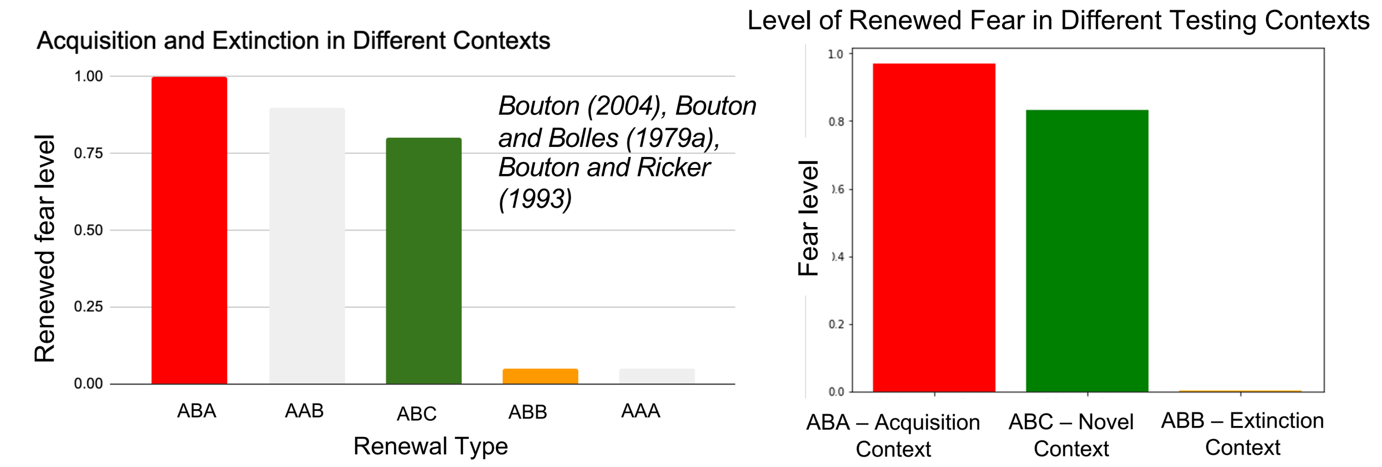
a. Fear Acquisition and Extinction across Trials



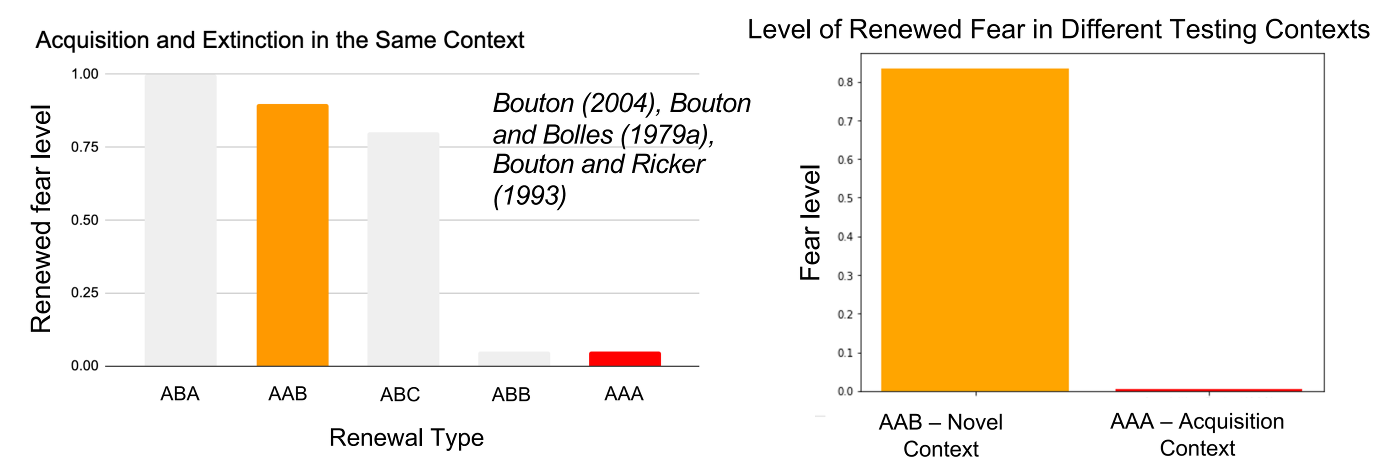
b. Spontaneous Recovery



c. Fear Renewal with Different Acquisition (A) and Extinction (B) Contexts



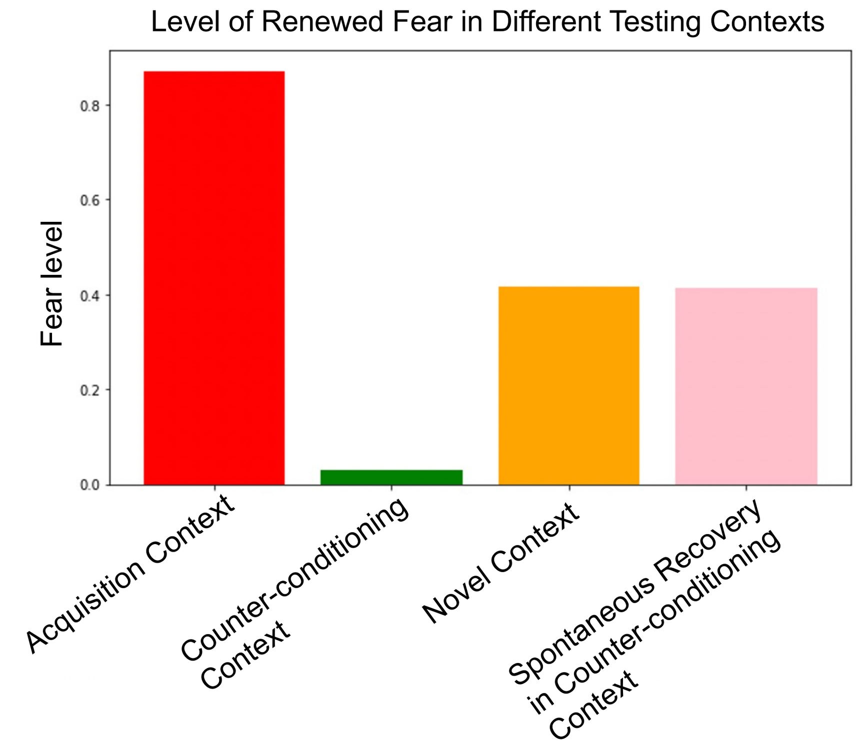
d. Fear Renewal with the same Acquisition (A) and Extinction (A) Contexts



Discussion

- Our computational model of fear conditioning constrained by key neuroscientific insights can simulate return of fear phenomena – specifically fear renewal in different context combinations and spontaneous recovery in the extinction context.
- It can also replicate experimental findings it was not explicitly tuned for; contextual conditioning, external activation of the extinction engram, and extinction in multiple contexts (not shown).
- It provides testable predictions on the return of fear following counterconditioning: counterconditioning might lead to a more permanent loss of fear, with lower context-specificity to the counterconditioning context (unlike extinction).

IV. Predictions for Return-of-Fear after Counterconditioning



References

1. Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & memory*, *11*(5), 485-494.
2. Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and motivation*, *10*(4), 445-466.
3. Rescorla, R. A. (2004). Spontaneous recovery. Learning & Memory, 11(5), 501-509.
4. Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: implications for fear conditioning, extinction and psychopathology. Nature reviews neuroscience, 14(6), 417-428
5. Zhang, X., Kim, J., & Tonegawa, S. (2020). Amygdala reward neurons form and store fear extinction memory. *Neuron*, *105*(6), 1077-1093.

LSA PSYCHOLOGY UNIVERSITY OF MICHIGAN