Effects of Oxytocin and Muscimol on a Spiking Neuron Model of Fear Conditioning, Expression, and Extinction

Peter Duggins (psipeter@gmail.com)

SYDE 556/750 Simulating Neurobiological Systems
April 25, 2016

1 Introduction

The amygdala, a heterogeneous collection of nuclei located in the temporal lobe, plays a critical role in the neural circuitry underlying fear learning and fear expression. *Fear conditioning* is the process by which animals learn associations between neutral stimuli (like auditory tones) that occur alongside unpleasant stimuli (like electric shocks). After such learning, animals will express a *fear response* (such as freezing) when the neutral stimulus is presented by itself. Animals have great difficulty unlearning fear conditioning: *extinction conditioning* occurs slowly and effectively suppresses the fear response orchestrated by the amygdala rather than overwriting the original association. Numerous theoretical models describe fear learning and extinction, and a wealth of electrophysiological and anatomical data corroborate the amygdala's role in these processes, making it one of the best-studied emotional systems in the brain.

Computational models of fear conditioning, expression, and extinction are valuable tools for understanding the relationships between low-level neuroanatomy and high-level functionality, but most models to date have focused on only one such level of analysis. For instance, Kolbeck [10]'s spiking neural network model performs first- and second-order fear conditioning, blocking, and extinction, but does not explore the complex circuitry within the amygdala, relegating the relevant computations to other brains areas such as the periaqueductal gray (PAG), hippocampus, cortex, and thalamus. In contrast, some biophysical models investigate how electrophysiology induces neural plasticity and learning withing the amygdala [9], while others study how competition between multicompartment neurons controls expression vs extinction of conditioned fear [12]. However, such models rarely incorporate these mechanisms into functional circuits that can learn to produce a behavioral fear response given input stimuli.

In this project, I propose a functional, anatomically-accurate model of the amygdala, then investigate the effects of two drugs known to affect fear learning and expression. These drugs not only affect fear behavior when injected directly into the amygdala [13, 19], but can safely be ingested or applied through a nasal spray. Recent studies suggest that these latter application methods may disrupt consolidation of fear conditioning and

improve the efficacy of extinction-training in humans, making them valuable resources in treating resilient fear disorders such as as phobias and PTSD [17, 1]. Computational models of these drugs will help clarify their functional interactions with amygdala fear circuits and potentially improve the design and testing of drug treatments. Here, I expand Kolbeck's fear conditioning model to include more anatomical detail, then show that the model reproduces behavioral data on how two drugs, oxytocin and muscimol, affect fear learning and expression.

2 Background

2.1 Fear conditioning

In traditional fear conditioning experiments, a rat is placed in a conditioning cage, then repeatedly presented with a neural sensory cue (auditory, olfactory, or visual) followed by an unpleasant cue (footshock). After enough of these pairings, the rat learns that the neural cue, or conditioned stimulus (CS) predicts the onset of the unpleasant cue, or unconditioned stimulus (US). When presented with the CS alone, the rat will exhibit the fear response (freezing).

Learning occurs through modification of connection weights between areas of the brain representing cue information from the environment, such as cortex, thalamus, and hippocampus, and the amygdala (as well as between subpopulations within the amygdala). These areas communicate information about CS and US to input populations in the amygdala, particularly the lateral amygdala (LA). This information is used to compute an "error" signal that modifies connection weights into/within amygdala[1]. Successful conditioning is evidenced not only by the rat's behavioral response, but by the increased electrical activity of amygdala populations, including LA and basal nuclei (BA), in response to CS presentation after training [15]. LA and BA connect to a third amygdala population, the central medial amygdala (CeM), an output population which projects to motor neurons in the brainstem and PAG, inhibiting them and causing freezing [5]

2.2 Context conditioning and fear extinction

During fear conditioning, rats also learn to associate information about the environment, or *context*, with the onset of the US. This may include persistent auditory, visual, or olfactory cues, such as the coloration pattern of the conditioning cage where fear learning took place. Context conditioning is assessed by training a rat as above, removing it for a recovery period, then returning it to the conditioning cage and observing the fear response in the absence of any presented CS or US. The rat will freeze (more often than baseline, but less often than when the CS is presented) due to the secondary associations it has formed between the context and the US. However, after spending enough time in this (or another) context, the rat begins to "unlearn" this contextual conditioning. Similarly, when a rat is conditioned, placed in a novel context, and repeatedly presented

[1] the precise source of this error population is disputed [11]

with the CS without an associated US, the rat gradually learns the CS no longer predicts the US in that context, and freezes less often.

In reality, however, animals do not unlearn fear conditioning. Instead, they learn to suppress the fear response in particular contexts which they have learned are "safe". The original conditioned association still exists, and will dominate the animal's behavior in any novel context. After *extinction conditioning* in which a previously-conditioned CS is paired with (no US) in a particular context, returning the animal to the original cage, or to a cage unlike either the conditioning cage or extinction cage, will result in normal fear responses to CS presentation.

Context and extinction conditioning occur most strongly in the BA, allowing the original association stored in LA connection weights to be maintained while permitting the LA and BA to compete for the appropriate fear/extinction response [4]. Studies have shown that distinct populations within BA become more/less active after context and extinction learning [8], implying that intra-BA circuits gate expression of fear versus extinction. Analogous to LA, it is believed that brain areas representing context information, such as hippocampus and cortex, project to these populations via different channels, and that these connection weights are modified during context and extinction training, leading to increased/decreased post-training responses in BA [11]. The GABAergic interneuron system in BA, including populations expressing peptide cholecystokinin (CCK) and parvalbumin (PV), seems particularly important in this learned, dynamic gating, and ensure that fear and extinction responses cannot occur simultaneously [6]. Connection weights between the interneuron population and fear/extinction-activated BA neurons may also be modified during context and extinction conditioning.

2.3 Drug Effects

Fear conditioning, expression, and extinction are strongly affected by changes in neurochemistry within the amygdala. Neuromodulators like dopamine, norepinephrine, and oxytocin (OXY), as well as drugs like muscimol (MUSC) and propanolol, regulate intra-LA and intra-BA inhibitory circuits through excitation and suppression of GABAergic interneurons (see Duvarci and Pare [6]). The differential expression of receptors within the amygdala, as well as recent techniques for targeted drug injections, allow researchers to study the dynamics of amygdala circuits using pharmacological techniques.

- OXY receptors are found exclusively in neurons in the lateral central amygdala (CeL) that develop inhibitory responses to CS after fear conditioning[19]. Applying oxytocin agonist $[Thr^4, Gly^7]$-oxytocin (TGOT) to this population excites these inhibitory cells, which project to CeM and reduce conditioned fear responses.

- MUSC, a GABA agonist, mimics the effects of inhibitory interneuron activity when injected into LA and BA. These regions are involved both in fear conditioning and fear response, so applying the drug reduces both fear learning and fear expression, depending on when in the training/testing regime it is injected [13]. Propanolol, a drug which has been used to treat phobias and PTSD, has similar effects [17].

- Dopamine and norepinephrine exert the opposite effect of MUSC: they inhibit inhibitory interneurons, increasing the activity of LA and BA compared to baseline. These neuromodulators have been shown to facilitate synaptic plasticity within LA and are implicated in the acquisition of conditioned fear, implying that fear learning may be under natural modulatory control by other brain areas [6].

3 Model Description

3.1 System Description

This model combines the functional properties of Kolbeck [10], including the implementation of transformations and learning rules in the Neural Engineering Framework (NEF, [7]), with a detailed anatomical model of the amygdala, including LA, BA, ITC, CeL, and CeM [6]. A schematic of the model is shown in figure 1, with a brief description in the figure caption. Below, I describe each of the subpopulations, the variables they represent, the functions they compute, and the learning they undergo.

3.1.1 Lateral Amygdala (LA)

`LA` is the primary input ensemble to the amygdala. `CS` and `US` information, represented as one-dimensional vectors, enter from external populations `C` and `U`, which convey information from the environment (`stim_C` and `stim_U`). These representations are input into separate dimensions of a two-dimensional `LA`, consistent with evidence that single neurons in `LA` encode both CS and US information [16].[2]

`error_cond` is an error population that controls classical fear conditioning. It receives input from `LA`, which is functionally transformed to compute the function $f(x) = US - CS$. When the first dimension of `LA` (representing response to CS) is small, but the second dimension (representing US) is large, `error_cond` will represent a positive value. This value is input to the Prescribed Error Sensitivity rule (PES), which updates the weights between `C` and `LA` (initialized to zero). Functionally, this learning trains `LA` to respond more strongly in the future (represent a larger value in the CS dimension) if a shock is present and the ensemble is not responding to it ($US - CS < 0$). Importantly, in keeping with empirical results on fear conditioning and extinction, `error_cond` only represents positive values[3], prohibiting training on negative error (large CS response while shock low) that would cause unlearning of the CS-US association.

`LA_inter` is a control population that receives two-dimensional input from `LA` and projects inhibition back to it. `LA_inter` is also stimulated by `stim_GABA` in the MUSC experiments. `LA_inter` sends recurrent negative feedback to `LA` using the Parisien transform [14, 18]. The Parisien transform forces all weights from `LA_inter` to `LA` to be negative and project directly onto the `LA.neurons`, bypassing the encoders to inject

[2] I experimented with implementing a differentiator here (using Bryan's two-PSTC method) to capture the physiological result that LA neurons only respond transiently to CS/US onset. However, I was unable to get this working alongside the PES learning and Parisien inhibition.

[3] All encoders and intercepts are positive; learning rate 5×10^{-4}

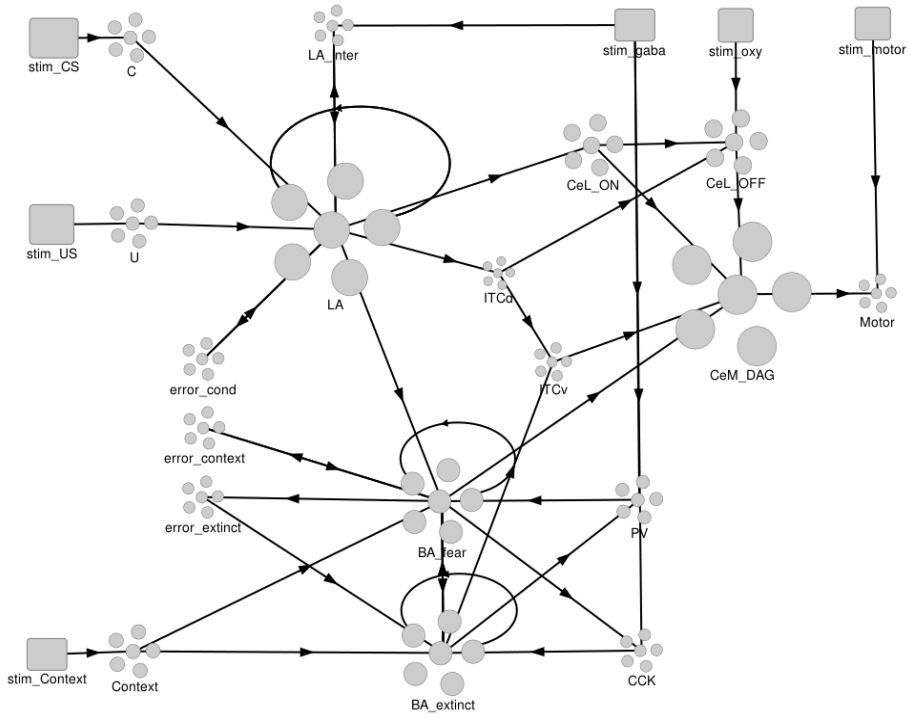

Figure 1: Schematic of the model visualized using `nengo_gui`. **LA**: CS and US information enters `LA` through intermediates `C` and `U`. If `LA` doesnt respond strongly enough to CS (a US is present without an appropriate response), the connection between `C` and `LA` is strengthened by `error_cond`, causing fear conditioning. `LA_inter` inhibits `LA` and `error_cond`, decreasing the fear response to CS as well as reducing the effective learning rate. **BA** Context information enters `BA_fear` and `BA_extinct` through separate connections. If `BA_fear` doesnt respond strongly enough to context, the connection between `context` and `BA_fear` is strengthened by `error_context`, causing contextual fear conditioning. If `BA_fear` responds too strongly to context (a strong response exists without a justifying US), the connection between `context` and `BA_extinct` is strengthened by `error_extinction`. `BA_fear` and `BA_extinct` compete through interneurons `CCK` and `PV`. **ITC, CeL, CeM**: `LA`, `BA_fear`, and `BA_extinct` project indirectly to `CeM` via the `ITC` and `CeL` ensembles. These populations create further competition, allowing `BA_extinct` to extinguish fear responses in contexts where fear conditioning has been unlearned. `CeM` inhibits a tonically active `motor` population, causing freezing when `LA` and/or `BA_fear` are active. **Drugs**: Interneurons are stimulated externally by MUSC and OXY, which effectively dampen the activity of the associated populations, reducing fear conditioning, expression, and extinction.

a uniformly negative bias onto LA and damping its activity. Functionally, GABAergic inhibition of LA through LA_inter decreases the response of LA to C. This both reduces the output that LA sends to CeM (suppressing conditioned freezing) and reduces the value represented in error_cond (slowing fear conditioning).

3.1.2 Basal Nuclei (BA)

BA populations have representations and transformations similar to LA, except that they learn context-US associations instead of CS-US associations. The main ensembles in BA, BA_fear and BA_extinct, develop distinct responses to fear and extinction training. stim_context stimulates an intermediate ensemble context, which feeds (through distinct connections) to the first dimension of both BA_fear and BA_extinct, which represents their respective responses to context: fear and extinction. BA_fear is also excited by LA [6] through a communication channel conveying CS to the first dimension of BA_fear US to the second. BA_fear and BA_extinct project indirectly to CeM via excitatory and inhibitory pathways in ITC (discussed below), inducing and reducing freezing respectively.

error_context is an error ensemble that controls fear conditioning to contextual information. Like error_cond, it receives context and US information from BA_fear and computes their difference using the function $f(x) = US - F$, where F is the dimension in BA_fear that represents the combined fear response to CS and context. error_context represents positive values, so it strictly increases the fear responses to context information. In contrast, error_extinct receives the same inputs from BA_fear, but computes the function $f(x) = F - US$, which is nonzero when the fear response is large but there is no associated shock, indicating a behavioral overreaction to a "safe" stimulus. error_extinct is used to learn the connection between context and BA_extinct, so that contexts in which the US is consistently absent activate an extinction response which ultimately inhibits freezing.

BA interneurons are also divided into two populations, CCK and PV. These are recurrently connected to BA_fear and BA_extinct as described in [6] and shown in figure 1. Unlike in LA and ITC/CeL, these connections have not yet been anatomically identified, but make functional sense: BA_fear inhibits BA_extinct through CCK, while BA_extinct inhibits BA_fear through PV. This mutual inhibition prevents the BA from learning/exhibiting fear and extinction responses simultaneously. As with LA_inter, CCK and PV are stimulated by stim_GABA, inhibiting their respective populations, reducing contextual fear/extinction responses, and reducing context/extinction learning.[4]

3.1.3 Intercalatec Cells (ITC or ICM) and Central Lateral Amygdala (CeL)

ITC ensembles act as intermediaries between LA/BA and CeM. ITCd is excited by LA and disinhibits CeM neurons through inhibitory projections to CeL (causing freezing), while

[4]There is also evidence that BA_fear and BA_extinct have recurrent excitatory connections with areas outside of BA, including prelimbic (PL) and infralimbic (IL) areas. These connections prolong the activity of these populations in response to CS/context [2]. I connected these populations recurrently, but found that small feedback did not increase response length, while large feedback broke the circuit functionally.

ITCv is excited by BA_extinct and inhibits CeM neurons (causing unfreezing). ITCd may also "override" the extinction response via inhibitory projections to ITCv. Functionally these populations add little to the model: I have included them for completeness with respect to amygdala anatomy, and to, in future work, reproduce empirical studies that target them specifically[5].

CeL acts as an intermediary between LA/ITCd and CeM. It is known to contain cells that show only excitatory (CeL_ON) and only inhibitory (CeL_OFF) responses to CS. These cells express distinct suites of intracellular proteins and neurotransmitter receptors: notably, only CeL_OFF expresses the OXY receptor. In the model, bound OXY excites this population through stim_oxy. As in BA, CeL_ON and CeL_OFF are mutually inhibitory, but it seems that only CeL_OFF, a GABAergic inhibitory population, is connected to CeM. Again, these ensembles serve little functional purpose, as all their transformations are either $f(x) = x$ or $f(x) = -x$; they are included for completeness and to model the known effects of OXY on fear response.[6]

3.1.4 Centeral Medial Amygdala (CeM) and Motor

The output population CeM is fed excitation by BA_fear and inhibition by ITCv and CeL_OFF. The population simply sums the input and outputs the function $f(x) = -x$ to the motor population. In turn, motor is fed a continuous positive signal by stim_motor, causing it to be tonically active unless inhibited by CeM. The value represented in motor represents the motor activity of the animal (as initiated by brainstem or PAG); freezing is equal to the ensemble's saturation value minus its current value.

3.2 Design Specification

Unless specified below, all model parameters, including radius, evaluation points, firing rates, and noise, are set equal to their nengo defaults.

- Stimuli (CS, US, context) are represented as one-dimensional vectors with constant values, but acquire noise and delay when passed through the intermediary C, U, and context ensembles.
- To ensure that the error populations represent positive values when CS and US are both on, I use $US = 2$, $CS = 1$. This requires that populations which represent US (U,LA,BA) have radius=2.
- The error and interneuron populations (as a result of the Parisien transform) have positive encoders and intercepts, forcing them to represent only positive numbers.
- Temporal filters are the standard post-synaptic filters, with τ_{PSTC} equal to one of five values based on the information passing through the connection: tau_stim for stimulating the intermediary stimulus populations; tau_drug for applying

[5] see section "Extinction Also Depends on Gating of BA Inputs to CeM by Intercalated Cells" in [6]

[6] I do not use the Parisien transform on these populations because some of them require the pure-inhibition transform, which is not yet implemented in Nengo. In this sense, what I refer to as "inhibition" in these populations is actually a negative transform rather than GABAergic inhibition

MUSC and OXY to the interneurons; `tau_recurrent` for recurrent inhibition from interneurons back to `LA` and `BA`; `tau_learn` for connection of error populations onto learning rules; and `tau` for all other ensemble-to-ensemble connections.

- Dynamics of represented variables, which include CS response, context response, extinction response, and summed response, are dictated by transformation described above, notably dynamically learned connections between inputs and `LA`/`BA`, recurrent inhibition by drug-modulated interneurons, and summation through various communication channels.

3.3 Implementation

Decoders were computed using the standard NEF least-squares optimization [7], and error was computed using using the PES learning rule [3]. The Parisien transform was computed using code available at `https://github.com/nengo/nengo/issues/921`. The model was implemented in `Nengo 2.1.0-dev` and `nengo_gui 0.2.0-dev` and run on a standard workstation. The code, including raw data (`.pkl` files), parameters (`.json` files) and figures, is available at `https://github.com/psipeter/fear`.

4 Results

4.1 Validation

I began by confirming that the model exhibits fear conditioning, expression, and extinction in the absence of any drugs, and that each ensemble appropriately represents and transforms its variables. During the training phase, the model is presented with $CS = 1$ for a duration of $t = 10$s, with $US = 2$ presented from $7 < t < 9$s, alongside the stimulus $context = 1$. This pairing is repeated ten times. In the testing phase, the model is either placed in a new context ($context = -1$) and presented with a sustained $CS = 1$ to test fear conditioning (labeled a "tone" experiment), or is returned to the original context but presented no CS ($context = 1$, $CS = 0$) to test context conditioning (labeled a "context" experiment). Finally, in the extinction phase, the model is returned to the original environment and presented with a sustained CS. Freezing is measured continuously during the simulation, along with the value of the major populations. Figure 2 shows the model behavior over ten trials for a tone experiment (top) and a context experiment (bottom).

As CS-US pairs are presented to the model during the training stage, the responses of `LA`, `BA_fear`, and `freeze` to CS increase as expected. The `error_cond` and `error_context` populations represent large numbers initially because `LA` and `BA_fear` have small responses to CS and context, but as conditioning changes the relevant connection weights, the ensembles' fear/context responses increase, necessitating smaller error signals for training. Around $t = 700$, the testing stage begins, and `C` turns on while `context` switches to a novel value. `LA`, `BA_fear`, and `freeze` remain active for the duration of the testing period. However, `error_extinct` also responds strongly because `BA_fear` exhibits a strong fear response despite no US. As the extinction phase progresses ($t > 1000$), this error signal modifies the connection weight between

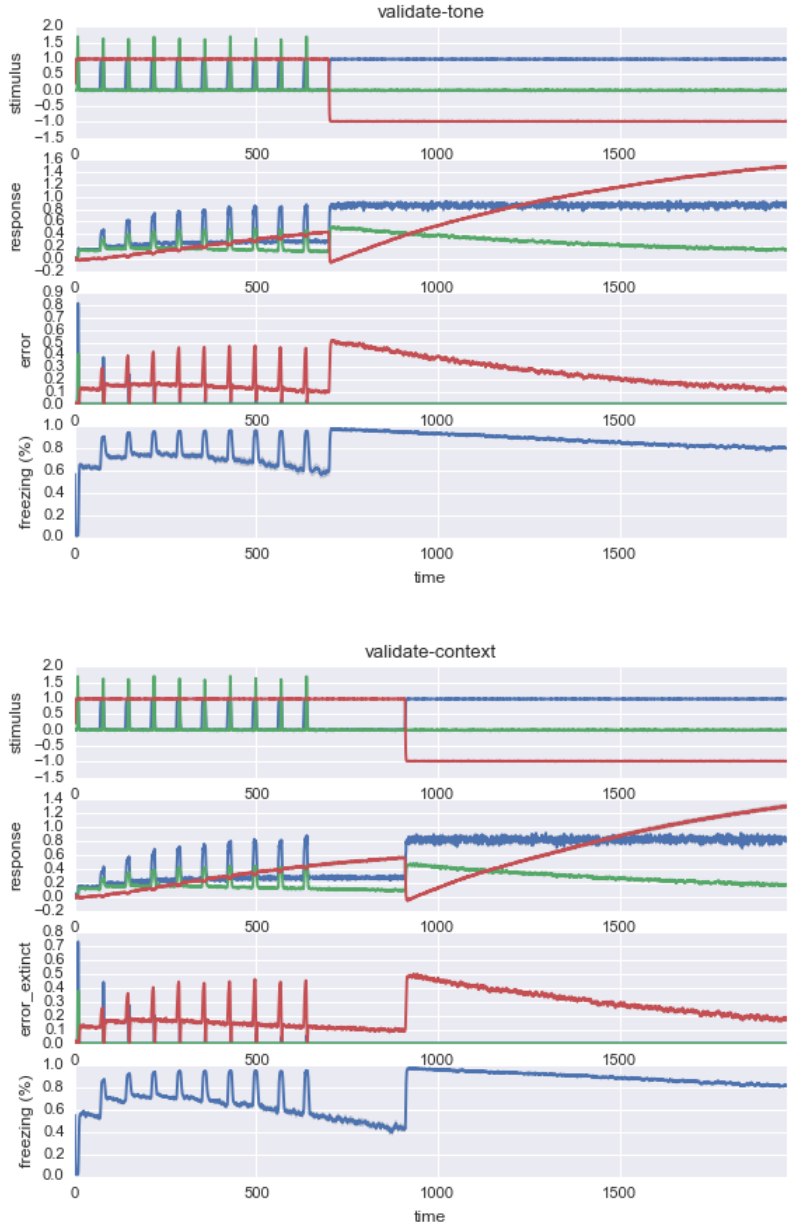

Figure 2: Fear conditioning, expression, and extinction in a typical tone (top) and context (bottom) experiment. Values represent the mean decoded values from the indicated ensembles, with 68% bootstrap confidence intervals plotted in gray ($n_{boot} = 3000$) **Legend**: *stimulus:* CS-blue, US-green, context-red; *response:* LA-blue, BA_fear-green, BA_extinct-red; *error:* error_cond-blue, error_context-green, error_extinct-red.

context and BA_extinct, causing the latter to respond more strongly. The increasing value in BA_extinct inhibits BA_fear through PV, causing freezing to slowly degrade. However, this extinction training does not affect LA, which continues to maximally respond to the presence of CS.

The behavior of the model in the context experiment is identical in the training stage. During testing, the model continues to exhibit freezing because the value of context remains the same as during context conditioning, ensuring a moderate level of activity in BA_fear. However, the magnitude of freezing is lower due to the quieting of LA with the removal of the CS. During extinction, error_extinct rises and BA_fear falls as before, suppressing the fear response.

4.2 Oxytocin

In this experiment, I apply OXY to CeL with the goal of impairing fear expression. In a similar experiment, Viviani et al. [19] induce fear conditioning in rats by placing them in the conditioning cage and giving them seven randomly-spaced footshocks over a twenty minute period. No CS is presented, so the rats' freezing response results soley from context conditioning. On the next day, TGOT, the OXY receptor agonist, is injected into the CeL of rats in experimental group. The control and experimental groups are returned to the conditioning cage and their freezing is measured. Finally, both groups are returned to the cage again on the last day to assess potential long-term effects of OXY application. Figure 3 (top) shows the behavioral result that rats in the experimental group froze significantly more often than in the control group on the testing day, but did not demonstrate significantly different fear responses than control a day later.

To recreate the experimental conditions, I removed the CS from the training and testing regimes, and observed the changes in freezing that occurred over time in four experimental groups: those that received no drug ("saline") vs TGOT (stim_GABA applied to CeL_OFF) during training vs testing. My hypothesis was that receiving OXY during training would not reduce the freezing during the testing phase, but that receiving OXY during testing would inhibit the conditioned contextual fear response. Figure 3 (bottom) shows the magnitude of freezing over the course of the experiment for the four experimental groups, as well as average freezing during each phase of the experiment. As predicted, OXY reduces test-phase freezing when applied during testing (sharp decline of the red and purple lines), but does not reduce freezing when applied solely during training (green line). All four populations undergo extinction at a similar rate.

4.3 Muscimol

In this experiment, I apply MUSC to GABAergic interneuron ensembles with the goal of impairing both fear conditioning and fear extinction. In a similar experiment, Muller et al. [13] induce fear conditioning in rats using a standard CS/US pairing (seven trials), then assess fear expression in both a tone condition (CS presented in a novel context) and a context condition (no CS presented, but testing in the conditioned context). Muller et al. inject saline/MUSC into the LA and BA of rats before training/testing, creating the four-factor that I experimentally performed above. Figure 4 shows the behavioral

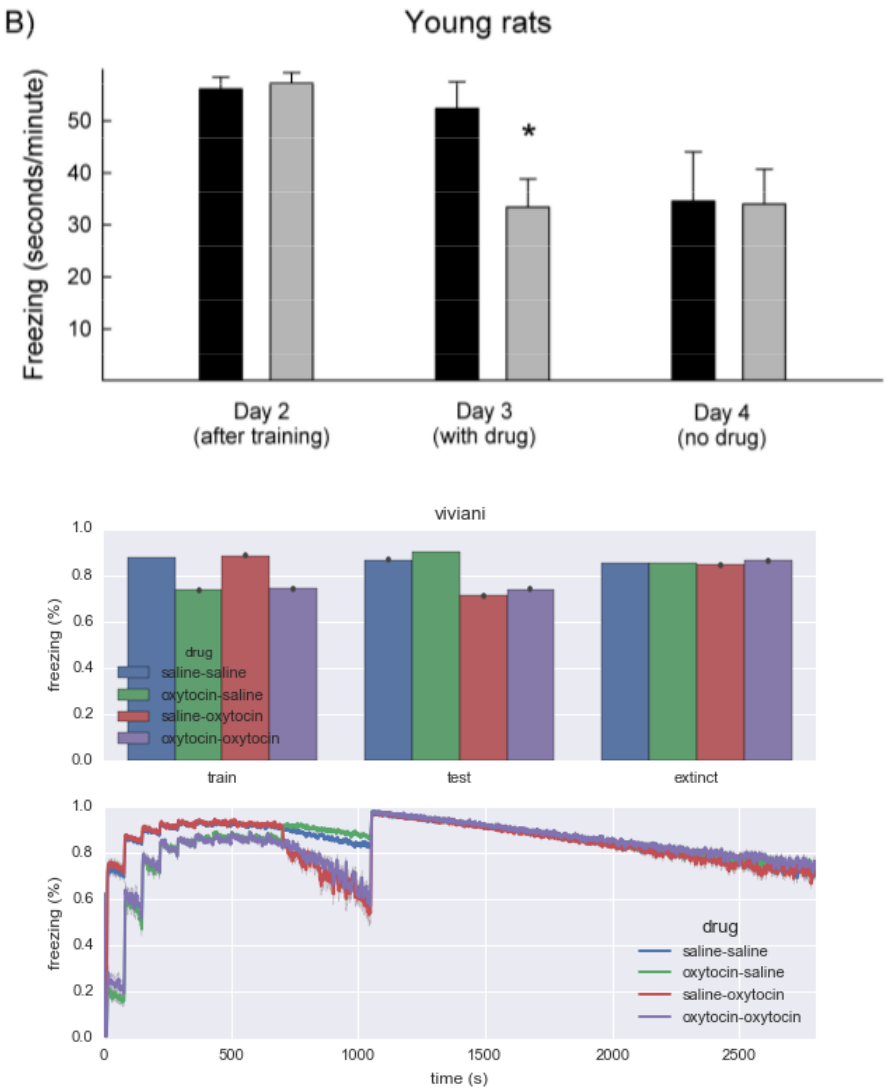

Figure 3: Top: contextual fear responses in rats which have been injected with oxytocin (gray bars) vs saline (black bars) prior to testing. Reproduced from the supplementary materials of Viviani et al. [19]. Bottom: Simulated freezing under the effects of four drug treatments. In agreement with Viviani et al, only oxytocin applied during testing significantly decreases freezing. Legend entries correspong to [training drug]-[testing drug]. 68% bootstrap confidence intervals plotted in gray ($n_{boot} = 3000$).

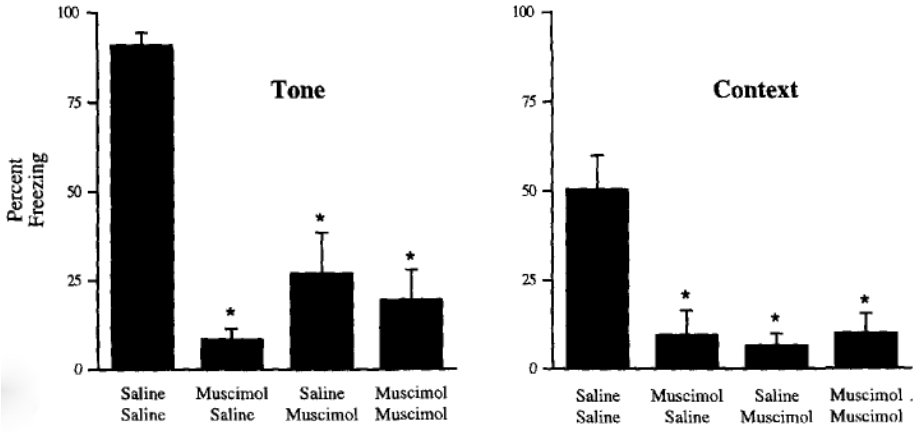

Figure 4: Conditioned and contextual fear responses in rats which have been injected with saline or muscimol prior to training and/or testing. Reproduced from Muller et al. [13]

result that rats rarely freeze when MUSC is applied during training or testing, both in the tone and the context condition.

As with OXY, I applied MUSC during training and/or testing by positively stimulating the interneuron populations `LA_inter`, `CCK`, and `PV`. These interneurons in turn inhibited `LA`, `BA_fear`, and `BA_extinct`, which indirectly reduced the values represented in the error populations (not shown). Figure 5 shows that the simulated effects of these drugs captures the behavioral result: applying MUSC inhibits both fear conditioning and expression.[7] Surprisingly, applying MUSC during training in the tone condition reduced freezing much less than in the context condition, and exhibits an interesting nonmonotonic behavior approximately twenty seconds into the testing phase ($t = 720$). I am unsure about the origin of this effect, although it is sensitive to parameters such as the magnitude of MUSC stimulation, the default inhibition in `LA_inter`, and the number of CS-US pairings during training. The differences between the other drug groups were to small to drawn any conclusions.

5 Discussion

Fear conditioning and expression are reduced by the pharmacological stimulation of GABAergic interneurons in specific subpopulations of the amygdala. The effects of drugs such as muscimol and oxytocin fear-related learning and behavior depend on where in the amygdala they are applied, as well as when in the training/recall process they are injected. Oxytocin only affects the CeL, a GABAergic population that expresses the OXY-receptor. Because CeL appears late in the fear learning/expression circuit,

[7] As expected giving the slower learning rate of context conditioning compared to fear conditioning, the baseline fear response was lower in the context condition than in the tone condition.

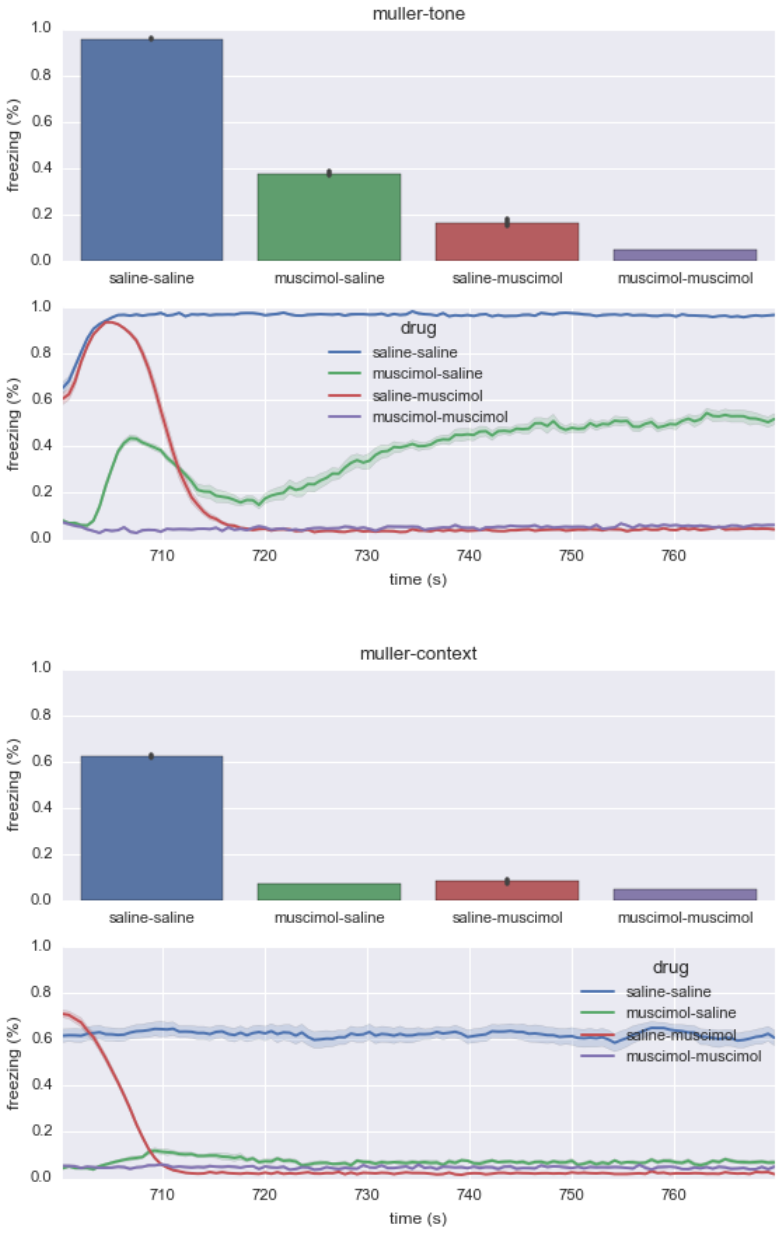

Figure 5: Simulated freezing under the effects of four drug treatments for the tone (top) and the context (bottom) experiments. In agreement with Muller et al, muscimol prevented fear learning when applied during training and prevented fear expression when applied during testing. Legend entries correspond to [training drug]-[testing drug]. 68% bootstrap confidence intervals plotted in gray ($n_{boot} = 3000$).

its activity only suppresses fear responses (through inhibition of the output population CeM), not fear conditioning. Muscimol, a GABA agonist, affects populations in the LA and BA which represent fear/context response. Because these populations influence the rate of learning in the model, applying Muscimol also affects fear conditioning when applied during training. The model's success suggests that simple approximations of these drugs interaction with interneurons (linear excitation) is sufficient to describe the behavioral results [6]. Although detailed biophysical models such as [9] clarify the mechanisms underlying learning and neurochemical stimulation, they may not be needed for functionally-useful and behaviorally-predictive descriptions of the amygdala fear circuit.

I found that changing the learning rate parameters in the model affected the speed of conditioning and extinction, and that altering the relative weights of populations projecting to CeM shifted the balance of fear vs. extinction in the animal's fear response. It could be interesting to model the differences between individuals' susceptibility to extinction training by varying these parameters. I also found that the Parisian transform was necessary to correctly model recurrent inhibition. Earlier versions of the model which used negative transforms, or negative transforms directly onto the neurons, either did not dampen the target populations at all, or had nonsensical responses to increase in `stim_GABA`. However, the transform is not always necessary, as can be seen by the effective action of the `ITC`/`CeL` pathways after oxytocin administration. Finally, I saw that the PES learning rule can be used to train connection weights based on post-synaptic activity if that activity feeds directly into the error population. This "creative" use of the learning rule bypasses the need for a BCM rule; however, it makes strong assumptions about the location/connectivity of this error population, which are not yet justified by empirical evidence from amygdala.

6 Conclusion

The amygdala circuitry underlying fear learning and expression is one of the best-studied emotional systems in the brain. Numerous simple and complex models have previously simulated these processes, both functionally and biophysically, but few models have included the relevant anatomical detail while providing a full input-output description of fear conditioning. New models are needed to simulate the interactions of amygdala subpopulations in these processes, to model the effects of drugs, and to eventually understand the neuromodulatory role of amygdala for other cognitive and emotional tasks. In this project, I showed that the empirical effects of two drugs could be reproduced with a spiking neuron model of the amygdala through the simulated excitation of the GABAergic interneuron system.

Several extensions of the model are immediately obvious. Incorporating greater biophysical detail in the neurons would allow a more realistic approximation of the electrophyiological effects of drug application on interneurons, as well as allow a wider range of drugs to be simulated. Using higher-dimensional representations would increase the psychological plausibility of the model, allowing for multiple visual/auditory/olfactory features to exist within the external stimuli, permitting a multidimensional comparison of learned fear cues with current inputs (via a dot product), and opening up possibilities

for more complex binding operations (circular convolution). A deeper investigation of the necessity of Parisien transforms in various parts of the network and the utility of the hPES rule would also be theoretically interesting. In the long term, I would like to interface the amygdala with other brain systems to study the modulatory role it plays in other cognitive processes, particularly motivation, attention, and memory consolidation/recall.

References

Acheson, D., Feifel, D., de Wilde, S., Mckinney, R., Lohr, J., and Risbrough, V. (2013). The effect of intranasal oxytocin treatment on conditioned fear extinction and recall in a healthy human sample. *Psychopharmacology*, 229(1):199–208.

Amano, T., Duvarci, S., Popa, D., and Paré, D. (2011). The fear circuit revisited: contributions of the basal amygdala nuclei to conditioned fear. *The journal of neuroscience*, 31(43):15481–15489.

Bekolay, T., Kolbeck, C., and Eliasmith, C. (2013). Simultaneous unsupervised and supervised learning of cognitive functions in biologically plausible spiking neural networks. In *Proceedings of the 35th annual conference of the cognitive science society*, pages 169–174.

Bouton, M. E., Westbrook, R. F., Corcoran, K. A., and Maren, S. (2006). Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biological psychiatry*, 60(4):352–360.

Ciocchi, S., Herry, C., Grenier, F., Wolff, S. B., Letzkus, J. J., Vlachos, I., Ehrlich, I., Sprengel, R., Deisseroth, K., Stadler, M. B., et al. (2010). Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature*, 468(7321):277–282.

Duvarci, S. and Pare, D. (2014). Amygdala microcircuits controlling learned fear. *Neuron*, 82(5):966–980.

Eliasmith, C. and Anderson, C. H. (2003). *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT press.

Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., and Lüthi, A. (2008). Switching on and off fear by distinct neuronal circuits. *Nature*, 454(7204):600–606.

Kim, D., Paré, D., and Nair, S. S. (2013). Mechanisms contributing to the induction and storage of pavlovian fear memories in the lateral amygdala. *Learning & Memory*, 20(8):421–430.

Kolbeck, C. (2013). A neurocomputational model of the mammalian fear conditioning circuit. Master's thesis, University of Waterloo, Waterloo, ON.

Krasne, F. B., Fanselow, M. S., and Zelikowsky, M. (2011). Design of a neurally plausible model of fear learning. *Front. Behav. Neurosci*, 5(41):10–3389.

Li, G., Amano, T., Pare, D., and Nair, S. S. (2011). Impact of infralimbic inputs on intercalated amygdala neurons: a biophysical modeling study. *Learning & Memory*, 18(4):226–240.

Muller, J., Corodimas, K. P., Fridel, Z., and LeDoux, J. E. (1997). Functional inactivation of the lateral and basal nuclei of the amygdala by muscimol infusion prevents fear conditioning to an explicit conditioned stimulus and to contextual stimuli. *Behavioral neuroscience*, 111(4):683.

Parisien, C., Anderson, C. H., and Eliasmith, C. (2008). Solving the problem of negative synaptic weights in cortical models. *Neural computation*, 20(6):1473–1494.

Quirk, G. J., Repa, J. C., and LeDoux, J. E. (1995). Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. *Neuron*, 15(5):1029–1039.

Romanski, L. M., Clugnet, M.-C., Bordi, F., and LeDoux, J. E. (1993). Somatosensory and auditory convergence in the lateral nucleus of the amygdala. *Behavioral neuroscience*, 107(3):444.

Soeter, M. and Kindt, M. (2015). An abrupt transformation of phobic behavior after a post-retrieval amnesic agent. *Biological psychiatry*, 78(12):880–886.

Tripp, B. and Eliasmith, C. (2016). Function approximation in inhibitory networks. *Neural Networks*, 77:95–106.

Viviani, D., Charlet, A., van den Burg, E., Robinet, C., Hurni, N., Abatis, M., Magara, F., and Stoop, R. (2011). Oxytocin selectively gates fear responses through distinct outputs from the central amygdala. *Science*, 333(6038):104–107.