

# MAPPING FEAR BEHAVIOR: NEURAL NETWORKS, VENTRAL TEGMENTAL AREA DOPAMINE, AND ORCHESTRATION OF CONDITIONED DEFENSIVE BEHAVIORS

AMANDA CHU

A dissertation

submitted to the faculty of the Department of  
Psychology & Neuroscience in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

Boston College  
Morrissey College of Arts and Sciences  
Graduate School

June 2024



## **ABSTRACT**

### **Mapping fear behavior: neural networks, ventral tegmental area dopamine, and orchestration of conditioned defensive behaviors**

Amanda Chu

Advisor: Michael A. McDannald, Ph.D.

The ability to appropriately respond to threats is critical for survival. Disruptions in the neural circuits underlying threat responding are studied in animal models and have clinical implications for anxiety disorders in humans. Pavlovian fear conditioning has been extensively used to study the behavioral and neural basis of defensive systems for threat in animals. In a typical procedure, a cue is paired with foot shock, and subsequent cue presentation elicits freezing, a behavior linked to predator detection. Studies have since shown a fear conditioned cue can elicit locomotion, a behavior that - in addition to jumping, and rearing - is linked to imminent or occurring predation. Yet, the full neural circuit for conditioned, activity-promoting behaviors (e.g. locomotion, jumping, and rearing) remains unclear. The overarching goal of this dissertation is to demonstrate that a fear conditioned cue elicits a variety of defensive behaviors and to probe the neural circuit responsible for the expression of such activity-promoting defensive behaviors. To address the lack of research on activity-promoting defensive behaviors, I conducted experiments to observe multiple behaviors during fear discrimination over a baseline of reward seeking and constructed temporal ethograms of behavior. To improve efficiency in behavior scoring for future projects, I devised and trained a machine learning pipeline using convolutional neural networks. To aid in the understanding of the full neural circuit for activity-promoting defensive behaviors, I

investigated the role of dopaminergic neurons of the ventral tegmental area in the expression of the defensive behaviors we observed during fear discrimination.

Ultimately, the findings in this dissertation contribute to our general understanding of fear behavior in animals and may inform therapeutic strategies for anxiety disorders.

## TABLE OF CONTENTS

<b>LIST OF TABLES</b> .....	viii
<b>LIST OF FIGURES</b> .....	ix
<b>LIST OF ABBREVIATIONS</b> .....	xi
<b>ACKNOWLEDGEMENTS</b> .....	xiii
<b>Chapter 1: Introduction to conditioned fear behaviors and its neural underpinnings in rats</b> .....	1
<b>1.1 Pavlovian fear conditioning</b> .....	2
<b>1.2 Defensive behaviors in fear settings</b> .....	3
<b>1.3 The need for more behavior and its challenges</b> .....	5
<b>1.4 Neural pathways involved in defensive behaviors: freezing and flight</b> .....	6
<b>1.5 Dissertation aims</b> .....	9
<b>Chapter 2: A fear conditioned cue orchestrates a suite of behaviors in rats</b> .....	12
<b>2.1 Introduction</b> .....	13
2.1.1 <i>Pavlovian fear conditioning and defensive behaviors</i> .....	13
2.1.2 <i>Chapter Aims</i> .....	15
<b>2.2 Materials and Methods</b> .....	17
2.2.1 <i>Subjects</i> .....	17
2.2.2 <i>Behavior apparatus</i> .....	17
2.2.3 <i>Pellet exposure and nose poke shaping</i> .....	18
2.2.4 <i>Cue pre-exposure</i> .....	18
2.2.5 <i>Pavlovian fear discrimination</i> .....	19
2.2.6 <i>Fear extinction</i> .....	19
2.2.7 <i>Calculating suppression ratio</i> .....	20
2.2.8 <i>Frame capture system</i> .....	21
2.2.9 <i>Post-acquisition frame processing</i> .....	21
2.2.10 <i>Anonymizing trial information</i> .....	22
2.2.11 <i>Behavior categories and definitions</i> .....	22
2.2.12 <i>Frame scoring system</i> .....	24
2.2.13 <i>Inter-observer reliability</i> .....	24
2.2.14 <i>Statistical analyses</i> .....	25
<b>2.3 Results</b> .....	26

2.3.1	<i>Experiment 1</i>	26
2.3.1.1	<i>Conditioned suppression reveals complete discrimination</i>	26
2.3.1.2	<i>Hand scoring with high inter-observer reliability</i>	34
2.3.1.3	<i>Temporal ethograms reveal shifting behavioral patterns over discrimination</i>	35
2.3.1.4	<i>Danger orchestrates a suite of behaviors</i>	38
2.3.1.5	<i>Associatively acquired behaviors generalize early</i>	40
2.3.1.6	<i>Sex informs the temporal pattern of freezing</i>	41
2.3.1.7	<i>Danger-elicited behaviors are independently expressed</i>	43
2.3.2	<i>Experiment 2</i>	45
2.3.2.1	<i>Conditioned suppression reveals complete discrimination during extinction with reward apparatus present</i>	45
2.3.2.2	<i>Danger-elicited locomotion peaks when foot shock would have occurred</i>	48
2.3.2.3	<i>Freezing is less danger-specific and is sensitive to time, test type, and order</i>	52
2.3.2.4	<i>Danger-elicited behaviors are independently expressed during extinction</i>	53
2.4	<b>Discussion</b>	56
<b>Chapter 3: Neural networks for comprehensive ethograms</b>		64
3.1	<b>Introduction</b>	65
3.1.1	<i>Overview of neural networks</i>	66
3.1.2	<i>DeepLabCut</i>	69
3.1.3	<i>B-SOiD</i>	70
3.1.4	<i>BehaviorDEPOT</i>	71
3.1.5	<i>DeepEthogram</i>	72
3.1.6	<i>Chapter aim</i>	73
3.2	<b>Materials and Methods</b>	75
3.2.1	<i>Datasets</i>	75
3.2.2	<i>Feature information from basic convolutional neural network</i>	75
3.2.3	<i>Feature information from ResNet18</i>	77
3.2.4	<i>Motion information from optical flow estimation</i>	79
3.3	<b>Results</b>	81
3.3.1	<i>ResNet18 model achieves proficient overall accuracy but lacks accuracy for specific behaviors of interest</i>	81
3.3.2	<i>Optical flow estimation</i>	83
3.4	<b>Discussion</b>	84

<b>Chapter 4: Investigating a role for ventral tegmental area dopaminergic neurons in fear expression</b> .....	85
<b>4.1 Forward</b> .....	86
<b>4.2 Introduction</b> .....	88
4.2.1 <i>The role of the periaqueductal gray in flight and avoidance</i> .....	88
4.2.2 <i>The role of cortical and amygdalar structures in flight and avoidance</i> .....	90
4.2.3 <i>A potential role for dopamine and the VTA in fear behaviors</i> .....	92
4.2.4 <i>The role of dopamine and the VTA in fear learning</i> .....	94
4.2.5 <i>The role of dopamine and the VTA in fear extinction and safety learning</i> .....	95
4.2.6 <i>Chapter aim</i> .....	96
<b>4.3 Materials and Methods</b> .....	97
4.3.1 <i>Subjects</i> .....	97
4.3.2 <i>Stereotaxic surgery</i> .....	97
4.3.3 <i>Behavior apparatus</i> .....	98
4.3.4 <i>Pellet exposure and nose poke shaping</i> .....	98
4.3.5 <i>Cue pre-exposure</i> .....	99
4.3.6 <i>Pavlovian fear discrimination</i> .....	99
4.3.7 <i>Fear extinction</i> .....	99
4.3.8 <i>Calculating suppression ratio</i> .....	100
4.3.9 <i>Frame capture system</i> .....	100
4.3.10 <i>Post-acquisition frame processing</i> .....	101
4.3.11 <i>Anonymizing trial information</i> .....	101
4.3.12 <i>Behavior categories and definitions</i> .....	102
4.3.13 <i>Frame scoring system</i> .....	103
4.3.14 <i>Inter-observer reliability</i> .....	104
4.3.15 <i>Statistical analyses</i> .....	104
4.3.16 <i>Perfusion and tissue collection</i> .....	105
4.3.17 <i>Histology</i> .....	105
4.3.18 <i>Cell quantification</i> .....	105
4.3.19 <i>Genotyping</i> .....	106
<b>4.4 Results</b> .....	107
4.4.1 <i>Caspase-mediated deletion of ventral tegmental area dopaminergic neurons leaves discriminative conditioned suppression intact</i> .....	107
4.4.2 <i>Histological verification of viral infusions</i> .....	111

4.4.3	<i>Hand scoring with high inter-rater reliability</i>	113
4.4.4	<i>Controls did not exhibit expected danger-specific behavioral patterns</i>	114
4.4.4.1	<i>Discrimination</i>	114
4.4.4.2	<i>Extinction</i>	119
4.4.5	<i>Deletion of VTA dopaminergic neurons potentially enhances jumping during extinction</i>	123
4.4.5.1	<i>Discrimination</i>	123
4.4.5.2	<i>Extinction</i>	123
<b>4.5</b>	<b>Discussion</b>	126
4.5.1	<i>Summary</i>	126
4.5.2	<i>Limitations and caveats</i>	126
4.5.3	<i>What our findings mean in the context of the existing literature</i>	128
4.5.3.1	<i>Revisiting the role of the dPAG and amygdala in the expression of activity-promoting defensive behaviors</i>	129
4.5.3.2	<i>Linking VTA dopamine to the amygdala and PAG in the expression of activity-promoting defensive behaviors</i>	130
<b>Chapter 5: Summary of results and discussion</b>		134
<b>5.1</b>	<b>Summary of results</b>	135
<b>5.2</b>	<b>Discussion</b>	136
5.2.1	<i>Re: Limitations of Chapter 4</i>	136
5.2.2	<i>Temporal organization of defensive behaviors: revisiting the brain circuit for activity-suppressing and activity-promoting defensive behaviors</i>	138
5.2.3	<i>Where does the VTA fit in the fear circuit?</i>	140
5.2.4	<i>Relevance to clinical research</i>	143
<b>5.3</b>	<b>Conclusion</b>	144
<b>References</b>		145

## LIST OF TABLES

### CHAPTER 2:

Table 2.1 ANOVA results for suppression ratios during discrimination.....	29
Table 2.2 Behavior definitions.....	33
Table 2.3 ANOVA results for suppression ratios during extinction.....	48

### CHAPTER 3:

Figure 3.1 Test accuracies for CNNs.....	81
--	----

### CHAPTER 4:

Figure 4.1 ANOVA results for suppression ratios during discrimination .....	110
Figure 4.2 ANOVA results for suppression ratios during extinction .....	110

## LIST OF FIGURES

### CHAPTER 1:

Figure 1.1 Representation of Pavlovian fear conditioning.....	5
Figure 1.2 Neural circuit for fear conditioning with an auditory cue.....	8

### CHAPTER 2:

Figure 2.1 Experimental design and nose poke suppression.....	28
Figure 2.1.1 Body weight and baseline nose poke rate.....	30
Figure 2.1.2 Nose poke x discrimination.....	30
Figure 2.2 Representative behaviors.....	32
Figure 2.3 Inter-rater reliability.....	35
Figure 2.4 Temporal ethograms during discrimination.....	37
Figure 2.5 Danger-elicited behaviors.....	40
Figure 2.6 Special case of freezing.....	42
Figure 2.7 Behavior-behavior correlations.....	44
Figure 2.8 Experimental design and nose poke suppression.....	47
Figure 2.9 Temporal ethograms during extinction.....	49
Figure 2.10 Inter-rater reliability.....	50
Figure 2.11 Danger elicits locomotion during extinction.....	51
Figure 2.11.1 Freezing separated by test by order.....	53
Figure 2.12 Behavior-behavior correlations during extinction.....	55

### CHAPTER 3:

Figure 3.1 Diagram of a simple neural network .....	67
Figure 3.2 Example of a convolutional neural network.....	68
Figure 3.3 Basic CNN model.....	76

Figure 3.4 Residual block.....	77
Figure 3.5 ResNet18 model.....	78
Figure 3.6 TinyMotionNet architecture.....	80
Figure 3.7 Confusion matrix for ResNet18.....	82

**CHAPTER 4:**

Figure 4.1 Experimental design and nose poke suppression .....	109
Figure 4.2 Th+ cell counts in SN and VTA across anterior and posterior levels.....	112
Figure 4.3 Inter-rater reliability.....	114
Figure 4.4 Temporal ethograms during discrimination.....	117
Figure 4.4.1 Behavior comparisons for danger vs. safety during discrimination.....	118
Figure 4.5 Temporal ethograms during extinction.....	121
Figure 4.5.1 Behavior comparisons for danger vs. safety during extinction.....	122
Figure 4.6 Caspase-mediated deletion of VTA DA increases jumping and scaling during extinction .....	125

**CHAPTER 5:**

Figure 5.1 Danger-elicited freezing and locomotion are routinely expressed in our laboratory.....	137
Figure 5.2 Linking the VTA to the cortico-amygdalar-PAG pathway for fear learning and expression of flight and other activity-promoting defensive behaviors.....	142

## LIST OF ABBREVIATIONS

<b>BA</b>	Basal Amygdala
<b>BLA</b>	Basolateral Amygdala
<b>Casp</b>	Caspase-3
<b>CeA</b>	Central Amygdala
<b>CeL</b>	Central Amygdala, Lateral division
<b>CeM</b>	Central Amygdala, Medial division
<b>CNN</b>	Convolutional Neural Network
<b>CR</b>	Conditioned Response
<b>Cre</b>	Cre recombinase
<b>CRF</b>	Corticotropin-Releasing Factor
<b>CS</b>	Conditioned Stimulus
<b>DA</b>	Dopamine
<b>DNN</b>	Deep Neural Network
<b>DP</b>	Dorsal Peduncular
<b>dPAG</b>	Dorsal Periaqueductal Gray
<b>dIPAG</b>	Dorsolateral Periaqueductal Gray
<b>DRN</b>	Dorsal Raphe Nucleus
<b>FPS</b>	Fear-Potentiated Startle
<b>FR-1</b>	Fixed Ratio-1
<b>GABA</b>	Gamma-Aminobutyric Acid
<b>GUI</b>	Graphical User Interface
<b>HDBSCAN</b>	Hierarchical Density-Based Spatial Clustering of Applications

<b>IHC</b>	Immunohistochemistry
<b>IL</b>	Infralimbic Cortex
<b>ITI</b>	Inter-trial Interval
<b>LA</b>	Lateral Amygdala
<b>L-DOPA</b>	L-3,4-dihydroxyphenylalanine
<b>mPFC</b>	Medial Prefrontal Cortex
<b>NAc</b>	Nucleus Accumbens
<b>PAG</b>	Periaqueductal Gray
<b>PTSD</b>	Post-Traumatic Stress Disorder
<b>ResNet</b>	Residual Network
<b>SN</b>	Substantia Nigra
<b>SOM</b>	Somatostatin
<b>TH</b>	Tyrosine Hydroxylase
<b>UMAP</b>	Uniform Manifold Approximation Projection
<b>US</b>	Unconditioned Stimulus
<b>VI-30</b>	Variable Interval, 30 s
<b>VI-60</b>	Variable Interval, 60 s
<b>vPAG</b>	Ventral Periaqueductal Gray
<b>VTA</b>	Ventral Tegmental Area
<b>YFP</b>	Yellow Fluorescent Protein

## ACKNOWLEDGEMENTS

I owe many thanks to all of the people I've met throughout my journey at Boston College. This experience has shaped me into the person I am today. I am grateful for the privilege to be able to move across the country in pursuit of a doctoral degree.

First, I would like to extend my deepest gratitude to my parents. Your brave choice to immigrate to the U.S. and your countless sacrifices in order to provide better opportunities for me and my siblings have given me the drive and ambition to power me through all these years of education. Without your sacrifices, I wouldn't be where I am today. Thank you for instilling in me the importance of education and knowledge.

I would also like to thank all the friends I've made at BC, especially the graduate students in behavioral neuroscience and my cohort. I'm especially thankful for Danielle, Eliza, Rebecca, and Emily. I'm so grateful that BC brought us together. Thank you for listening to my vent sessions about all my science troubles and supporting me throughout graduate school. You all have truly been such a bright spot during dark times.

Thank you to all my loved ones back home. Christine and Nick – our unwavering friendship through life's biggest changes has been a testament of how much I value your presence in my life.

Thank you to the McDannald lab: all past and present members. Thank you Jasmin, Mahsa, and Emma for being such supportive colleagues and friends. Emma – thanks for being a great lab mate and conference buddy. Good luck with your last two years of grad school! (...it gets worse, sorry). Big thanks to all of the undergraduates who I have worked with over the years: Aleah, Nick, David, Christa, Summer, Jacob,

Anaise, and Katie. The experiments in this dissertation would not have been possible without all of you. A special thanks to Nick and Aleah – it has been such a joy to mentor and work alongside you two over the past several years (even though I'll always act like it wasn't). I am so proud of the scientists you both have become and can't wait to see the great things you'll do in the coming years.

I would like to thank all of the mentors who have had such a positive influence on my development as a scientist. Thank you to Dr. Drew Fox, who took a chance on me as an undergrad at UC Davis – you inspired me to pursue a doctoral degree and sparked my love for neuroscience research. Thank you to my committee members: Dr. Jamie Maguire, Dr. Stefano Anzellotti, Dr. Gorica Petrovich, and Dr. Mike McDannald. Stefano, words cannot express how thankful I am for all that you have contributed to my project and more importantly, to my knowledge of all things neural networks. It is a bit crazy to think that I started this project with zero knowledge about neural networks – I (and the McDannald Lab) are eternally grateful! This project will have lasting effects for the lab. Gorica, I want to thank you for being my science role model throughout graduate school. Your scientific and emotional support during these years have truly helped me through hard times.

Mike, thank you for everything. Your unrelenting support and optimism have always been a beacon of light for me during graduate school. Your presence in the lab was sorely missed while you were on sabbatical in Thailand – it made me realize how much of a positive impact your buzzing energy and enthusiasm had on me and the rest of the lab. You have shaped me into the scientist I am, and for that, I will always be so grateful.

And finally, thank you to my partner, Tyler. One of the best things about BC is that it brought me to you. Your constant love and support throughout the years has carried me through all the ups and downs of graduate school and life. As I end this chapter of my life, I am so excited to start a new one with you.

Dedicated to my parents.  
Mama and Baba, I hope I have made you proud.

**Chapter 1: Introduction to conditioned fear behaviors and its neural underpinnings in rats**

## 1.1 Pavlovian fear conditioning

Detecting and assessing threats is essential to one's survival. Importantly, appropriately responding to dangerous situations in the environment ensures self-preservation. Disruption in threat response processes is a hallmark of many stress-related disorders, particularly in Post-Traumatic Stress Disorder (PTSD) and other anxiety disorders. Therefore, studying the brain basis of threat processing and responding has been an important effort in recent decades to understand the basic neural processes of fear and anxiety, and ultimately, to facilitate the development of therapeutic strategies and treatments for those affected by anxiety disorders. Anxiety disorders is an umbrella term that encompasses a range of disorders such as PTSD, generalized anxiety disorder, and phobias. Because of the broad range of types of anxiety disorders, anxiety manifests behaviorally in very different ways depending on the type of anxiety and also on the individual. For example, while experiencing a bout of severe anxiety, one person may be physically stuck or "frozen" in place by their anxiety, while another person may be hyperactive. Both types of responses hinder their abilities to function in the present moment. Thus, when modeling fear and anxiety in animals in a laboratory setting, it is important that researchers try to capture the variation in threat behavioral responding that we experience as humans.

One way to model fear and anxiety in animal laboratory settings is via Pavlovian fear conditioning (Pavlov, 1927). In Pavlovian fear conditioning, a neutral stimulus, such as an auditory tone, is paired with an inherently aversive stimulus, such as a foot shock (unconditioned stimulus or US). Through this pairing, the neutral stimulus becomes a conditioned stimulus (CS) that acquires the ability to elicit behavioral responses

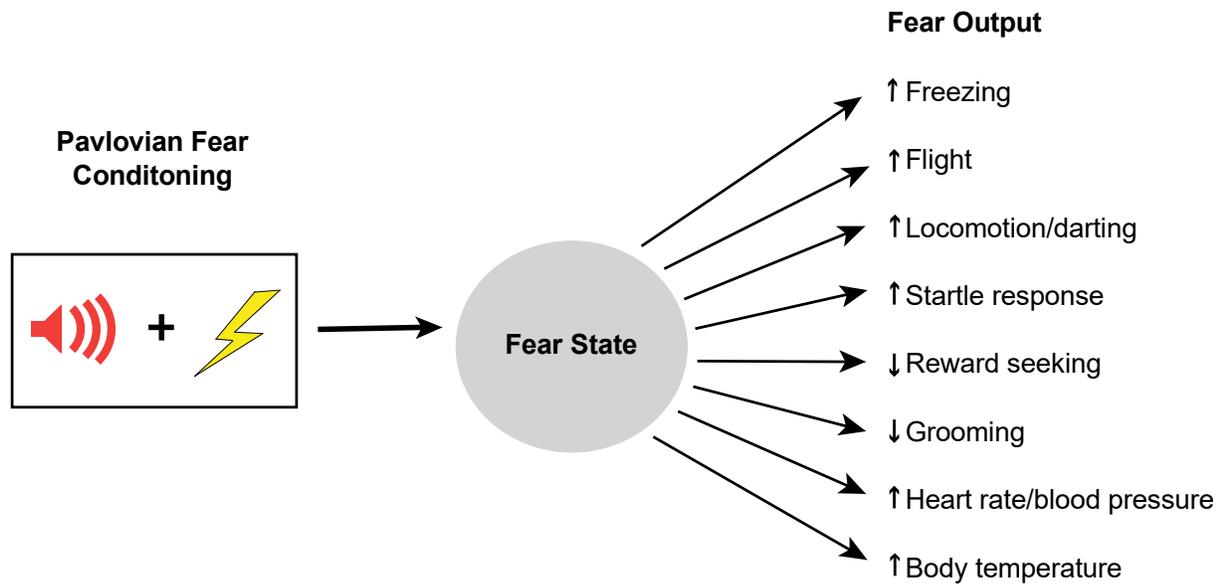
appropriate to the US (i.e. defensive behaviors). In fear conditioning, the idea is that behavioral responding evoked by the fear conditioned cue (CS) represents a state of fear in the animal (Estes and Skinner, 1941), for the presentation of the CS predicts an aversive stimulus (i.e. foot shock). For decades, neuroscience laboratories have used Pavlovian conditioning and its variations in combination with neuroscience techniques, as a way to study the neural and behavioral basis of fear in animals. Importantly, in a variety of fear conditioning settings, researchers have observed and measured different behavioral responses in rodents (Figure 1.1), which I will elaborate more on in the following sections.

## **1.2 Defensive behaviors in fear settings**

Perhaps the most ubiquitous finding in fear conditioning studies is that rodents freeze in response to a fear conditioned cue (Blanchard and Blanchard, 1969; Bolles and Collier, 1976). Freezing is thought to be part of a set of species-specific defensive reactions or SDRs (Bolles, 1970) and is an evolutionarily conserved defense mechanism to avoid predation. Because freezing is consistently observed in fear conditioning settings, it is widely used as a measure of fear in animals – so much so, that a failure to freeze during the presentation of a shock-associated cue is interpreted as a failure of the animal to learn that the cue predicts foot shock. Although freezing is a common fear behavior, it is not the only one. Bolles (1970) also identified fleeing and fighting as other SDRs. Relatedly, Bolles and Collier (1976) observed that rats locomote more and freeze less when tested in a longer box during fear conditioning, regardless if they were trained in either a short or long box. In support, flight and avoidance behaviors have been studied in fear-eliciting settings (Blanchard and

Blanchard, 1968). More recent studies have identified a conditioned, “active” response in rats called darting, which is characterized by quick locomotive movements (Gruene et al., 2015; Mitchell et al., 2022). Conditioned flight behavior has also been observed in mice during tone-white noise paired fear conditioned cues (Fadok et al., 2017; Borkar et al., 2024).

Given these findings, I was curious as to what exact behaviors rats exhibit in our laboratory’s fear discrimination setting. Prior to my first project, our laboratory had only measured conditioned nose poke suppression during discrimination (Walker et al., 2018; Wright et al., 2019; Ray et al., 2020) as suppression of reward seeking is a consequence of fear conditioning (Estes and Skinner, 1941). So in order to observe and measure multiple behaviors during fear discrimination, I devised a novel camera system in the laboratory to capture images of rats before, during, and after cue presentation. The experiments I carried out to fully investigate the behavioral patterns during discrimination and extinction are described in detail in Chapter 2.



**Figure 1.1.** Representation of Pavlovian fear conditioning

Pavlovian fear conditioning is used in behavioral neuroscience to study fear in animals. Fear conditioning (left) is thought to produce a state of fear (middle) which has physiological and behavioral consequences (right). Many fear outputs have been studied in fear conditioning.

### 1.3 The need for more behavior and its challenges

As emphasized previously, an important goal in animal fear research is to reveal the specific brain circuits, regions, and cell types that underlie fear and anxiety. Given past literature, it is apparent that there are many defensive behaviors expressed in fear settings; however, in recent years, most neuroscience laboratories use and measure freezing as an indicator for fear. This “freezing trend” in fear conditioning research is extremely limiting, since fear can manifest in a multitude of ways in both humans and rodents. In support, recent studies in mice and rats have shown that activity-promoting behaviors such as flight and darting, occur in fear conditioning settings (Gruene et al., 2015; Fadok et al. 2017; Mitchell et al., 2022). Nevertheless, the view that freezing is the only reliable conditioned fear behavior remains dominant (Kim et al., 2019). In order

to help bridge the gap between animal models and human research, it is necessary for the field of fear research to once again expand its behavioral repertoire to include other fear-related behaviors other than freezing.

Expanding our behavioral toolbox in neuroscience research comes with challenges. One pressing challenge is that measuring multiple behaviors consistently in experiments is time-consuming, and many laboratories may not have the time or labor power to execute these experiments. The recent applications of machine learning tools in behavioral neuroscience presents a promising solution to this problem. Several groups have created machine learning pipelines for use in behavioral neuroscience research, and their work has inspired my own approach in developing an automated pipeline in our laboratory. The machine learning tools I will elaborate on later in this document use convolutional neural networks (CNNs). CNNs are a class of deep neural networks that are commonly used for computer vision tasks such as face recognition and object or action classification. When used for data labeling, CNNs provide faster outputs than human scorers with similar or even better accuracy levels. Thus, CNNs are an ideal approach to automatically classify behavior in experiments that collect large amounts of videos or images of animals.

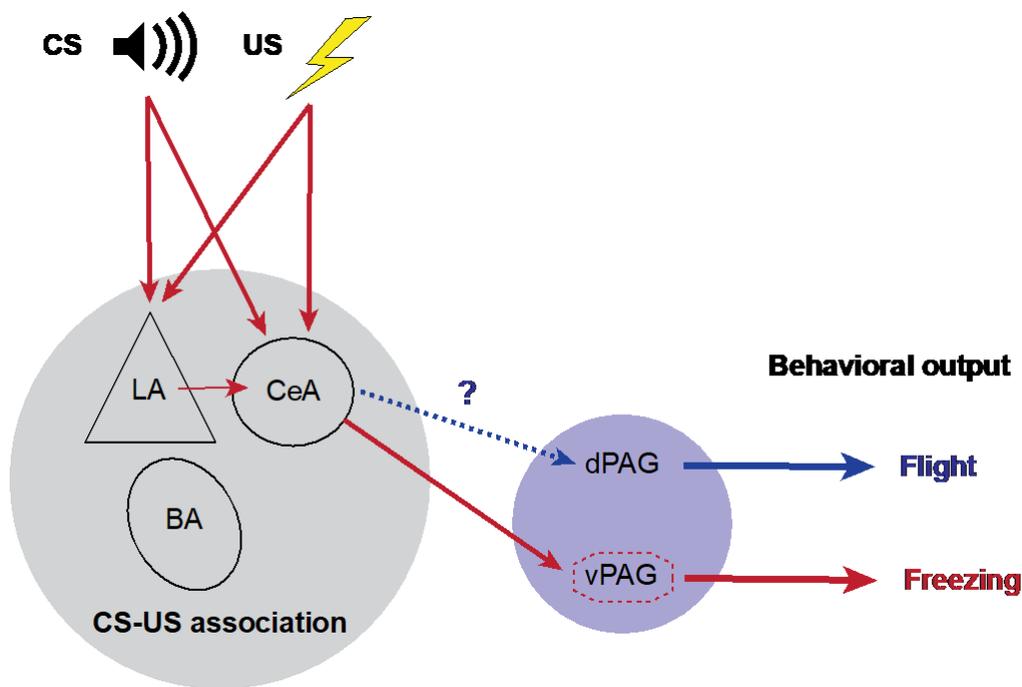
#### **1.4 Neural pathways involved in defensive behaviors: freezing and flight**

The brain regions that have been established in contributing to defensive behaviors such as freezing and flight include the periaqueductal gray (PAG), central amygdala (CeA), and lateral amygdala (LA) - nested in the larger basolateral complex of the amygdala (BLA). The ventral PAG (vPAG) is necessary for freezing behavior as

lesions to the vPAG impair freezing during fear-eliciting settings (LeDoux et al., 1988; Fanselow, 1994). In order to produce defensive responses like freezing, the PAG receives inputs from the central amygdala, which is commonly known as the “center of fear” to many (LeDoux, 2000). The neural circuit that produces conditioned freezing is well established (Figure 1.2). In fear conditioning with an auditory cue, the freezing circuit includes the thalamus which receives auditory input (CS) and relays information to LA and BLA. The LA and BLA integrate relevant, external stimuli such as an aversive foot shock (US) with the auditory cue (CS), then relays this CS-US association to the CeA, whose projections to the vPAG produce conditioned freezing (LeDoux, 2000). The more recent view is that the CeA itself can integrate relevant stimuli in both fear and appetitive settings (Kong and Zweifel, 2021). Further, studies lesioning the LA, BLA, and CeA demonstrate the necessity of the amygdala in producing freezing during both auditory and contextual fear conditioning (Goosens and Maren, 2001; Maren and Fanselow, 1996; Kim et al., 1993; Phillips and LeDoux, 1992).

Additionally, it has been shown that the dorsal PAG (dPAG) is involved in flight behavior (Fanselow, 1994; Morgan et al., 1998; Deng et al., 2016). Researchers have lesioned the dPAG and found reduced flight behavior in both fear conditioning and naturalistic threat (unconditioned fear) settings. However, the full brain circuit for producing flight behavior remains unclear. A recent study proposes a novel pathway from the dorsal peduncular subdivision (DP) of the medial prefrontal cortex (mPFC) to the medial subdivision of the central amygdala (CeM) to be responsible for producing flight in mice during a high-threat fear conditioning paradigm (Borkar et al., 2024). This study demonstrates that the projections from the CeM to the dPAG and dorsolateral

PAG (dPAG) are the likely candidate for completing the flight pathway. Although, altogether, these studies implicate the mPFC, the amygdala, and dPAG in producing flight, there are not many studies directly manipulating the projections to the dPAG that are necessary and sufficient to produce fear conditioned flight behavior. In general, pathway-specific manipulations for flight behavior in fear conditioning settings are far and few.



**Figure 1.2.** Neural circuit for fear conditioning with an auditory cue

During fear conditioning, multiple pairings of CS-US produce conditioned defensive behaviors (i.e. freezing and flight). Information about the auditory cue (CS) and foot shock (US) are integrated in the lateral amygdala (LA) and there is evidence the central amygdala (CeA) is also capable of CS-US integration. Outputs from the amygdala to downstream targets like the PAG produce defensive responses. The circuit for freezing (dark red arrows) is well established, amygdalar inputs to the vPAG produces conditioned freezing in animals. The neural circuit for conditioned flight (dark blue arrows) is less well known.

There is a body of work that implicates the ventral tegmental area (VTA) and dopamine (DA) in fear learning (Nader and LeDoux, 1999; Fadok et al., 2009; Tang et al., 2020; Jo et al., 2018), fear extinction (Luo et al., 2018; Salinas-Hernández et al., 2023), safety learning (Yau and McNally, 2022), and the expression of activity-promoting fear behaviors (Borowski and Kikkinidis, 1996; Reis et al, 2004; Ribeiro de Oliveira et al., 2008). I will discuss in detail the literature implicating VTA dopamine in flight as well as other activity-promoting fear behaviors (like avoidance and jumping) during various fear settings in Chapter 4. These studies demonstrate that VTA dopaminergic cells likely play an important modulatory role in the brain circuitry responsible for producing activity-promoting fear behaviors in both conditioned and unconditioned fear settings. Given this literature, I conducted an experiment to determine whether VTA dopaminergic neurons contribute to the expression of the defensive behaviors we observe during fear discrimination.

## **1.5 Dissertation aims**

The aims of this dissertation are three-fold: first, to demonstrate that a fear conditioned cue elicits a variety of defensive behaviors in our laboratory's behavioral setting (Chapter 2), second, to develop a machine learning pipeline to automate behavior scoring for future research projects (Chapter 3), and lastly, to demonstrate a role of ventral tegmental area dopaminergic neurons in the expression of fear conditioned defensive behaviors (Chapter 4).

In Chapter 2, the first experiment I conducted observed the behavioral patterns of rats responding to 3 auditory cues that predicted foot shock with different probabilities: a

safety cue ( $p=0.00$ ), an uncertainty cue ( $p=0.25$ ), and a danger cue ( $p=1.00$ ), over the course of fear discrimination. Prior to this project, our laboratory had used only nose poke suppression as a measure of fear (Walker et al., 2018; Wright et al., 2019; Ray et al., 2020). With the installation of cameras and a novel recording system, we were able to capture behavioral data at a sub-second level for the first time in our laboratory's history. We hypothesized that freezing would be the dominant behavior expressed during the danger cue as conditioned suppression and freezing are highly correlated (Bouton and Bolles, 1980). Instead, we found that a fear conditioned cue (danger cue) not only elicits freezing, but also a variety of behaviors including locomotion, jumping, and rearing. In a following experiment, I was interested in whether the danger cue would evoke the activity-promoting behaviors (locomotion, jumping, and rearing) observed from Experiment 1, in an extinction setting where foot shocks are absent. We found that the danger cue evoked timed locomotion during extinction with and without the reward apparatus present.

In Chapter 3, I will elaborate on the convolutional neural networks (CNNs) we developed and trained to automatically quantify behavior. The hand-scored behavior frames from the first experiment in Chapter 2 were used for training the CNNs described in Chapter 3. In this chapter, I will describe different machine learning tools that currently exist for application in behavioral neuroscience research and directly compare our approach to existing tools. Additionally, I will describe the overall architecture of the CNNs we developed and compare their performance with our human raters.

Lastly, I was interested in whether dopaminergic neurons in the VTA contributed to the expression of the fear conditioned, activity-promoting defensive behaviors we

observed in Chapter 2. The experiments I conducted to answer this question are described in Chapter 4, where I selectively deleted dopaminergic cells in the VTA prior to fear discrimination, and quantified behavior during the last session of discrimination (when discrimination was complete) as well as during extinction (with reward present). The results suggest that deleting dopaminergic cells in the VTA potentially promotes danger cue-specific jumping and scaling during extinction, although there are many considerations to this finding.

In the final chapter, I will summarize the main findings from my experiments and discuss the results in the context of previous findings. I will explain the limitations of the current experiments I conducted, and propose future research that is necessary to paint a complete picture of defensive behaviors in fear conditioning. The work presented here aims to contribute to our overall understanding of fear behavior.

## **Chapter 2: A fear conditioned cue orchestrates a suite of behaviors in rats**

*This chapter was published in the following research article:*

Amanda Chu, Nicholas T Gordon, Aleah M DuBois, Christa B Michel, Katherine E Hanrahan, David C Williams, Stefano Anzellotti, Michael A McDannald (2024) A fear conditioned cue orchestrates a suite of behaviors in rats. eLife 13:e82497

## 2.1 Introduction

### 2.1.1 *Pavlovian fear conditioning and defensive behaviors*

Animals evolved defensive systems to detect and avoid predation. The predatory imminence continuum (PIC), a prominent theory of defensive behavior, identifies three defensive modes based on the proximity to predation: pre-encounter (leaving the safety of the nest), post-encounter (predator detected), and circa-strike (predation imminent or occurring) (Fanselow and Lester, 1988). Pavlovian fear conditioning has been extensively used to reveal the behavioral and neural underpinnings of defensive systems in rats (Bolles and Collier, 1976; Fanselow, 1993; Killcross et al., 1997; McNally et al., 2011). In a typical Pavlovian fear conditioning procedure, a rat is placed in a neutral context, and played an auditory cue whose termination coincides with foot shock delivery. Each PIC mode is characterized by a unique set of behaviors and, critically, each mode is thought to be captured by a unique epoch of a Pavlovian fear conditioning trial (Fanselow et al., 2019). The post-encounter mode is characterized by freezing, and is captured by cue presentation. Circa-strike is characterized by locomotion, jumping, and rearing, and is captured by shock delivery.

Freezing to a fear conditioned cue may be the most ubiquitous finding in all of behavioral neuroscience (Blanchard and Blanchard, 1969; Bolles and Collier, 1976; Maren et al., 1997; Anagnostaras et al., 1999; Wilensky et al., 1999; Quirk, 2002; Koo et al., 2004; Rogers and Kesner, 2004; Iordanova et al., 2006; Shumake et al., 2014; Foilb et al., 2016; Furlong et al., 2016). The relationship between freezing and Pavlovian fear conditioning is so strong that failing to observe freezing in defensive settings has been used to support assertions that Pavlovian fear conditioning did not occur (Zambetti et

al., 2021). Cued fear as freezing has been further entrenched by historical observations that locomotion, jumping, and rearing (theorized circa-strike behaviors) are not elicited by fear conditioned cues (Fanselow et al., 2019). Instead, activity-promoting defensive behaviors are restricted to shock delivery (Fanselow, 1982) or to other sudden changes in stimuli (Fadok et al., 2017; Totty et al., 2021). Yet, locomotion, jumping, and rearing all readily occur in defensive settings (Blanchard et al., 1986; Holland, 1979; Dielenberg and McGregor, 2001). Most relevant, a fear conditioned cue can elicit locomotion, rapid forward movements termed “darting” (Gruene et al., 2015; Mitchell et al., 2022).

The ability of a fear conditioned cue to elicit locomotion has been called into question (Trott et al., 2022). Trott et al. noted that in prior studies locomotion was greatest at cue onset – the time point most distal from shock delivery (Gruene et al., 2015; Fadok et al., 2017). Moreover, prior studies did not use associative controls (but see Totty et al., 2021) – essential to making claims that cue-elicited behaviors were due to a predictive relationship with foot shock. Using between-subjects designs in mice, Trott et al. ascribe the majority of cue-elicited locomotion to non-associative cue properties. The foundational study demonstrating the need for proper associative controls in *any* form of conditioning used Pavlovian fear conditioning (Rescorla, 1967). Not just all-or-none, the magnitude of a fear conditioned, cue-elicited response can scale with foot shock probability (Rescorla, 1968; Ray et al., 2020). Rescorla 1968, and many foundational associative learning studies (Kamin, 1969; Rescorla and Wagner, 1972), relied on experiments that did not measure ‘fear’ with freezing, but with suppression of operant responding for reward (now termed conditioned suppression) (Estes and Skinner, 1941). Drawing from Rescorla 1968, our laboratory has devised a

robust, within-subjects Pavlovian fear conditioning procedure in which three cues predict unique foot shock probabilities: danger ( $p=1$ ), uncertainty ( $p=0.25$ ), and safety ( $p=0$ ). Measuring conditioned suppression, we consistently observe complete behavioral discrimination: danger elicits greater suppression than safety, and uncertainty elicits suppression intermediate to danger and safety (Wright et al., 2015; DiLeo et al., 2016; Walker et al., 2018; Ray et al., 2022).

### 2.1.2 Chapter Aims

The goal of Experiment 1 was to construct comprehensive, temporal ethograms of rat behavior during discriminative Pavlovian fear conditioning, consisting of a danger, uncertainty, and safety cue. This would allow us to determine what behaviors come under the control of a fear conditioned cue, and how these behaviors are temporally organized. We had the ability to reveal freezing as the exclusive conditioned behavior, as prior studies have found positive relationships between conditioned freezing and conditioned suppression (Bouton and Bolles, 1980; Mast et al., 1982). Yet, we also had the ability to detect additional behaviors, as brain manipulations that impair conditioned freezing can have little or no impact on conditioned suppression (McDannald, 2010; McDannald and Galarce, 2011). A sub-goal was to compare behaviors elicited by the deterministic danger cue, and the probabilistic uncertainty cue. The goal of Experiment 2 was to reveal which of these danger-elicited behaviors transferred to an extinction context in which shock and reward were not present. For Experiment 2, we simplified the discrimination procedure to include only the danger and safety cues.

Twenty-four rats (12 female; Experiment 1) and sixteen rats (8 female, Experiment 2) received Pavlovian fear discrimination. TTL-triggered GigE cameras were

installed in behavioral boxes and programmed to capture frames at sub-second temporal resolution prior to and during cue presentation. 86,400 frames (Experiment 1) and 25,600 frames (Experiment 2) were hand scored for nine discrete behaviors reflecting reward (Holland, 1977), activity-suppressing fear (Blanchard and Blanchard, 1969; Fanselow, 1982), and activity-promoting fear (Blanchard et al., 1986; Dielenberg and McGregor, 2001; Gruene et al., 2015). Complete temporal ethograms were constructed during early, middle, and late conditioning sessions (Experiment 1), and for the two types of extinction tests (Experiment 2). Danger responding was compared to baseline and to safety, which served as an unpaired control cue. Behaviors elicited by the danger cue were considered associative (due to pairing with foot shock) if they differed both from baseline and from the safety cue. The temporal profile of responding was determined by tracking behavior change over cue presentation.

## 2.2 Materials and Methods

### 2.2.1 Subjects

For Experiment 1, twenty-four adult Long Evans rats (12 female) weighing 196-298g arrived from Charles River Laboratories on postnatal day 55. Rats were single-housed on a 12-hr light cycle (lights off at 6:00pm) and maintained at their initial body weight with standard laboratory chow (18% Protein Rodent Diet #2018, Harlan Teklad Global Diets, Madison, WI). Water was available *ad libitum* in the home cage. For Experiment 2, sixteen adult Long Evans rats (8 female) were housed and maintained as described above. All protocols were approved by the Boston College Animal Care and Use Committee and all experiments were carried out in accordance with the NIH guidelines regarding the care and use of rats for experimental procedures.

### 2.2.2 Behavior apparatus

The apparatus for Pavlovian fear discrimination consisted of four individual chambers with aluminum front and back walls, clear acrylic sides and top, and a grid floor. LED strips emitting 940 nm light were affixed to the acrylic top to illuminate the behavioral chamber for frame capture. 940 nm illumination was chosen because rats do not detect light wavelengths exceeding 930 nm (Nikbakht and Diamond, 2021). Each grid floor bar was electrically connected to an aversive shock generator (Med Associates, St. Albans, VT). An external food cup, and a central port equipped with infrared photocells were present on one wall. Auditory stimuli were generated with an Arduino-based device and presented through two speakers mounted on the ceiling.

### *2.2.3 Pellet exposure and nose poke shaping*

Rats were food-restricted and specifically fed to maintain their body weight throughout behavioral testing. Each rat was given four grams of experimental pellets in their home cage in order to overcome neophobia. Next, the central port was removed from the experimental chamber, and rats received a 30-minute session in which one pellet was delivered every minute. The central port was returned to the experimental chamber for the remainder of behavioral testing. Each rat was then shaped to nose poke in the central port for experimental pellet delivery using a fixed ratio schedule in which one nose poke into the port yielded one pellet. Shaping sessions lasted 30 min or until approximately 50 nose pokes were completed. Each rat then received 6 sessions during which nose pokes into the port were reinforced on a variable interval schedule. Session 1 used a variable interval 30 s schedule (poking into the port was reinforced every 30 s on average). All remaining sessions used a variable interval 60 s schedule. For the remainder of behavioral testing, nose pokes were reinforced on a variable interval 60 s schedule independent of cue and shock presentation.

### *2.2.4 Cue pre-exposure*

Each rat was pre-exposed to the three cues to be used in Pavlovian discrimination in one session. Auditory cues consisted of repeating motifs of broadband click, phaser, or trumpet. This 37 min session consisted of four presentations of each cue (12 total presentations) with a mean inter-trial interval (ITI) of 2.5 min. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

### 2.2.5 *Pavlovian fear discrimination*

#### Experiment 1

Each rat received sixteen, 48-minute sessions of fear discrimination. Each session consisted of 16 trials, with a mean ITI of 2.5 min. Auditory cues were 10 s in duration. Each cue was associated with a unique foot shock probability (0.5 mA, 0.5 s): danger,  $p=1.00$ ; uncertainty,  $p=0.25$ ; and safety,  $p=0.00$ . Foot shock was administered 2 s following the termination of the auditory cue on danger and uncertainty-shock trials. Auditory identity was counterbalanced across rats. Each session consisted of four danger trials, two uncertainty-shock trials, six uncertainty-omission trials, and four safety trials. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

#### Experiment 2

Each rat received twelve, 48-minute sessions of fear discrimination. Each session consisted of 8 trials, with a mean ITI of 3.5 min. Auditory cues were 10 s in duration. Each cue was associated with a unique foot shock probability (0.5 mA, 0.5 s): danger,  $p=1.00$  and safety,  $p=0.00$ . Foot shock was administered 2 s following the termination of the auditory cue on danger trials. Auditory identity was counterbalanced across rats. Each session consisted of four danger trials and four safety trials. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

### 2.2.6 *Fear extinction*

For Experiment 2, each rat received two types of extinction test: one with the reward apparatus present and one with the reward apparatus absent. Test type was

counterbalanced across rats. Extinction sessions were 48 minutes in duration, and consisted of 8 trials, with a mean ITI of 3.5 min. Auditory cues were 10 s in duration. Foot shocks were not delivered after danger cue termination. Auditory identity was counterbalanced across rats. Each session consisted of four danger trials and four safety trials. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

### *2.2.7 Calculating suppression ratio*

Time stamps for cue presentations, shock delivery, and nose pokes (photobeam break) were automatically recorded by the Med Associates program. Baseline nose poke rate was calculated for each trial by counting the number of pokes during the 20-s pre-cue period and multiplying by 3. Cue nose poke rate was calculated for each trial by counting the number of pokes during the 10-s cue period and multiplying by 6. Nose poke suppression was calculated as a ratio:  $(\text{baseline poke rate} - \text{cue poke rate}) / (\text{baseline poke rate} + \text{cue poke rate})$ . A suppression ratio of '1' indicated complete suppression of nose poking during cue presentation relative to baseline. A suppression ratio of indicated '0' indicates equivalent nose poke rates during baseline and cue presentation. Gradations in suppression ratio between 1 and 0 indicated intermediate levels of nose poke suppression during cue presentation relative to baseline. Negative suppression ratios indicated increased nose poke rates during cue presentation relative to baseline.

### *2.2.8 Frame capture system*

Behavior frames were captured using Imaging Source monochrome cameras (DMK 37BUX28; USB 3.1, 1/2.9" Sony Pregius IMX287, global shutter, resolution 720x540, trigger in, digital out, C/CS-mount). Frame capture was triggered by the Med Associates behavior program. The 28V Med Associates pulse was converted to a 5V TTL pulse via Adapter (SG-231, Med Associates, St. Albans, VT). The TTL adapter was wired to the camera's trigger input. Captured frames were saved to a PC (OptiPlex 7470 All-in-One) running IC Capture software (Imaging Source). For Experiment 1, frame capture began precisely 5 s before cue onset and continued throughout 10-s cue presentation. Frames were captured at a rate of 5 per second, with a target of capturing 75 frames per trial ( $5 \text{ frames/s} \times 15 \text{ s} = 75 \text{ frames}$ ), and 1200 frames per session ( $75 \text{ frames/trial} \times 16 \text{ trials} = 1200 \text{ frames}$ ). For Experiment 2, frame capture began 5 s before cue onset and continued throughout 10-s cue presentation and 5 s after cue termination. Frames were captured at a rate of 5 per second, with a target of capturing 100 frames per trial ( $5 \text{ frame/s} \times 20 \text{ s} = 100 \text{ frames}$ ), and 800 frames per session ( $100 \text{ frames/trial} \times 8 \text{ trials} = 800 \text{ frames}$ ).

### *2.2.9 Post-acquisition frame processing*

#### Experiment 1

We aimed to capture 1200 frames per session, and selected sessions 2, 8, and 16 for hand scoring. A Matlab script sorted the 1200 frames into 16 folders, one for each trial, each containing 75 frames. Each 75-frame trial was made into a 75-slide PowerPoint presentation to be used for hand scoring.

## Experiment 2

We aimed to capture 800 frames per session, and selected extinction session 1 and 2 for hand scoring. A Matlab script sorted the 800 frames into 8 folders, one for each trial, each containing 100 frames. Each 100-frame trial was made into a 100-slide PowerPoint presentation to be used for hand scoring.

### *2.2.10 Anonymizing trial information*

For Experiment 1, a total of 1,152 trials of behavior were scored from the 24 rats over the 3 sessions of discrimination (16 trials per session). For Experiment 2, a total of 256 trials were scored from 16 rats over the 2 extinction sessions (8 trials per session). We anonymized trial information in order to score behavior without bias. The numerical information from each trial (session #, rat # and trial #) was encrypted as a unique number sequence. A unique word was then added to the front of this sequence. The result was that each of the trials was converted into a unique word+number sequence. For example, trial ac01\_02\_07 (rat #1, session #2 and trial #7) would be encrypted as: abundant28515581. The trials from Experiment 1 were randomly assigned to 5 observers. 256 trials from Experiment 2 were randomly assigned to 7 observers. The result of trial anonymization was that observers were completely blind to subject, trial type, and session number. Further, random assignment meant that the 16 or 8 trials composing a single session were scored by different observers.

### *2.2.11 Behavior categories and definitions*

Frames were scored as one of ten mutually exclusive behavior categories, defined as follows:

Background. Specific behavior cannot be discerned because the rat is turned away from the camera or position of forepaws is not clear, or because the rat is not engaged in any of the other behaviors.

Cup. Any part of the nose above the food cup but below the nose port.

Freeze. Arched back and stiff, rigid posture in the absence of movement, all four limbs on the floor (often accompanied by hyperventilation and piloerection). Side to side head movements and up and down head movements that do not disturb rigid posture are permitted. Activity such as sniffing or investigation of the bars is not freezing. Freezing, as opposed to pausing, is likely to be 3 or more frames (600+ ms) long.

Groom. Any scratching, licking, or washing of the body.

Jump. All four limbs off the floor. Includes hanging which is distinguished when hind legs are hanging freely.

Locomote. Propelling body across chamber on all four feet, as defined by movement of back feet. Movement of back feet with front feet off the floor is rearing.

Port. Any part of the nose in the port. Often standing still in front of the port but sometimes tilting head sideways with the body off to the side of the port.

Rear. One or two hind legs on the grid floor with both forepaws off the grid floor and not on the food cup. Usually (not always) stretching to full extent, forepaws usually (not always) on top of side walls of the chamber, often pawing walls; may be accompanied by sniffing or slow side-to-side movement of head. Does not include grooming movements or eating, even if performed while standing on hind legs.

Scale. All four limbs off the floor but at least two limbs on the side of the chamber.

Standing on the food cup counts as scaling.

Stretch. Body is elongated with the back posture 'flatter' than normal. Stretching is often accompanied by immobility, like freezing, but is distinguished by the shape of the back.

### *2.2.12 Frame scoring system*

Frames were scored using a specific procedure. Frames were first watched in real time in Microsoft PowerPoint by setting the slide duration and transition to 0.19 s, then playing as a slideshow. Behaviors clearly observed were noted. Next, the observer went through all the frames scoring one behavior at a time. A standard scoring sequence was used: port, cup, rear, scale, jump, groom, freeze, locomote, and stretch. When the specific behavior was observed in a frame, that frame was labeled. Once all behaviors had been scored, the video was re-watched for freezing. The unlabeled frames were then labeled 'background'. Finally, all background frames were checked to ensure they did not contain a defined behavior.

### *2.2.13 Inter-observer reliability*

#### Experiment 1

To assess inter-observer reliability, we selected 12 trials from outside session 2, 8, and 16, six from females and six from males. Each of our five observers scored these 12 trials, interweaving the 12 comparison trials with the primary data trials. As a result, each observer scored 900 comparison frames which were then used to assess inter-observer reliability.

## Experiment 2

Inter-observer reliability was assessed as described in Experiment 1. 8 trials from outside extinction session 1 and 2 were selected for comparison. Each observer scored 800 comparison frames which were then used to assess inter-observer reliability.

### *2.2.14 Statistical analyses*

Analysis of variance (ANOVA) was performed for body weight, baseline nose poke rate, suppression ratios, and specific behaviors. Sex was used as a factor for all analyses. Cue, session, and time were used as factors when relevant. Univariate ANOVA following MANOVA used a Bonferroni-corrected p value significance of 0.0055 (0.05/9) to account for the nine quantified behaviors. Multiple analysis of variance (MANOVA) was performed for the nine quantified behaviors with factors of sex, cue, and time. Pearson's correlation coefficient was used to examine the relationship between baseline nose poke rate and body weight, baseline nose poke rate and cue discrimination, as well as the relationship between danger cue-elicited behaviors during early and late cue presentation in session. Within-subject comparisons were made using 95% bootstrap confidence intervals with the Matlab bootci function. Comparisons were said to differ when the 95% bootstrap confidence interval did not contain zero. Between subject's comparisons were made using independent samples t-test.

## 2.3 Results

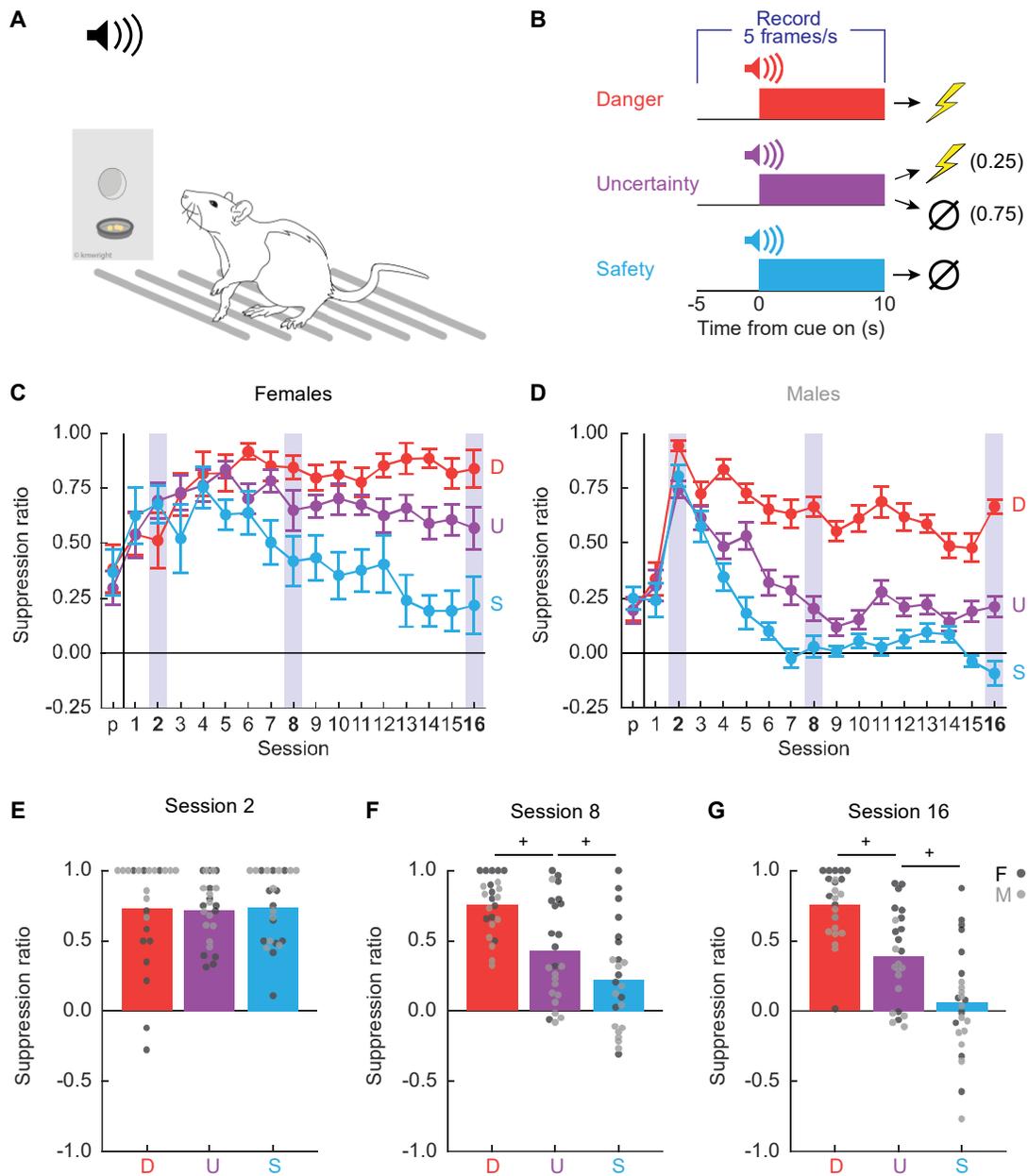
### 2.3.1 Experiment 1

#### 2.3.1.1 Conditioned suppression reveals complete discrimination

Twenty-four Long Evans rats (12 females) were trained to nose poke in a central port for food reward. Nose poking was reinforced on a 60-s variable interval schedule throughout behavioral testing. Independent of the poke-food contingency, auditory cues were played through overhead speakers, and foot shock delivered through the grid floor (Figure 2.1A). The experimental design consisted of three cues predicting unique foot shock probabilities: danger ( $p=1$ ), uncertainty ( $p=0.25$ ), and safety ( $p=0$ ) (Figure 2.1B). Behavior chambers were equipped with TTL-triggered cameras capturing 5 frames/s starting 5 s prior to cue presentation and continuing throughout the 10-s cue. TTL-triggered capture yielded 75 frames per trial, and 1200 frames per session. We aimed to capture 28,800 frames each session (1200 frames x 24 rats).

Our laboratory routinely observes complete behavioral discrimination between danger, uncertainty, and safety in female and male rats measuring conditioned suppression (Walker et al., 2018; Wright et al., 2019; Ray et al., 2022). Suppression ratios are calculated using baseline and cue nose poke rates:  $(\text{baseline} - \text{cue}) / (\text{baseline} + \text{cue})$ . Suppression ratios provide a continuous behavior measure, from no suppression (ratio = 0) to total suppression (ratio = 1). To determine if we observed complete behavioral discrimination in these 24 rats, we performed ANOVA for suppression ratios [factors: cue (danger vs. uncertainty vs. safety), session (17 total: 1 pre-exposure and 16 discrimination), and sex (female vs. male)]. Complete behavioral discrimination emerged over testing (Figure 2.1C, D). ANOVA found a significant main

effect of cue and a significant cue x session interaction ( $F_s > 6$ ,  $p_s < 0.0001$ ; see Table 2.1 for specific values). Sex effects were apparent; ANOVA found a significant main effect of sex, as well as a significant cue x sex interaction and a cue x session x sex interaction ( $F_s > 3$ ,  $p_s < 0.05$ ; Table 2.1). Female suppression ratios were higher to each cue across all discrimination sessions: danger ( $t_{22} = 3.36$ ,  $p = 0.003$ ), uncertainty ( $t_{22} = 7.14$ ,  $p = 3.67 \times 10^{-7}$ ), and safety ( $t_{22} = 4.40$ ,  $p = 0.0002$ ).



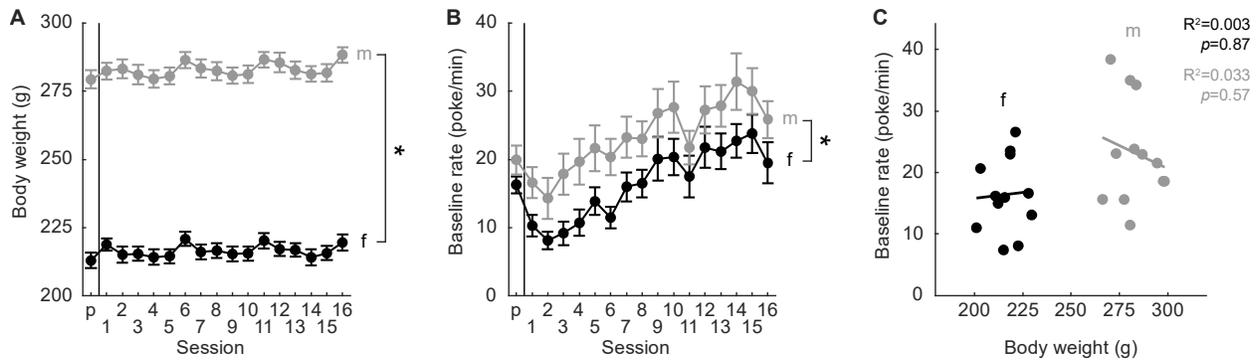
**Figure 2.1.** Experimental design and nose poke suppression

**(A)** Conditioned suppression procedure during which rats nose poke for food, while cues are played overhead and shocks delivered through floor. **(B)** Fear discrimination consisted of 10-s auditory cues predicting unique foot shock probabilities: danger (red;  $p=1$ ), uncertainty (purple;  $p=0.25$ ), safety (blue;  $p=0$ ). Five video frames were captured per second, starting 5-s prior to cue onset and continuing through cue presentation. Mean  $\pm$  SEM suppression ratios for danger (red), uncertainty (purple), and safety (blue) from pre-exposure through discrimination session 16 are shown for **(C)** females, and **(D)** males. Mean + individual suppression ratios for each cue are shown for **(E)** session 2, **(F)** session 8, and **(G)** session 16. Individuals represented by black (female) and gray (male) dots. +95% bootstrap confidence interval does not contain zero.

	Suppression ratio	
Cue	<b><math>F_{(2,44)} = 58.44</math></b>	<b><math>p = 4.10 \times 10^{-13}</math></b>
Cue x sex	<b><math>F_{(2,44)} = 3.62</math></b>	<b><math>p = 0.035</math></b>
Session	<b><math>F_{(16,352)} = 12.12</math></b>	<b><math>p = 2.63 \times 10^{-25}</math></b>
Session x sex	<b><math>F_{(16,352)} = 5.26</math></b>	<b><math>p = 6.05 \times 10^{-10}</math></b>
Cue x session	<b><math>F_{(32,704)} = 6.58</math></b>	<b><math>p = 5.51 \times 10^{-24}</math></b>
Cue x session x sex	<b><math>F_{(32,704)} = 1.68</math></b>	<b><math>p = 0.012</math></b>
Sex	<b><math>F_{(1,22)} = 59.08</math></b>	<b><math>p = 1.14 \times 10^{-7}</math></b>

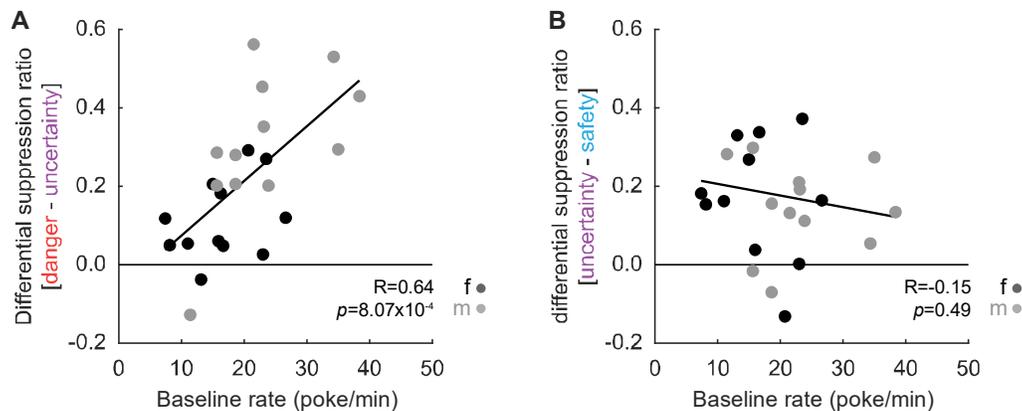
**Table 2.1.** Experiment 1 ANOVA results for suppression ratio. Significant main effects and interactions are bolded.

Sex differences in body weight and baseline nose poke rate existed prior to and throughout discrimination, with males weighing more and poking more than females (Figure 2.1.1). It is therefore possible that sex indirectly moderates conditioned suppression through effects on body weight or baseline nose poke rate. To determine this, we performed analysis of covariance (ANCOVA) for suppression ratios [factors: cue (danger vs. uncertainty vs. safety) and session (17 total: 1 pre-exposure and 16 discrimination)] using body weight or baseline nose poke rate as the covariate. ANCOVA with body weight found neither a significant body weight x cue interaction ( $F_{(2,44)} = 2.97, p=0.062$ ) nor a significant body weight x cue x session interaction ( $F_{(32,704)} = 1.40, p=0.074$ ). However, ANCOVA with baseline nose poke rate found a significant baseline x cue interaction ( $F_{(2,44)} = 5.49, p=0.007$ ) but not a significant baseline x cue x session interaction ( $F_{(32,704)} = 0.79, p=0.79$ ). Irrespective of sex, higher baseline nose poke rates predicted greater discrimination of danger and uncertainty (Figure 2.1.2).



**Figure 2.1.1.** Body weight and baseline nose poke rate

**(A)** Analysis of variance (ANOVA) for body weight (g) [factors: sex and session] revealed significant main effects of session ( $F(16,352) = 29.58$ ,  $p = 1.90 \times 10^{-55}$ ), sex ( $F(1,22) = 287.54$ ,  $p = 4.07 \times 10^{-14}$ ), and a significant session  $\times$  sex interaction ( $F(16,352) = 2.20$ ,  $p = 0.005$ ). Mean  $\pm$  SEM body weights in grams (y-axis) of males (gray) and females (black) from pre-exposure through session 16. **(B)** Baseline nose poke rates (poke/min) decreased during discrimination sessions 1 and 2, then increased over the remaining sessions. Males poked at higher baseline levels across all sessions. Analysis of variance (ANOVA) for baseline nose poke rate (poke/min) [factors: sex and session] revealed significant main effects of session ( $F(16,352) = 19.30$ ,  $p = 4.44 \times 10^{-39}$ ) and sex ( $F(1,22) = 5.10$ ,  $p = 0.034$ ). Mean baseline pose rate (y-axis) of males and females from sessions 1-16. **(C)** Baseline nose poke rate plotted against body weight for all individuals. There was no relationship between the two measures in either female or male rats. \*paired samples t-test  $p < 0.05$ .

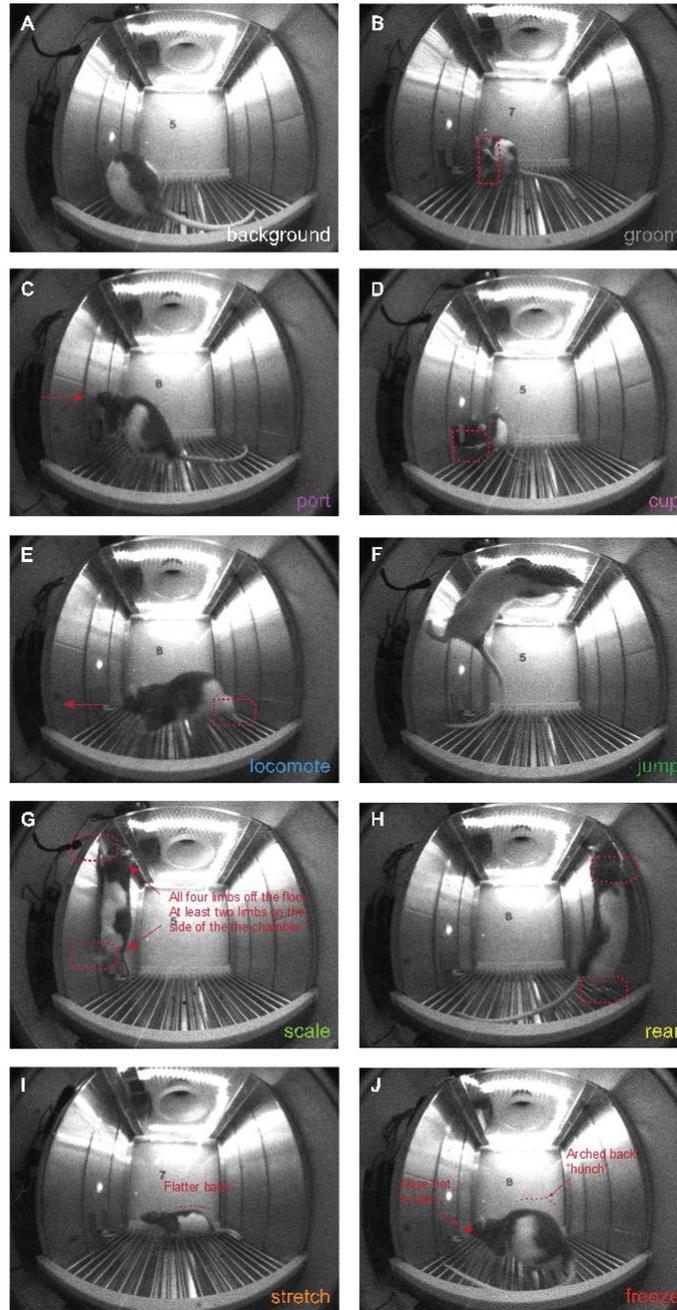


**Figure 2.1.2.** Nose poke  $\times$  discrimination

Correlations between baseline nose poke rate and differential suppression ratios for **(A)** danger and uncertainty, and **(B)** uncertainty and safety are shown. Individuals represented by black (female) and gray (male) circles. R and p values from Pearson's correlation coefficient.

Constructing behavioral ethograms for all 16 discrimination sessions would have required hand scoring 460,800 frames. To make scoring feasible and capture the emergence of discrimination, we selected sessions 2, 8, and 16. Suppression generalized to all cues during Session 2 (Figure 2.1E). Behavioral discrimination emerged by session 8 (Figure 2.1F), and was at its most complete during session 16 (Figure 2.1G). Patterns were confirmed with 95% bootstrap confidence intervals (BCIs) which found no suppression ratio differences for any cue pair during session 2 (all 95% BCIs contained zero), but differences between all cue pairs during sessions 8 and 16 (no 95% BCIs contained zero).

Frames were hand scored for nine discrete behaviors: cup, freezing, grooming, jumping, locomotion, port, rearing, scaling, and stretching, plus “background” (Definitions in Table 2.2). Behavior categories and their definitions were based on prior work in appetitive conditioning (Holland, 1977), foot shock conditioning (Fanselow, 1982; Blanchard et al., 1986), as well as our own observations. Representative behavior frames are shown in Figure 2.2. Videos 1-4 show example danger trials for four different rats (females in Videos 1 & 3, males in Videos 2 & 4).



**Figure 2.2.** Representative behaviors

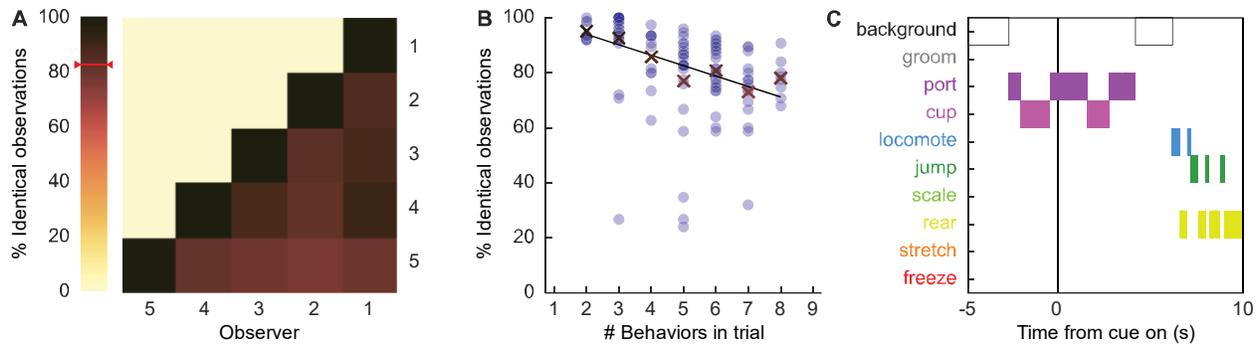
Representatives frames are shown for: **(A)** background, **(B)** groom, **(C)** port, **(D)** cup, **(E)** locomote, **(F)** jump, **(G)** scale, **(H)** rear, **(I)** stretch, and **(J)** freeze.

Behavior (in alphabetical order)	Definition
Background	Specific behavior cannot be discerned because the rat is turned away from the camera or position of forepaws is not clear, or because the rat is not engaged in any of the other behaviors
Cup	Any part of the nose is above the food cup but below the nose port.
Freeze	Arched back and stiff, rigid posture in the absence of movement, all four limbs on the floor (often accompanied by hyperventilation and piloerection). Side to side head movements and up and down head movements that do not disturb rigid posture are permitted. Activity such as sniffing or investigation of the bars is not freezing. Freezing, as opposed to pausing, is likely to be 3 or more frames in duration (600+ ms).
Groom	Any scratching, licking, or washing of the body.
Jump	All four limbs off the floor. Includes hanging which is distinguished when hind legs are hanging freely.
Locomote	Propelling the body across the chamber on all four feet, as defined by movement of back feet. Movement of back feet with front feet off the floor is rearing.
Port	Any part of the nose in the port or just outside the port. The nose blocks the infrared light in the port. Often standing still in front of the port but sometimes tilting head sideways with the body off to the side of the port.
Rear	One or two hind legs on the grid floor with both forepaws off the grid floor and not on the food cup. Usually (not always) stretching to full extent, forepaws usually (not always) on top of side walls of the chamber, often pawing walls; may be accompanied by sniffing or slow side-to-side movement of head. Does not include grooming movements or eating, even if performed while standing on hind legs.
Scale	All four limbs off the floor but at least two limbs on the side of the chamber. Standing on the food cup counts as scaling.
Stretch	Body is elongated with the back posture 'flatter' than normal. Stretching is often accompanied by immobility, like freezing, but is distinguished by the shape of the back.

**Table 2.2. Behavior definitions**  
Definitions are provided for each behavior scored.

### 2.3.1.2 *Hand scoring with high inter-observer reliability*

Frames were systematically hand scored by five observers blind to rat identity, session number, and trial type (see Materials and Methods for hand scoring approach and trial anonymization). A comparison data set consisting of 12 trials (900 frames) was also scored by each observer. A correlation matrix compared % identical observations for the 900 comparison frames for each observer-observer pair (Figure 2.3A). Mean % identical observation was 82.83%, with a minimum observer-observer pair agreement of 75.89% and a maximum of 90.56%. Previous studies scoring the presence or absence of freezing have reported inter-observer reliability as an R value: 0.93 (Parnas et al., 2005), 0.96 (Pickens et al., 2010), and 0.97 (Jones and Monfils, 2016). Another study simply reported >95% inter-observer agreement (Badrinarayan et al., 2012). These values exceed our mean % identical observation. However, we hand scored nine discrete behaviors. We observed a negative relationship between the number of behavior categories present and % identical observations ( $R^2 = 0.17$ ,  $p = 2.27 \times 10^{-6}$ , Figure 2.3B). Mean percent identical observation was 95% when two behavior categories were present, and 92.5% when three behavior categories were present. Even when eight behavior categories were present, a mean percent identical observation of 78% was achieved. Our approach yielded high inter-observer reliability across trials with few and many behavior categories present. An example of a single-trial ethogram resulting from hand scoring is shown in Figure 2.3C (female rat, discrimination session 8, uncertainty cue). Videos 1-4 show example danger trials for four different rats (females in Videos 1 & 3, males in Videos 2 & 4).



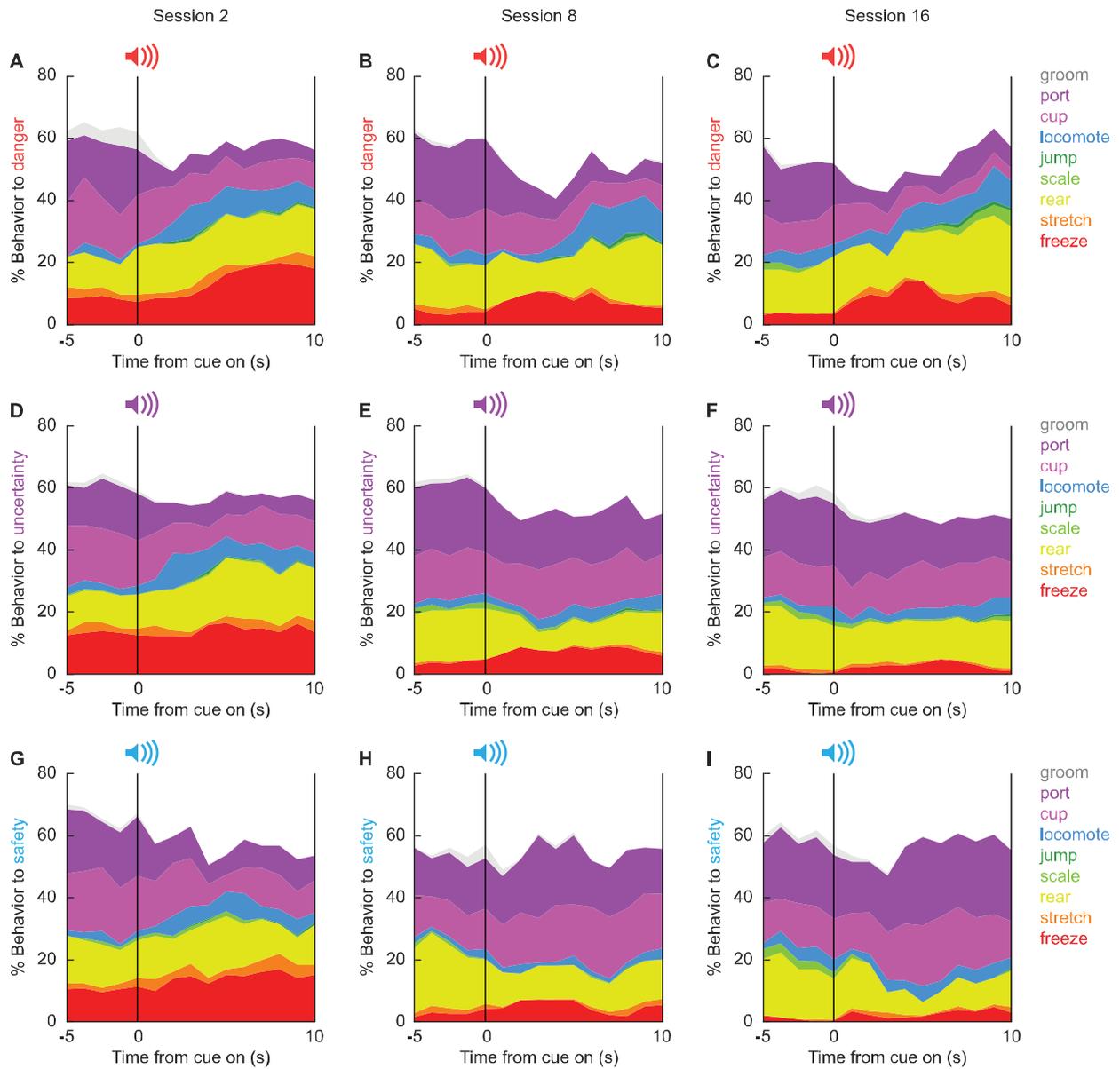
**Figure 2.3.** Inter-rater reliability

**(A)** Percentage of identical observations between observer-observer pairs. **(B)** Percentage of identical observations as a function of the number of behaviors present in a trial. **(C)** Example ethogram from a single uncertainty cue presentation, taken from a female during session 8.

### 2.3.1.3 Temporal ethograms reveal shifting behavioral patterns over discrimination

The 86,400 scored frames allowed us to construct temporal ethograms for danger (Figure 2.4A-C), uncertainty (Figure 2.4D-F), and safety (Figure 2.4G-I) during sessions 2 (Figure 2.4, column 1), 8 (Figure 2.4, column 2), and 16 (Figure 2.4, column 3). Shifts in the composition of behavior from baseline to cue presentation were apparent across all ethograms. During session 2 (column 1), behavioral shifts lacked cue-specificity. Temporal ethograms revealed danger, uncertainty, and safety to equally suppress grooming, port, and cup behavior, but increase freezing, and locomotion. Generalized cue control of behavior was supported by multiple analysis of variance (MANOVA) for all nine behavior categories [factors: cue (danger vs. uncertainty vs. safety), time (15 1-s bins: 5-s baseline → 10-s cue), and sex (female vs. male)] revealing a significant main effect of time ( $F_{(126,2772)} = 2.37, p=5.93 \times 10^{-15}$ ), but neither a significant main effect of cue ( $F_{(18,74)} = 1.00, p=0.47$ ) nor a significant cue x time

interaction ( $F_{(252,5544)} = 1.12, p=0.11$ ). Cue-specific shifts in behavior were apparent by session 8 (column 2), and continued to session 16 (column 3). Now, MANOVA revealed significant main effects of cue (session 8,  $F_{(18,74)} = 3.39, p=0.0001$ ; session 16,  $F_{(18,74)} = 4.44, p=0.000002$ ), and significant cue x time interactions (session 8,  $F_{(252,5544)} = 1.52, p=3.31 \times 10^{-8}$ ; session 16,  $F_{(252,5544)} = 1.52, p=4.74 \times 10^{-7}$ ).



**Figure 2.4.** Temporal ethograms during discrimination

Mean percent behavior from 5s prior through 10-s cue presentation is shown for the danger cue during sessions **(A)** 2, **(B)** 8, **(C)** and 16; the uncertainty cue during sessions **(D)** 2, **(E)** 8, and **(F)** 16; and the safety cue during sessions **(G)** 2, **(H)** 8, and **(I)** 16. Behaviors are groom (gray), port (dark purple), cup (light purple), locomote (blue), jump (dark green), scale (light green), rear (yellow), stretch (orange), and freeze (red).

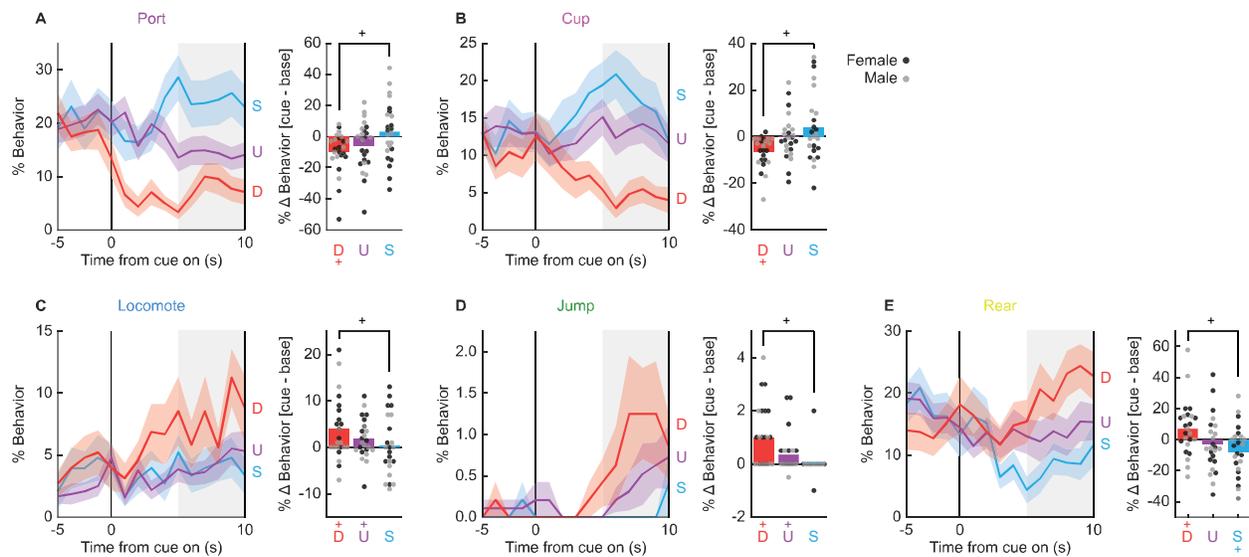
#### 2.3.1.4 *Danger orchestrates a suite of behaviors*

A central question driving this study was what behaviors come under the specific control of the fear conditioned, danger cue? To determine this, we focused on session 16, when discrimination was at its most complete. We first performed MANOVA for the 5-s baseline period [factors: cue (danger vs. uncertainty vs. safety), time (5, 1-s bins), and sex (female vs. male)]. As expected, MANOVA returned no main effect of cue, time, nor a cue x time interaction ( $F_s < 1.5$ ,  $p_s > 0.1$ ). Univariate ANOVA results were subjected to Bonferroni correction ( $p < 0.0055$ ,  $0.05/9 = 0.0055$ ) to account for the nine separate analyses. Like for MANOVA, univariate ANOVA for each of the nine behaviors showed no main effect of cue, time, nor a cue x time interaction. In contrast to all other behaviors, univariate ANOVA for baseline freezing showed a main effect of sex ( $F_{(1,22)} = 10.37$ ,  $p = 0.004$ ). ANOVA for freezing across the baseline and cue periods revealed a significant sex x cue x time interaction ( $F_{(28,616)} = 1.94$ ,  $p = 0.003$ ). Females only froze during early danger presentation while males froze for the duration of danger presentation. The unique freezing pattern warrants separate consideration, which we return to later.

MANOVA was then performed for the 10-s cue period [factors: cue (danger vs. uncertainty vs. safety), time (10, 1-s bins), and sex (female vs. male)]. MANOVA returned significant main effects of cue and time, as well as a significant cue x time interaction ( $F_s > 1.3$ ,  $p_s < 0.005$ ). Of most interest, univariate ANOVA found a significant main effect of cue for six of the nine behaviors: port ( $F_{(2,44)} = 32.15$ ,  $p = 2.47 \times 10^{-9}$ , Figure 2.5A), cup ( $F_{(2,44)} = 18.40$ ,  $p = 0.00002$ , Figure 2.5B), locomote ( $F_{(2,44)} = 6.33$ ,  $p = 0.004$ , Figure 2.5C), jump ( $F_{(2,44)} = 10.90$ ,  $p = 0.0001$ , Figure 2.5D), rear ( $F_{(2,44)} = 8.64$ ,  $p$

= 0.001, Figure 2.5E), and freeze ( $F_{(2,44)} = 13.86, p = 0.00002$ ). Danger suppressed port and cup behavior (Figure 2.5A, B line graphs), but promoted locomotion, jumping, and rearing (Figure 2.5C, D, E line graphs). Danger-specific control of behavior was most apparent in the last 5 s of cue presentation (Figure 2.5, shaded region).

Claiming danger specificity requires that % behavior during the danger cue differs from baseline as well as the safety cue. To test this, we subtracted mean % behavior during the 5-s baseline from mean % behavior during the last 5 s of cue presentation, giving  $\% \Delta$  danger,  $\% \Delta$  uncertainty, and  $\% \Delta$  safety for each subject. We constructed 95% BCIs for each cue/behavior. 95% BCIs for  $\% \Delta$  danger did not contain zero for each of the five behaviors (Figure 2.5), meaning that levels of behavior during cue presentation differed from baseline. Danger presentation decreased port and cup behavior below baseline, but increased locomotion, jumping, and rearing over baseline. 95% BCIs for  $\% \Delta$  uncertainty revealed increased locomotion and jumping, while 95% BCIs for  $\% \Delta$  safety revealed only decreased rearing. To demonstrate danger-specificity, we subtracted  $\% \Delta$  safety from  $\% \Delta$  danger. We then constructed 95% BCIs for the difference score for each behavior. Confirming danger specificity (greater changes for danger than for safety), 95% BCIs did not contain zero for each of the five behaviors. Thus, danger specifically and selectively suppressed reward-related port and cup behavior, but promoted locomotion, jumping, and rearing.



**Figure 2.5.** Danger-elicited behaviors

Line graphs show mean  $\pm$  SEM percent behavior from 5s prior through 10-s cue presentation for danger (red), uncertainty (purple), and safety (blue) for (A) port, (B) cup, (C) locomote, (D) jumping, and (E) rearing. Bar plots show mean change in behavior from baseline (5 s prior to cue) compared to last 5 s of cue. Individuals represented by black (female) and gray (male) dots. +95% bootstrap confidence interval for danger vs. safety (black), danger vs. baseline (red), or safety vs. baseline (blue) comparison does not contain zero (black).

### 2.3.1.5 Associatively acquired behaviors generalize early

By the end of session 16 each rat had received 96 total foot shocks. It is possible that danger-specific control of multiple behaviors was only observed in session 16 because rats received far more cue-shock pairings than a typical Pavlovian conditioning procedure employs. Session 2 provided a comparison to numbers of cue-shock pairings more typical of fear conditioning studies; rats had received 12 total foot shocks by session's end. The key question was whether pattern of danger-elicited behaviors in session 2 resembled the pattern in session 16, or if a fundamentally different pattern was observed. To determine this, we performed univariate ANOVA for danger [factors: session (2 vs. 16) and time (15, 1-s bins)] for each of the five behaviors showing

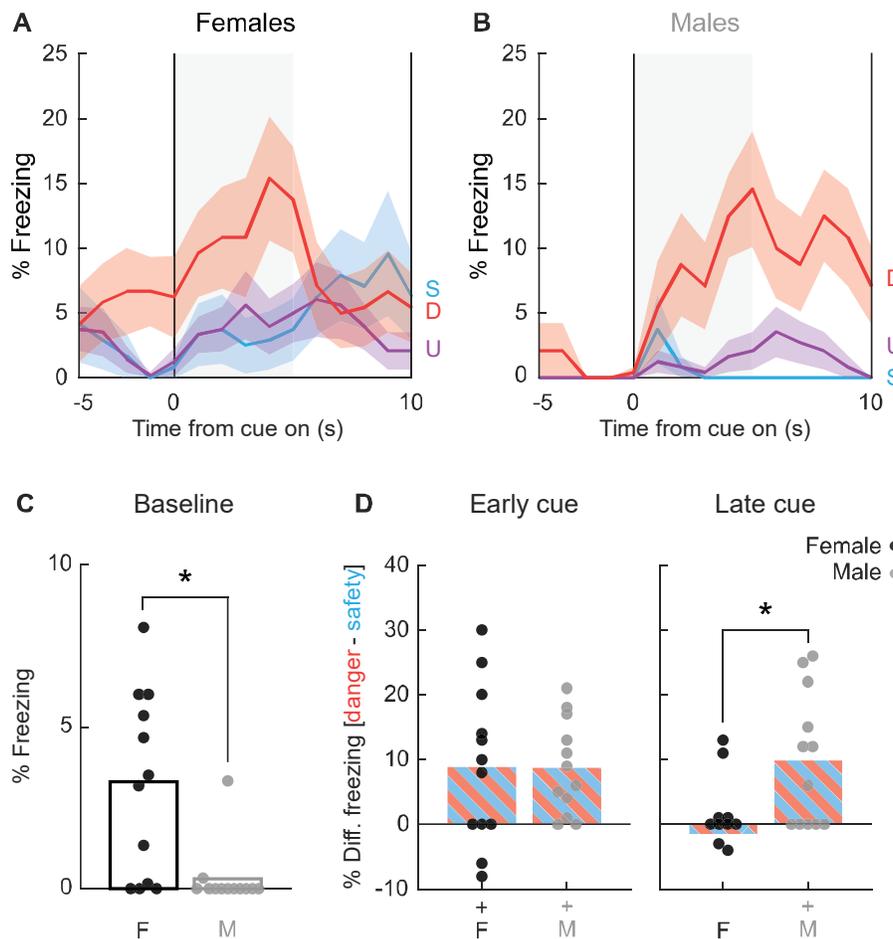
session 16 selectivity. Confirming near identical temporal patterns of behavior expression during sessions 2 and 16, ANOVA found no significant session x time interaction for any behavior [port ( $F_{(14,322)} = 0.45, p=0.96$ ), cup ( $F_{(14,322)} = 0.61, p=0.86$ ), locomote ( $F_{(14,322)} = 1.09, p=0.37$ ), jump ( $F_{(14,322)} = 1.23, p=0.25$ ), and rear ( $F_{(14,322)} = 0.92, p=0.54$ )]. Thus, danger orchestrated a suite of behaviors even early in discrimination. Recall that early discrimination (session 2) was marked by non-specific cue control of behaviors. This would mean that associatively acquired behaviors initially generalized to uncertainty and safety – and that discrimination consisted of restricting behavior to danger. In support, univariate ANOVA for session 2 [factors: cue (danger vs. uncertainty vs. safety), time (15, 1-s bins), and sex (female vs. male)] found no cue x time interaction for any of the five, danger-specific behaviors (all  $F_s < 1.2$ , all  $p_s > 0.3$ ).

#### *2.3.1.6 Sex informs the temporal pattern of freezing*

We return to the case of freezing; the most measured overt fear conditioned behavior. We again focus on session 16 during which discrimination was most complete. Female and male rats differed in the temporal pattern and cue-specificity of freezing. Females showed higher baseline freezing levels, a rapid increase in freezing that was specific to danger in the first 5 s, then became non-specific and declined back to baseline levels in the last 5 s (Figure 2.6A). By contrast, males show little baseline freezing and danger-specific freezing increases that persisted throughout cue presentation (Figure 2.6B). Baseline freezing differences were confirmed with independent samples t-test ( $t_{22} = 3.22, p=0.0039$ ; Figure 2.6C). Confirming sex differences in the temporal pattern of freezing, differential freezing to danger and safety was equivalent in females and males during early cue presentation ( $t_{22} = 0.02, p=0.98$ ;

Figure 2.6D left), but differed during late cue ( $t_{22} = 2.80, p=0.01$ ; Figure 2.6D right).

Generalized freezing to all cues was observed during session 2, with freezing increases more evident in males. Thus, discrimination consisted of restricting freezing to danger in males, and selectively freezing to early danger presentation in females.



**Figure 2.6.** Special case of freezing

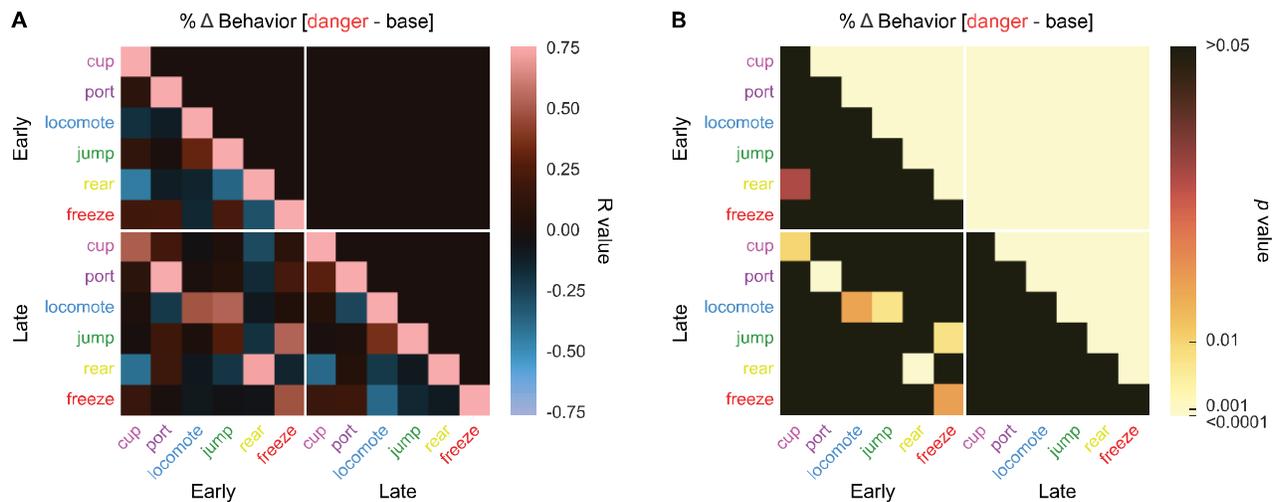
Line graphs show mean  $\pm$  SEM percent freezing from 5s prior through 10-s cue presentation for danger (red), uncertainty (purple), and safety (blue) for **(A)** females and **(B)** males. **(C)** Percent freezing during baseline (5s prior to cue) is shown for females (black) and males (gray). **(D)** Mean differential freezing to danger and safety is shown.

### 2.3.1.7 Danger-elicited behaviors are independently expressed

Danger suppression of reward-related port and cup behaviors could simply be the byproduct of danger-elicited freezing. Such a relationship has previously been reported (Bouton and Bolles, 1980; Mast et al., 1982). To examine the relationship between reward-related behaviors and freezing, in addition to other possible behavior-behavior relationships, we calculated  $\% \Delta$  behavior for early (first 5 s) and late (last 5 s) danger presentation for the six danger-elicited behaviors: cup, port, locomote, jump, rear, and freeze. We constructed 12 x 12 matrices containing the R values (Figure 2.7A) and  $p$  values (Figure 2.7B) for the Pearson's correlation coefficient for each behavior-behavior comparison during the two danger periods. Surprisingly, only one behavior-behavior relationship was observed during the early danger presentation period (Figure 2.7A, upper left quadrant). Early rearing and early cup behavior were negatively correlated ( $R = -0.43$ ,  $p = 0.03$ , but note this would not survive Bonferroni correction). Even more, no behavior-behavior relationships were observed during late danger presentation (Figure 2.7A, lower right quadrant). These results suggest the six behaviors are more or less expressed independently of one another.

Maybe our analysis cannot detect behavior-behavior relationships? To test this, we compared behaviors across the early and late danger periods. Now, the correlation matrix revealed a band of positive R values cutting diagonally across the bottom left quadrant. 5 of the 6 behaviors showed positive early-late relationships with themselves: cup ( $R = 0.51$ ,  $p = 0.01$ ), port ( $R = 0.87$ ,  $p = 2.67 \times 10^{-8}$ ), locomote ( $R = 0.48$ ,  $p = 0.017$ ), rear ( $R = 0.71$ ,  $p = 7.92 \times 10^{-5}$ ), and freeze ( $R = 0.48$ ,  $p = 0.017$ ). In other words, changes in cup behavior evident during early danger presentation persisted to late

danger presentation. Jumping was an exception to this trend, as there was no relationship between early and late jumping levels to danger. Overall, danger-elicited behaviors were expressed independently of one another.



**Figure 2.7.** Behavior-behavior correlations

**(A)** A correlation matrix for the six cue specific behaviors (port (dark purple), cup (light purple), locomote (blue), jump (dark green), rear (yellow), and freeze (red) comparing mean percent behavior during early (first 5 s) and late (last 5 s) cue is shown. Lighter red values indicate positive R values, lighter blue values indicate negative R values. Black indicates  $R = 0$ . P values associated with each associated R value are shown in **(B)**. Black indicates p values greater than 0.05, while increasingly lighter values indicate lower p values.

### 2.3.2 *Experiment 2*

In Experiment 2, we aimed to answer two questions: 1) were the danger-elicited behaviors during discrimination in Experiment 1 dependent on foot shock delivery, and 2) were these behaviors due to the presence of the reward apparatus? To answer this, rats received danger vs. safety discrimination, then were given extinction tests with reward apparatus absent or present. During extinction testing, we captured and hand scored behavior frames before, during, and following cue presentation.

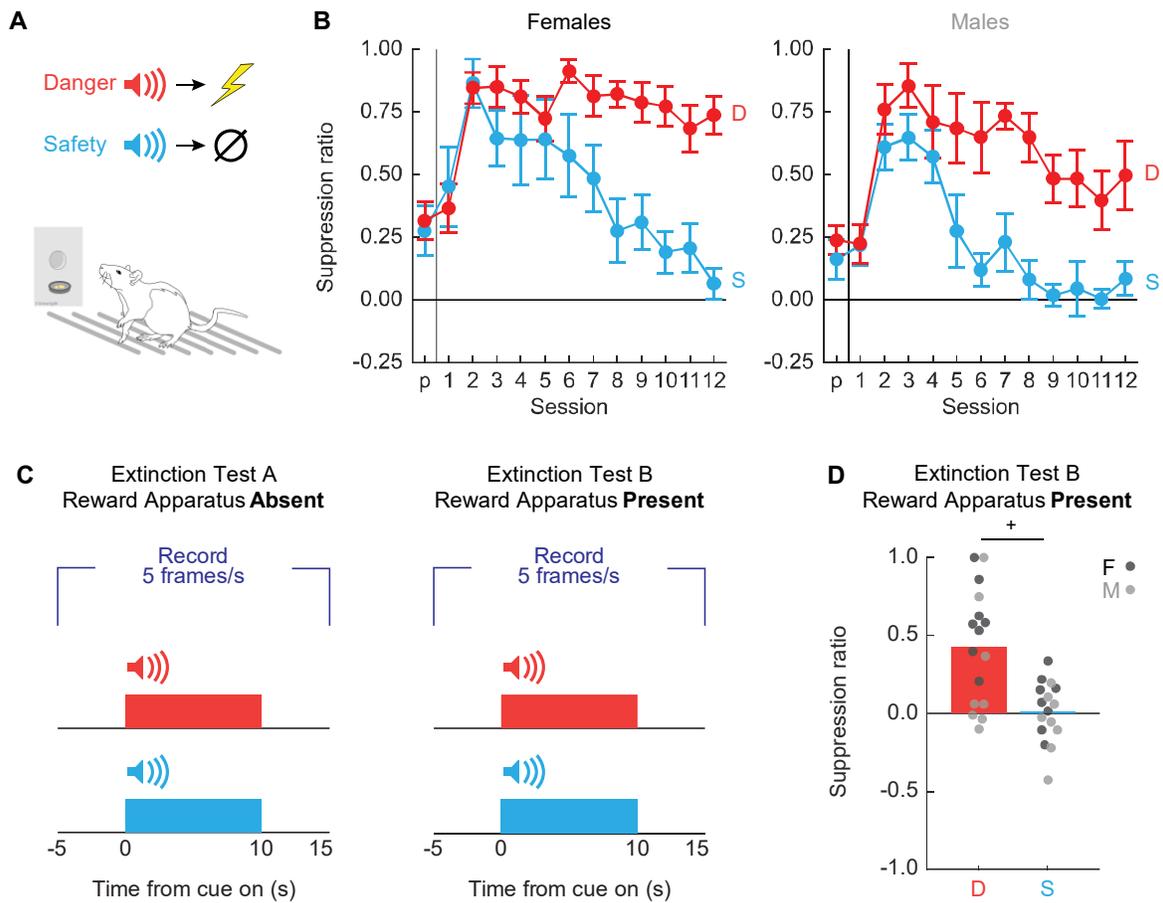
#### 2.3.2.1 *Conditioned suppression reveals complete discrimination during extinction with reward apparatus present*

Sixteen Long Evans rats (8 females) were trained to nose poke in a central port for food reward as in Experiment 1. Nose poking was reinforced on a 60-s variable interval schedule throughout behavioral testing. Independent of the poke-food contingency, auditory cues were played through overhead speakers, and foot shock delivered through the grid floor (Figure 2.8A). The experimental design consisted of two cues deterministically predicting foot shock: danger ( $p=1$ ), and safety ( $p=0$ ) (Figure 2.8A).

To determine if we observed complete behavioral discrimination, we performed ANOVA for suppression ratios [factors: cue (danger vs. safety), session (13 total: 1 pre-exposure and 12 discrimination), and sex (female vs. male)]. Complete behavioral discrimination emerged over testing (Figure 2.8B). ANOVA found a significant main effect of cue and a significant cue x session interaction ( $F_s > 8$ ,  $p_s < 0.0001$ ; see Table 2.3 for specific values). No significant main effect or interactions with sex were observed. Following the 12 discrimination sessions, each rat received two extinction test

sessions. In both test sessions each cue was presented 4 times. In one test session the nose poke and food cup were absent while in the other test session the nose poke and food cup were present (Figure 2.8C). Test order was fully counterbalanced. 95% BCIs for differential suppression ratio (danger - safety) during extinction test with the reward present revealed complete discrimination (Figure 2.8D). The 95% BCI did not contain zero [lower bound = 0.24, upper bound = 0.60].

We captured 25,600 total frames (800 frames/test x 16 rats x 2 tests) during extinction testing. Frames were hand scored for nine discrete behaviors: cup, freezing, grooming, jumping, locomotion, port, rearing, scaling, and stretching, plus “background” as in Experiment 1, with the exception that if a trial did not have the reward apparatus present, then food cup and nose poke were not scored.



**Figure 2.8.** Experimental design and nose poke suppression

**(A)** Conditioned suppression procedure during which rats nose poke for food, while danger (red;  $p=1$ ), and safety (blue;  $p=0$ ) cues are played overhead and shocks delivered through floor. **(B)** Mean  $\pm$  SEM suppression ratios for danger (red), and safety (blue) from pre-exposure through discrimination session 12 are shown for (left) females, and (right) males. **(C)** Rats received one extinction test with reward apparatus absent (left), and another with reward apparatus present (right), counterbalanced. Five video frames were captured per second, starting 5-s prior to cue onset and continuing through 5-s after cue offset. **(D)** Mean + individual suppression ratios for each cue are shown for extinction with reward apparatus present. Individuals represented by black (female) and gray (male) dots. +95% bootstrap confidence interval does not contain zero.

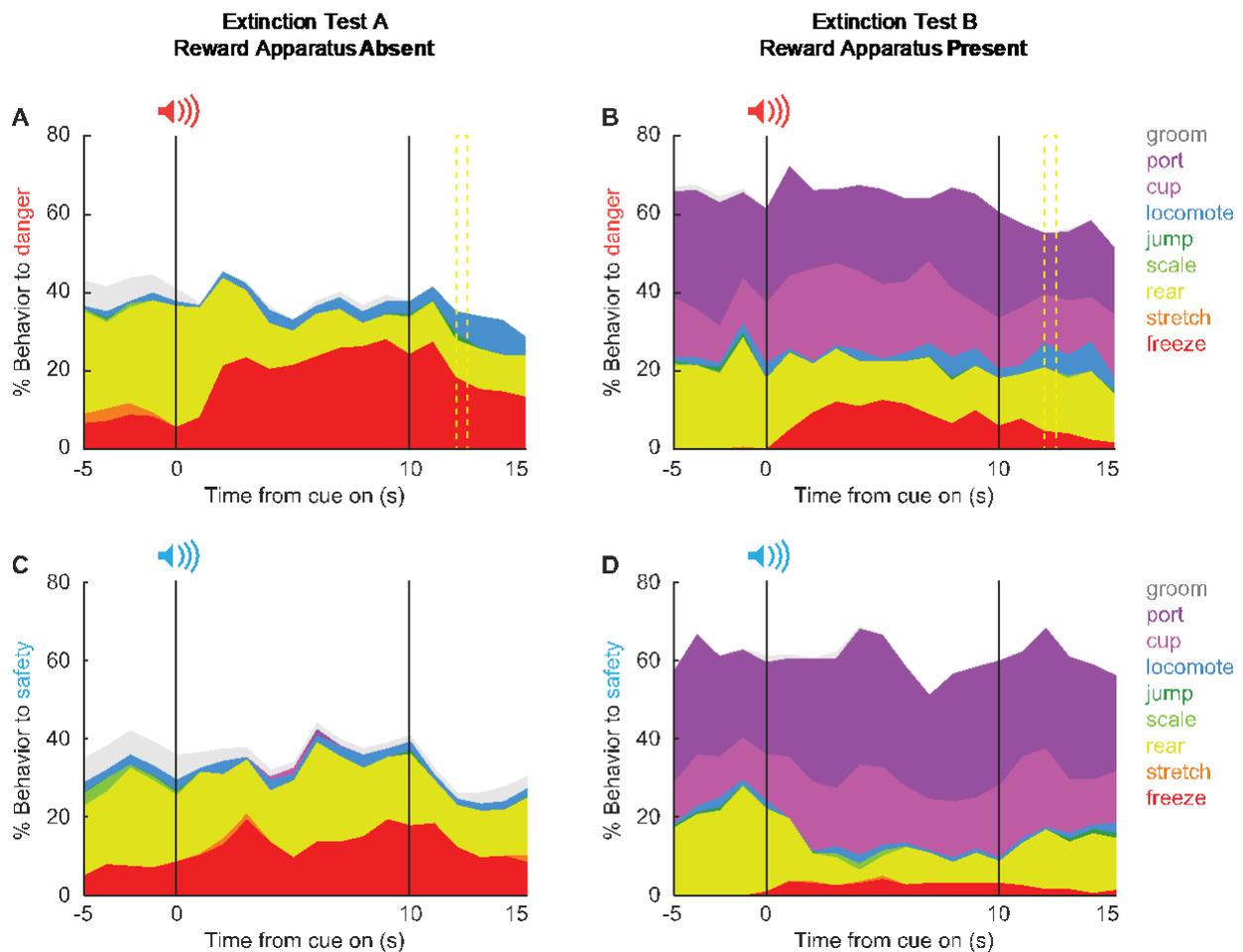
	Suppression ratio	
Cue	<b>F<sub>(1,14)</sub> = 75.98</b>	<b>p = 4.98 x 10<sup>-7</sup></b>
Cue x sex	F <sub>(1,14)</sub> = 0.26	p = 0.62
Session	<b>F<sub>(12,168)</sub> = 20.79</b>	<b>p = 2.04 x 10<sup>-27</sup></b>
Session x sex	F <sub>(12,168)</sub> = 1.54	p = 0.12
Cue x session	<b>F<sub>(12,168)</sub> = 8.54</b>	<b>p = 1.50 x 10<sup>-12</sup></b>
Cue x session x sex	F <sub>(12,168)</sub> = 1.22	p = 0.28
Sex	F <sub>(1,14)</sub> = 4.48	p = 0.053

**Table 2.3.** Experiment 2 ANOVA results for suppression ratio.

Significant main effects and interactions are bolded.

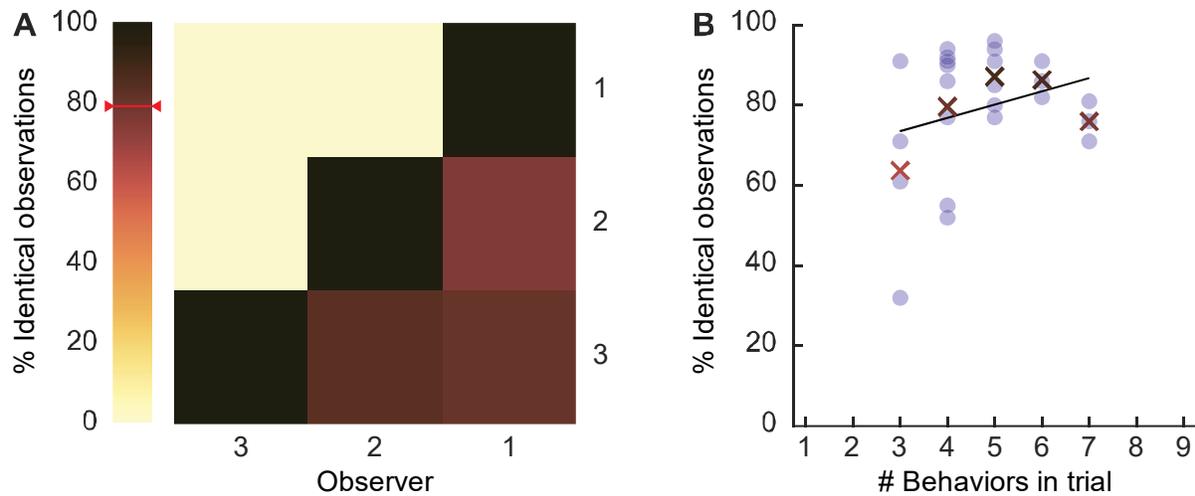
### 2.3.2.2 *Danger-elicited locomotion peaks when foot shock would have occurred*

The 25,600 scored frames allowed us to construct temporal ethograms for danger (Figure 2.9A, B) and safety (Figure 2.9C, D), during the extinction test with reward apparatus absent (Figure 2.9, column 1), and during the extinction test with the reward apparatus present (Figure 2.9, column 2). Hand scoring showed high inter-rater reliability even when many behaviors were present in a single trial (Figure 2.10). Mean % identical observation was 79.25%. Cue-specific changes in behavior during and following cue presentation were evident. In support, MANOVA [factors: sex (female vs. male), test type (absent vs. present), order (absent first vs. present first), cue (danger vs. safety), and time (20 1-s bins: 5-s baseline → 10-s cue → 5-s post cue)] for the seven behaviors common to both tests [groom, locomote, jump, scale, rear, stretch, and freeze] revealed a significant main effect of time ( $F_{(133,1596)} = 2.14$ ,  $p = 9.44 \times 10^{-12}$ ) and a significant cue x time interaction ( $F_{(133,1596)} = 1.46$ ,  $p = 0.001$ ).



**Figure 2.9.** Temporal ethograms during extinction

Mean percent behavior from 5s prior through 5s following cue offset is shown for the danger cue during extinction with **(A)** reward apparatus absent, and **(B)** reward apparatus present; and the safety cue during extinction with **(C)** reward apparatus absent, and **(D)** reward apparatus present. Behaviors are groom (gray), port (dark purple), cup (light purple), locomote (blue), jump (dark green), scale (light green), rear (yellow), stretch (orange), and freeze (red). Note, port and cup are not shown for A and C because the food cup and nose port were absent.

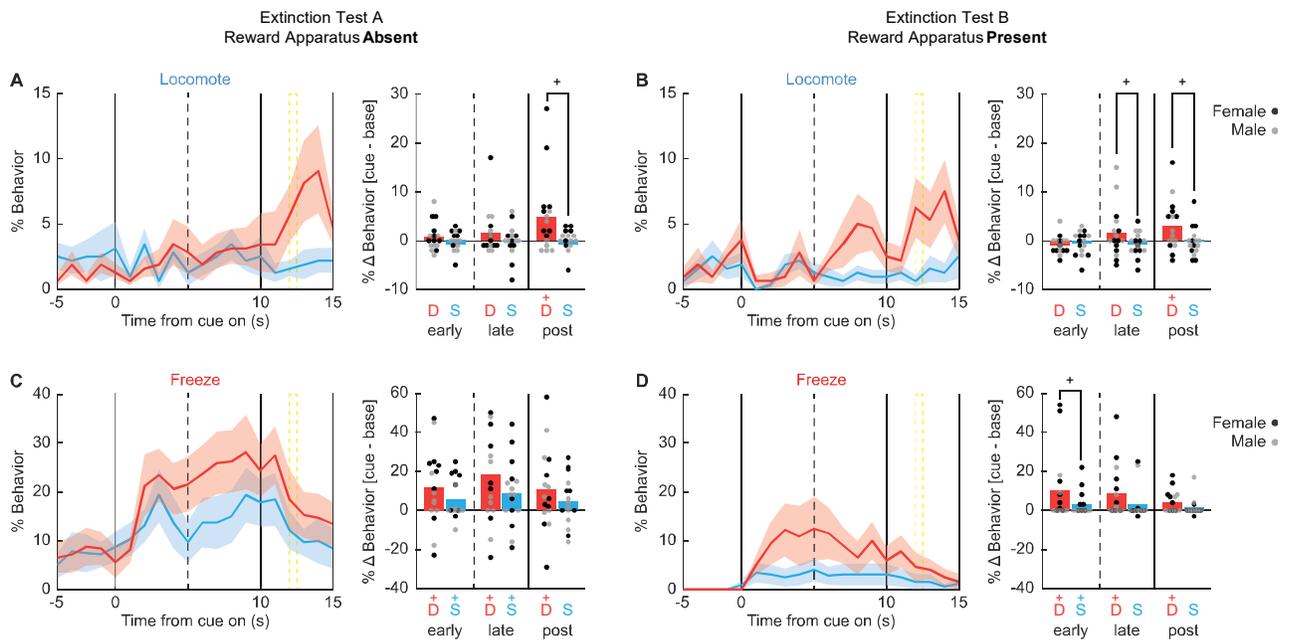


**Figure 2.10.** Inter-rater reliability

**(A)** Percentage of identical observations between observer-observer pairs. **(B)** Percentage of identical observations as a function of the number of behaviors present in a trial.

Of the seven behaviors, danger only increased locomotion during both test types (Figure 2.11A, B). In support, univariate ANOVA for locomotion [Bonferroni-corrected  $p < 0.007$  ( $0.05/7 = 0.007$ ); factors: sex (female vs. male), test type (absent vs. present), order (absent first vs. present first), cue (danger vs. safety), and time (20 1-s bins: 5-s baseline  $\rightarrow$  10-s cue  $\rightarrow$  5-s post cue)] found a significant cue x time interaction ( $F_{(19,228)} = 3.12$ ,  $p = 0.000026$ ). Danger-elicited locomotion was most prominent following cue offset, around the time shock would have occurred. 95% BCIs revealed danger-elicited locomotion to exceed baseline and safety cue levels during the 5-s post-cue periods for both the Absent (Figure 2.11A) and Present test (Figure 2.11B). Additionally, the 95% BCI revealed danger-elicited locomotion to exceed safety-elicited locomotion during the late cue period during the Present test, though danger-elicited locomotion did not exceed baseline (Figure 2.11B). Locomotion never increased during safety trials (all

95% BCIs contain zero). Danger-elicited locomotion occurred regardless of test order, as ANOVA revealed no significant order interactions ( $F_s < 1.5$ ,  $p_s > 0.2$ ). Sex partially mediated the temporal expression of locomotion, with ANOVA finding a significant sex  $\times$  cue  $\times$  time interaction ( $F_{(19,228)} = 2.34$ ,  $p = 0.002$ ). Females showed more robust post-cue, danger locomotion during both test types. Males showed more robust danger-elicited locomotion during the late cue period during the Present test. The results reveal that danger-elicited locomotion transfers to extinction settings when both foot shock and the reward apparatus were absent.



**Figure 2.11.** Danger elicits locomotion during extinction

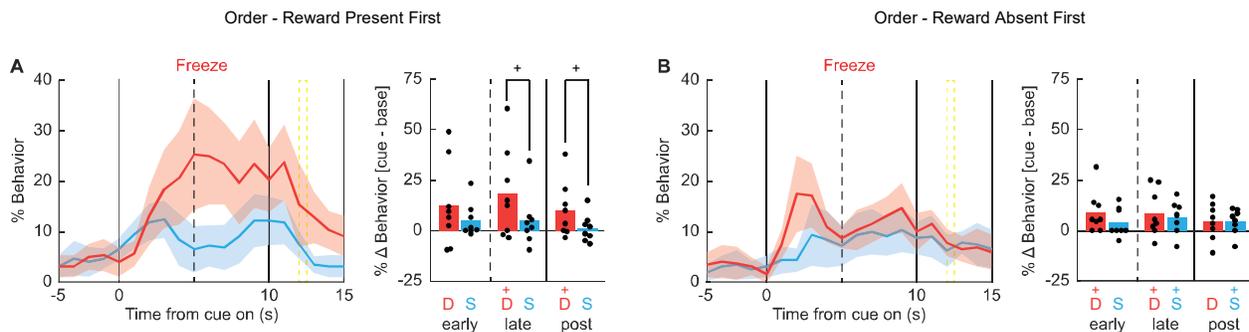
Line graphs show mean  $\pm$  SEM percent behavior from 5s prior through 10-s cue presentation for danger (red), and safety (blue) for locomotion during the **(A)** reward apparatus absent extinction test, and **(B)** reward apparatus present extinction test. Bar plots show mean change in behavior from baseline (5 s prior to cue) compared to early (first 5 s), late (last 5 s) and post (5 s after offset) cue periods. Individuals represented by black (female) and gray (male) dots. The same is shown for freezing **(C&D)**. +95% bootstrap confidence interval for danger vs. safety (black), danger vs. baseline (red), or safety vs. baseline (blue) comparison does not contain zero (black).

### 2.3.2.3 Freezing is less danger-specific and is sensitive to time, test type, and order

Unlike locomotion, there was lesser evidence of danger-specific freezing during extinction testing (Figure 2.11C, D). Most notably, univariate ANOVA [correction and factors identical to locomotion] found that the cue x time interaction failed to achieve significance ( $F_{(19,228)} = 1.25, p = 0.011$ ). When organizing % freezing by test type, there was no period (early cue, late cue, and post cue) during which freezing increases over baseline were selective to danger (Figure 2.11C, D). The only period during which freezing to danger exceeded freezing to safety was the early cue period when the reward apparatus was present (Figure 2.11, right). Though even during this period increases in freezing to safety were observed. Instead, freezing tended to generalize to safety; meaning it was cue-evoked but not cue-specific. Additionally, freezing was more prominent during extinction testing with the reward apparatus absent. In support, univariate ANOVA revealed significant main effects of time ( $F_{(19,228)} = 5.13, p = 3.64 \times 10^{-10}$ ) and test ( $F_{(1,12)} = 21.20, p = 0.001$ ). Like freezing, neither rearing nor jumping showed evidence of danger-specificity with univariate ANOVA for each finding no significant cue x time interaction ( $F_s < 1.5, p_s > 0.2$ ).

However, order mediated the specificity of danger-elicited freezing. Rats receiving the Present test first showed selective and differential freezing to danger (Figure 2.11.1A). Rats receiving the Absent test first showed no evidence of selective and differential freezing to danger (Figure 2.11.1B). In support, univariate ANOVA returned a significant order x cue x time interaction ( $F_{(19,228)} = 2.14, p = 0.002$ ). Of note, no significant order x cue x time interaction was observed for locomotion ( $F_{(19,228)} =$

1.03,  $p = 0.43$ ). The same rats that showed robust danger-elicited locomotion across both test types showed danger-elicited freezing that was sensitive to test order.



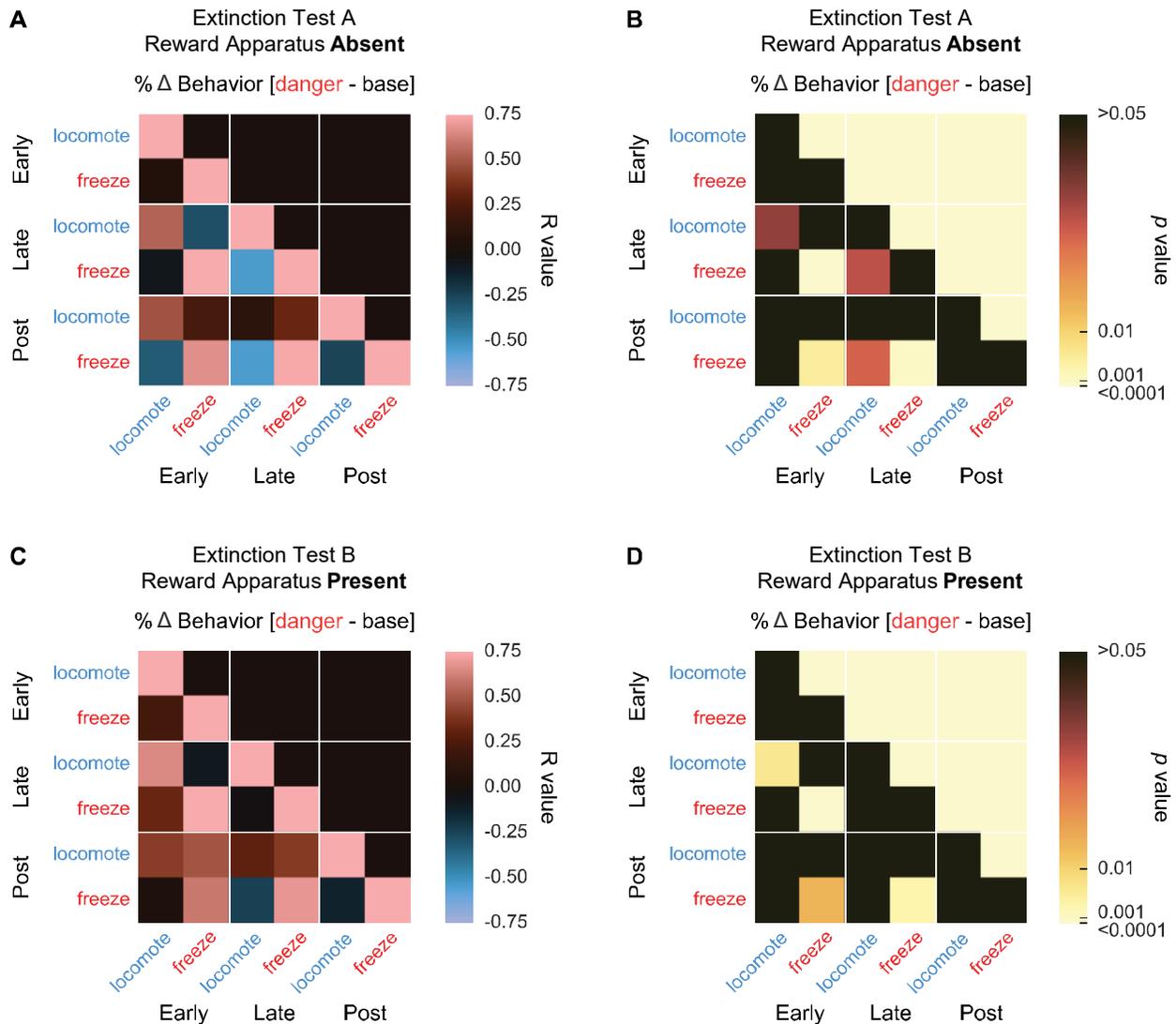
**Figure 2.11.1.** Freezing separated by test order

Line graphs show mean  $\pm$  SEM percent freezing from 5s prior through 10-s cue presentation for danger (red), and safety (blue) for rats receiving the reward apparatus present extinction test (**A**) first and (**B**) second. Mean responding is taken from both test types. Bar plots show mean change in behavior from baseline (5 s prior to cue) compared to early (first 5 s), late (last 5 s) and post (5 s after offset) cue periods. Individuals represented by black dots. +95% bootstrap confidence interval for danger vs. safety (black), danger vs. baseline (red), or safety vs. baseline (blue) comparison does not contain zero (black). Female and male data are collapsed as sex did not influence the significant interaction of order.

#### 2.3.2.4 Danger-elicited behaviors are independently expressed during extinction

We were interested to see if there were relationships between danger-elicited locomotion and freezing during extinction testing. To determine this we calculated Pearson's correlation coefficients (R value) for individual freezing and locomotion (% behavior over baseline) during early, late, and post-danger cue periods in extinction sessions with reward apparatus absent (Figure 2.12A and B) and with reward apparatus present (Figure 2.12C and D). As in Experiment 1, we found no evidence of inhibitory relationships between locomotion and freezing. That is, no comparison found a negative

R value. This was true both within and between trial periods. Instead, and like for Experiment 1, correlational analyses reveal significant, positive relationships within behaviors across trial periods. These positive relationships were more prominent during extinction testing with the reward apparatus present. In particular, freezing was positively correlated across all trial periods during the present extinction sessions [early-late  $R = 0.82$ ,  $p = 1.05 \times 10^{-4}$ ; early-post  $R = 0.60$ ,  $p = 0.015$ ; post-late  $R = 0.68$ ,  $p = 0.0036$ ]. These results demonstrate that opposing danger-elicited behaviors are independently expressed during extinction.



**Figure 2.12.** Behavior-behavior correlations during extinction

**(A)** A correlation matrix for locomote (blue) and freeze (red) comparing mean percent behavior during early (first 5 s), late (last 5 s), and post (5 s after) cue period is shown for the reward absent extinction test. Lighter red values indicate positive R values, lighter blue values indicate negative R values. Black indicates R = 0. P values associated with each associated R value are shown in **(B)**. Black indicates p values greater than 0.05, while increasingly lighter values indicate lower p values. Same shown for behavior correlations during reward present extinction test **(C&D)**.

## 2.4 Discussion

We set out to quantify behaviors elicited by a fear conditioned, danger cue. Consistent with virtually all studies of Pavlovian fear conditioning (but see Amoranth et al., 1999), we observed danger-elicited freezing over the course of acquisition (Experiment 1) and during extinction testing (Experiment 2). Yet, freezing was not the dominant danger-elicited behavior. Instead, danger orchestrated a suite of behaviors. During Experiment 1, danger suppressed reward behavior directed toward the site of food delivery and the location of the rewarded action. Even more, danger elicited locomotion, jumping, and rearing. During Experiment 2, danger again suppressed rewarded action and continued to elicit locomotion. During both Experiments, freezing was most prominent at danger cue onset. Locomotion was most prominent towards danger cue offset (Experiment 1), and when shock would have occurred in extinction (Experiment 2).

Before discussing our results further, an important limitation should be raised. 40-50% of frames could not be assigned to a specific behavior and were labeled as background. This was due to three main factors. First, in order to objectively hand score many behaviors, we developed mutually-exclusive, specific definitions. Our strict definitions meant erring on the side of labeling a behavior background if there was uncertainty in judgment. Second, use of a single, side view camera meant the observer could not view a rat's forelimbs or face when the rat was turned away from the camera. If forelimb and face position could not be determined the frame was labeled background. Finally, transition behaviors (e.g., switching from rearing to locomotion) and other behaviors (e.g., sniffing and turning) that did not fit into one of the nine

behavior definitions were labeled background. The upside of this limitation is high confidence in behavior judgments and high inter-rater reliability for those judgments.

Studies assessing auditory fear conditioning in a neutral context routinely report freezing to account for >80% of behavior during cue presentation (Bolles and Collier, 1976; Maren et al., 1997; Anagnostaras et al., 1999; Wilensky et al., 1999; Quirk, 2002; Koo et al., 2004; Rogers and Kesner, 2004; Iordanova et al., 2006; Shumake et al., 2014; Foilb et al., 2016; Furlong et al., 2016). The sheer number of demonstrations, and number of groups independently observing dominant freezing, puts us firmly in the minority. Placing us further in the minority, we observe danger-elicited locomotion, jumping, and rearing. These behaviors are characterized by lateral and vertical movements, polar opposites to the immobility that characterizes freezing. A commonality of the studies above was use of contexts in which only cues and shocks were delivered, with shocks being more intense than shocks used in our studies. These experimental settings favor freezing. It is likely that our inclusion of competing, reward behaviors and use of lower shock intensities permitted a broader range of danger-elicited behaviors to be observed (Holland, 1979; Mitchell et al., 2022).

Indeed, we are not the first group to observe locomotion, jumping, or rearing in defensive settings in rats. Using more traditional Pavlovian fear conditioning designs, Shansky and colleagues have observed “darting”, rapid forward movements across the test chamber, to a fear conditioned cue (Gruene et al., 2015). While more readily observed in female rats, darting can be observed in males under certain experimental conditions (Mitchell et al., 2022). Our definition of locomotion aligns well with the Shansky lab definition of darting. We found that danger-elicited locomotion was equally

apparent during extinction, and more robust than freezing. While locomotion was observed across all rats, female locomotion was better timed to shock delivery. Our results join previous studies in demonstrating a fear conditioned cue promotes movement in rats (Bolles and Collier, 1976; Totty et al., 2021).

Jumping is elicited in rats by hypoxia (decreasing oxygen levels in the air) – a life threatening condition (Spiacci et al., 2015). More relevant to our study, the Blanchards observed jumping in defensive settings in rats (Blanchard et al., 1986). In their procedure, a rat was placed at the end of an inescapable hallway, then a human experimenter slowly approached. Rats initially froze when the experimenter was distant (1-2 meters away), but switched from freezing to jumping as the experimenter drew near (<1 meter). Our observation of danger-elicited jumping during fear acquisition, and its preferential expression at the end of danger presentation, mirrored the defensive jumping pattern observed in the Blanchard's study.

Rearing is elicited by visual cues predicting moderate foot shock. Holland (1979) found that a mix of rearing and freezing are acquired to a visual cue paired with low intensity foot shock. A visual cue paired with intense foot shock exclusively produces freezing. The foot shock intensity we used in both experiments (0.5 mA) is more consistent with the low intensities in the Holland (1979) experiment. Dielenberg and McGregor found that rats exposed to a recently worn cat collar, with an opposing box to hide in, show “vigilant rearing” to the cat collar (Dielenberg and McGregor, 2001). Rearing was never observed in a control condition. While we cannot claim vigilance, we find that danger promotes rearing during fear acquisition.

Temporal ethograms revealed that during fear acquisition, jumping and rearing were most prevalent at the end of cue presentation – when foot shock was imminent. This was in contrast to freezing which was prominent during early danger presentation for both females and males, but only shown by males at the end of cue presentation. Though unlike Experiment 1, in which foot shock was present, we found no evidence of danger-elicited jumping and rearing during fear extinction. Because jumping and rearing are vertical behaviors, they may be avoidant or escape responses. The rat is trying to leave the floor before the shock comes on. This interpretation is supported by the finding that these responses peaked just before shock presentation in Experiment 1. Removing foot shock in Experiment 2 may have removed the impetus for avoidance and escape. However, it could be equally likely that these behaviors are more sensitive to context change. Experiment 2 also found that freezing transferred less well to the extinction context in which reward was absent.

Our findings accord well with the PIC theory of defensive behavior (Fanselow and Lester, 1988). PIC theory identifies three defensive modes: pre-encounter (for example, leaving the safety of the nest to forage), post-encounter (predator detected), and circa-strike (predation inevitable or occurring). Analogs to PIC modes are identified in a Pavlovian fear conditioning trial (Fanselow et al., 2019). Pre-encounter mode may correspond to leaving the home cage and being placed in the experimental chamber where foot shocks occur. Post-encounter mode corresponds to presentation of the fear conditioned cue. Circa-strike mode is said to correspond to foot shock delivery. It is argued that circa-strike behaviors (locomotion, jumping, and rearing) are not observed towards the end of danger presentation because rats do not time shock delivery. In

support, extending cue duration in traditional cued fear conditioning paradigms does not result in shifts from freezing to locomotion, jumping, and rearing towards cue offset (Fanselow et al., 2019).

Here we find expected patterns of defensive behavior in unexpected epochs of Pavlovian conditioning trials. Early danger freezing by all rats (females and males) gives way to a late mix of danger-elicited behaviors that included locomotion, jumping, and rearing (Experiment 1) or late locomotion (Experiment 2). Why do we observe late danger control of circa-strike behaviors? Hunger and the availability of a rewarded action may provide an impetus for shock timing. Timing would allow rats to minimize the display of defensive behaviors and maximize reward seeking. In support, presenting long duration danger cues in a conditioned suppression setting results in timing of fear responding. With experience, rats show little suppression of reward seeking to danger onset, which ramps towards shock delivery (Rosas and Alonso, 1996). Supporting the minimization of defensive behavior in reward settings, foot shocks signaled by danger will strongly suppress reward seeking only early in fear conditioning. Shock-induced suppression quickly wanes and with experience, shock delivery will paradoxically facilitate reward seeking (LaBarbera and Caul, 1976; Strickland et al., 2021). Shock timing information is readily apparent in the ventrolateral periaqueductal gray, a brain region central to defensive behavior (Fanselow, 1993; Carrive et al., 1997; Mobbs et al., 2007; McDannald, 2010; Tovote et al., 2016; Arico et al., 2017). Populations of ventrolateral periaqueductal gray neurons show little responding upon danger presentation, but ramp firing towards shock delivery (Ozawa et al., 2017; Wright and McDannald, 2019; Wright et al., 2019). Our results support the PIC theory of defensive

behavior but demonstrate that the relationship between defensive mode and Pavlovian conditioning trial epoch is not fixed, but depends on the experimental setting.

A secondary goal of Experiment 1 was to compare defensive behaviors elicited by a deterministic, danger cue and a probabilistic, uncertainty cue. In our setting, uncertainty is not simply a diminished version of danger. Indeed, uncertainty only promoted a subset of danger-elicited behaviors: locomotion and jumping. Most surprising was the inability of uncertainty to suppress reward behaviors directed towards the food cup and port. This is particularly puzzling because using suppression ratios, we found uncertainty to produce robust suppression of nose poking. What is going on here? Food cup, port and poke behavior lie on a rewarded action continuum. Food cup means the rat is in the area of food delivery – but is most distal from the rewarded action. Port means the rat is around or in the site of the rewarded action, but only poke requires the rat to be fully engaged in performing the rewarded action (nose all the way into the port). Danger suppresses all reward behavior regardless of proximity to rewarded action. By contrast, uncertainty selectively suppresses reward behavior most proximal to the rewarded action.

By comprehensively quantifying behavior and constructing temporal ethograms, we found a fear conditioned cue to independently control at least six distinct behaviors during fear acquisition and three distinct behaviors during fear extinction. Though our study was exclusively behavioral, we feel our results have implications for investigations into the neural basis of fear learning and the organization of defensive behavior. Most important is that a fear conditioned cue does not simply elicit freezing. Behaviors elicited by a fear conditioned cue are the product of many factors: species, sex, age, behavioral

setting, and experimenter-determined parameters (CS/US type, duration, and intensity; trial number, inter-trial interval, and more). In our view, freezing is a common – not dominant – defensive behavior because the field has favored behavioral settings and experimenter-determined parameters that maximize the expression of “fear” through freezing (McDannald, 2023). Here we show that a relatively simple modification of the rat’s behavioral setting – access to a rewarded action – is sufficient to de-emphasize freezing and promote the expression of many additional behaviors. Most prominent of these is locomotion. Even more, Pavlovian fear conditioning over a baseline of reward seeking reveals a temporally organized sequence of cue-elicited defensive behaviors predicted by PIC theory. The independent expression of these behaviors is appealing for studies attempting to link discrete behavioral sequelae of “fear” to distinct neural circuits, breathing new life into a classic Pavlovian fear conditioning procedure (Estes and Skinner, 1941).

## **Video Captions**

### **Video 1.** Behavior during a single danger trial

Video shows the 75 sequential frames for a danger trial. Frames 1-25 are background and 26-75 are danger cue presentation. Observer judgment is shown in the top right for each frame. The specific trial is 23\_16\_12, (female rat 23, session 16, trial 12).

### **Video 2.** Behavior during a single danger trial

Video shows the 75 sequential frames for a danger trial. Frames 1-25 are background and 26-75 are danger cue presentation. Observer judgment is shown in the top right for each frame. The specific trial is 24\_16\_16, (male rat 24, session 16, trial 16).

### **Video 3.** Behavior during a single danger trial

Video shows the 75 sequential frames for a danger trial. Frames 1-25 are background and 26-75 are danger cue presentation. Observer judgment is shown in the top right for each frame. The specific trial is 5\_16\_11, (female rat 5, session 16, trial 11).

### **Video 4.** Behavior during a single danger trial

Video shows the 75 sequential frames for a danger trial. Frames 1-25 are background and 26-75 are danger cue presentation. Observer judgment is shown in the top right for each frame. The specific trial is 4\_16\_3, (male rat 4, session 16, trial 3).

## **Chapter 3: Neural networks for comprehensive ethograms**

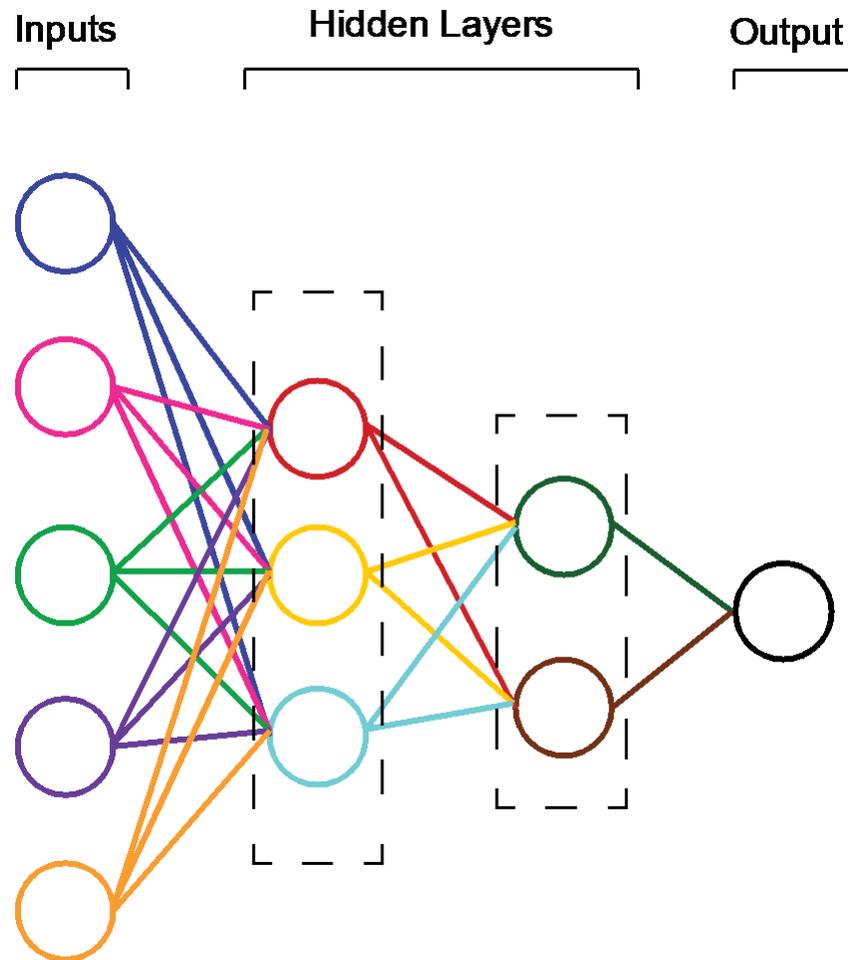
### 3.1 Introduction

As stated in Chapter 1, a central goal of my research, and more broadly, the field of neuroscience is to uncover neural circuits for specific behaviors. Neuroscience laboratories use different behavioral assays to assess emotional and cognitive processes in animals. In practice, measuring behavior in research laboratories poses a challenge to behavioral neuroscientists. To infer emotional and cognitive processes in animals, researchers must measure behaviors that can be objectively and reliably measured. To achieve this level of objectivity and reliability, the richness of behavioral repertoire measured in experiments is often compromised. For example, researchers limit themselves to measuring a single behavior in their experiments (i.e. nose poke), and this behavior is often automatically measured by a photosensor, such as in our laboratory, to facilitate accurate and quick behavioral data collection paired with sophisticated neuroscience techniques (e.g. single-unit recording, optogenetics, and calcium imaging). In this scenario, the behavior is quickly and reliably measured; however, it is at the cost of behavioral complexity. More recently, there is a growing interest in more comprehensive behavior quantification (Chu et al., 2024). However, to observe multiple behaviors in any given experiment, requires a large amount of time and labor to manually annotate an extensive range of behaviors.

To overcome this challenge, several groups have recently developed open-source, machine learning tools to automate the behavior classification process (Mathis et al., 2018; Hsu and Yttri, 2021; Gabriel et al., 2022; Bohoslav et al., 2021). The development of such tools has the potential to contribute to more thorough investigations between behavior and its neural substrates in animal models.

### 3.1.1 Overview of neural networks

Before elaborating on different machine learning pipelines in the following sections, I will first provide a brief overview of neural networks and deep learning. Artificial intelligence (AI) systems, such as speech recognition on our smartphones or personalized show recommendations on our Netflix app, use a technique called “deep learning.” Deep learning is how these AI systems learn, and consists of something called neural networks. A neural network is a model of learning that reflects the way the human brain works and mimics the way neurons are organized within the brain. Neural networks are composed of layers of *nodes*, and each node is analogous to a neuron in the brain. The first layer of a neural network is called the *input layer*, followed by *hidden layers*, and the last layer is referred to as the *output layer* (Figure 3.1). Each node performs a calculation (which can vary based on the type of neural network), and each calculation output is passed onto nodes in the subsequent layer. Neurons or nodes in a neural network are capable of receiving multiple inputs from the preceding layer, and each connection or “synapse” has an associated *weight* ( $w$ ) with it. Weights are important because they are the primary way a neural network is trained.

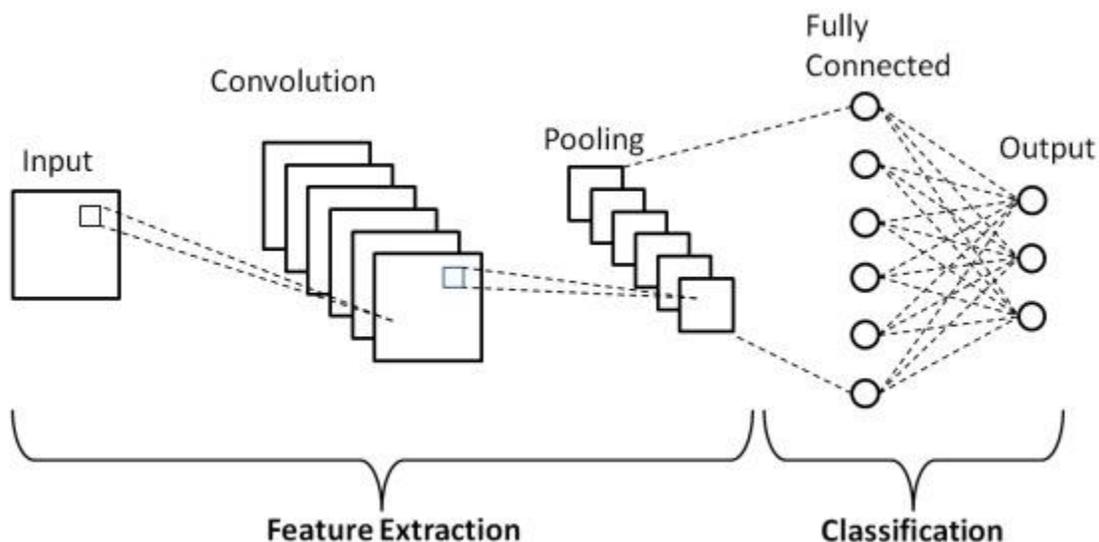


**Figure 3.1.** Diagram of a simple neural network

Schematic of a simple neural network. Each circle represents a neuron or node, which connects to every neuron in the subsequent layer. The connections represent a calculation with an associated weight.

Algorithms such as DeepLabCut and the other tools I will discuss later, use *convolutional neural networks* (CNNs) in their deep learning models. CNNs are commonly used for computer vision tasks such as object and action classification. A convolutional neural network is distinct since it has *convolutional layers*, which are layers that perform a convolution (Figure 3.2). In a convolution operation, a filter or *kernel* containing a square of *weights* slides over the input data (i.e. a pixel image) and

computes multiplication between the weight and input data at each position. The products of all the weights and input data are summed to produce a single output value, which serves as the input to the subsequent layer. In contrast, a *fully connected layer* is when each neuron in one layer is “connected” to every neuron in the subsequent layer, where the dot product of each input neuron and each weight is calculated. Typically, a network consisting of convolutional layers will have a fully connected layer preceding the output layer, to ensure that the output is receiving all the information about different features from previous layers (Figure 3.2). CNNs are ideal for large datasets, since it reduces the number of parameters by sharing weights, unlike a model that uses only fully connected layers, where every node has an associated weight for every single input. Sharing weights in CNNs improves computational efficiency and generalization to new datasets.



**Figure 3.2.** Example of a convolutional neural network

Convolutional neural networks perform a convolution on inputs. Feature extraction involves convolutions and pooling. Classification involves fully connected layers so that the final output layer contains the information from the feature extraction layers. *Source: Phung and Rhee (2019).*

Another important distinction to make is between *unsupervised learning* and *supervised learning*. Some of the machine learning tools I describe in the following sections use unsupervised learning, while the other tools use supervised learning. Unsupervised learning is a method in machine learning where the algorithm learns patterns in unlabeled data, and therefore does not require human-determined behavior labels. In contrast, supervised learning requires human-determined labels that it is trained on, in order to predict behavior labels. It is also important to know that neural networks calculate *losses*. In supervised learning tasks, a loss is the difference between the model's predicted outputs and the actual outputs. The goal of training a neural network is to minimize the loss, and therefore minimize the difference between predicted and actual outputs.

In the next sections, I will discuss four machine learning pipelines developed for automated behavior classification: DeepLabcut, B-SOiD, BehaviorDEPOT, and finally, DeepEthogram. These tools are increasingly being used by behavioral neuroscience laboratories, and inspired my own unique approach to building a machine learning workflow in our research laboratory.

### 3.1.2 *DeepLabCut*

Mathis et al. (2021) developed DeepLabCut (DLC) to track animal poses in various behavioral settings. DLC is a deep neural network (DNN) that extracts posture data from each frame. Its network architecture consists of a pre-trained deep residual network containing 50 layers (ResNet50) and deconvolutional layers. The ResNet50 model was pre-trained on ImageNet, a publicly available database containing over a

million images of various objects such as a pencil, keyboard, and types of animals. Deconvolutional layers are used to “upsample” or increase the image resolution after the resolution is decreased due to the convolutions performed in the ResNet50 layers. A single image (size 224x224) containing the user’s region-of-interest (i.e. a paw) that are manually labeled with body parts (i.e. digit joints and wrists) serves as input to the DNN and produces a distinct readout layer per body part that predicts the probability that a body part is in a particular pixel. It is important to note that DLC does not require users to train the DNN, as their network is already pre-trained.

As mentioned above, users must manually set body part-tracking points. DLC is able to accurately (>90%) estimate body postures in video frames. A drawback of DLC is that although it automatically extracts locations of body parts from videos, it does not automatically categorize each set of poses into specific behavior labels, leaving it up to the user to manually assign behavior labels to a set of poses. For this reason, DLC was not ideal for my research purposes since manually annotating body poses with specific behaviors would be almost as time-consuming as hand annotation of raw images.

### 3.1.3 *B-SOiD*

To overcome this potential hurdle for practical use in behavioral neuroscience, Hsu and Yttri (2021) developed B-SOiD, an unsupervised machine learning algorithm that extracts pose estimation data to predict behaviors. B-SOiD utilizes a common method used in unsupervised learning, called clustering, in order to detect underlying patterns in the dataset and organize the data points based on similarities. Specifically, B-SOiD uses uniform manifold approximation projection (UMAP) for dimension

reduction and clustering of the spatiotemporal relationships of poses. Then, it uses a technique called hierarchical density-based spatial clustering of applications (HDBSCAN) to separate the UMAP clusters. In other words, B-SOiD takes a high-dimensional dataset, reduces the dimensions, and subsequently, clusters the data points and separates the clusters based on similarity. These UMAP clusters then serve as inputs into their unsupervised machine learning algorithm for training. Finally, this trained algorithm can be used for behavior classification of brand new datasets. Thus, B-SOiD uses unsupervised machine learning to determine pose relationships in a given dataset that then predicts behavior labels for individual video frames, providing a potential link between pose estimation and behavior classification. Since B-SOiD uses unsupervised machine learning to determine behaviors given pose information, it will discover potentially new behaviors in any given dataset. For my research project, I wanted to set and define my own list of behaviors and train neural networks on these specific behaviors, rather than discover new behaviors in our data. Therefore, B-SOiD was not the ideal approach for our research needs.

#### *3.1.4 BehaviorDEPOT*

Gabriel et al. (2022) developed BehaviorDEPOT, a behavioral analysis software package that is based on pose tracking. It uses DLC for the pose estimation portion, and assigns each pose to a heuristics that detect behaviors. Unlike other similar resources, BehaviorDEPOT provides some heuristics that detect behaviors tailored for common behavioral assays used in neuroscience such as fear conditioning and open-field test. By organizing and saving behavioral data, it claims to be user-friendly and compatible for subsequent behavioral analysis. BehaviorDEPOT contains different

behavior analysis functions for specific behavioral paradigms such as velocity-based freezing heuristic, novel object exploration heuristic, and open field test analysis function. These built-in heuristics and functions were trained on animal behavioral datasets that were hand-scored by human raters. Additionally, these functions are useful to groups that utilize these specific behavioral paradigms in their research, as they provide a Graphical User Interface (GUI) to set various parameters within the behavior setting. However, it is unclear whether BehaviorDEPOT is optimal for behavioral assays such as our fear discrimination procedure, where we seek to record multiple behaviors in one setting. Authors show that BehaviorDEPOT works well for behavioral paradigms where the researcher tracks only one or two behaviors-of-interest (i.e. freezing versus not freezing in a traditional fear conditioning paradigm).

### 3.1.5 *DeepEthogram*

Unlike DeepLabCut, B-SOiD, and BehaviorDEPOT, DeepEthogram (Bohnslav et al., 2021) is an open-source software that uses raw pixels, instead of pose estimations, to automatically classify behavior via supervised learning and deep convolutional neural networks. The user must manually label behavior frames using their GUI in order to train DeepEthogram's deep neural networks. After hand annotation and training, the user then can use DeepEthogram to automatically predict labels for unlabeled behavior frames. DeepEthogram achieved >85% accuracy across datasets collected in various behavioral paradigms (e.g. mouse open-field test, elevated-plus maze, homecage, etc.), with each paradigm consisting between 3-8 total behaviors scored. Authors show that DeepEthogram can be used to predict multiple behaviors such as freezing, locomotion, grooming, and more. The software is made in such a way that there are no real limits to

the number of behaviors-of-interest the researcher can measure, although the researchers test DeepEthogram's efficacy with less than 10 behaviors-of-interest for each animal experiment. DeepEthogram was a promising approach for our laboratory's uses; however, to a behavioral neuroscientist with minimal computational knowledge, it was an extremely time-consuming effort to install and use DeepEthogram practically. At the beginning of this endeavor, DeepEthogram had just been made publicly available and had not yet been widely used by many neuroscience laboratories. During this time, my attempts to install and use DeepEthogram were much easier in theory than in practice. Instead, I opted to develop and train a deep neural network from scratch, in order to gain better knowledge of how neural networks work.

### *3.1.6 Chapter aim*

For my research, I was particularly interested in looking at the full range of behaviors expressed during our laboratory's fear discrimination procedure. As emphasized in Chapter 2, it is clear that rats not only freeze in response to a fear conditioned cue, but they also express activity-promoting behaviors – most notably locomotion, jumping, and rearing. Given this richness in behavioral responding, it is important to our laboratory that future projects look at the full breadth of behaviors expressed in our behavioral setting. However, the quantification of nine discrete behaviors in Experiments 1 and 2 from Chapter 2 was a Herculean task, requiring training of several raters. Hand scoring over 100,000 frames from the experiments in Chapter 2 took many months in total, with multiple raters. In order to improve efficiency for future behavior-focused projects in our laboratory, I aimed to establish a machine learning pipeline, by creating and training convolutional neural networks (CNNs) that

automatically generate behavior labels for individual frames. With the help of my collaborator, Dr. Stefano Anzellotti, we set out to develop and train CNNs to classify the nine discrete behaviors (plus background) from the experiments mentioned in Chapter 2.

Given the recent developments in automation of animal behavior classification, the goal of this chapter was inspired by the modeling approaches used by the aforementioned groups, especially DeepEthogram (Bohnslav et al., 2021). In this chapter, I will describe how we extract feature data from raw pixels of a single frame and motion data from sequential frames in order to generate a single behavior label output. We used similar neural network models from DeepLabCut and DeepEthogram to create and train deep convolutional neural networks for automated behavior scoring. Our model consists of a deep residual network containing 18 layers (ResNet18) and an optical flow neural network (TinyMotionNet) to generate behavior labels.

The aim described in this chapter was to develop and train convolutional neural networks to accurately quantify our laboratory's behaviors of interest. In order to do so, our model must achieve two things: 1) achieve test accuracy levels that match that of our human observers (79 - 83% for 10 behaviors), and 2) generalize well to a new dataset.

## 3.2 Materials and Methods

### 3.2.1 Datasets

86,400 hand-scored frames collected from 24 rats (see Chapter 2, Materials and Methods) served as the training dataset (22 animals) and testing dataset (2 animals) for the neural networks. Each frame was a 640x480 pixels JPEG image that varied slightly in brightness throughout the dataset.

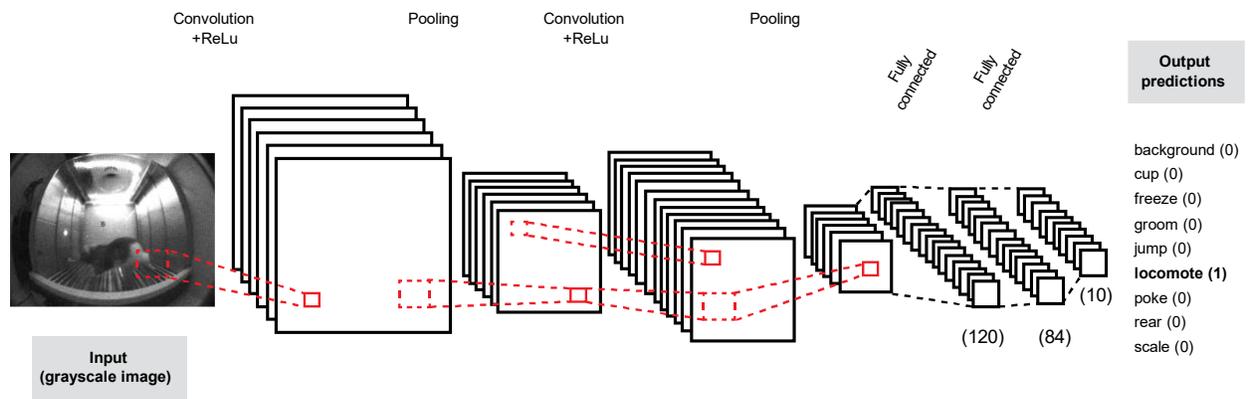
### 3.2.2 Feature information from basic convolutional neural network

First, we created a basic CNN consisting of 2 convolutional layers and 3 fully connected (FC) layers (Figure 3.3). For this basic CNN, we used *maxpooling2d* between the convolutional layers. Briefly, max pooling reduces the dimensions of the output of one convolutional layer; this reduced output then serves as the input to the next convolutional layer. Max pooling is useful since it reduces the computational load when training the model and also may reduce *overfitting* (when a model's predictions are too specific to the training dataset and does not generalize well to new data). Additionally, *rectified linear unit function* (or ReLU function) was used to perform non-linear transformations in order to recognize non-linear relationships between input and output variables within the data. Briefly, ReLU function takes in positive and negative input values, and returns zero for any negative values, leaving only zero and positive values. This mathematical computation allows for more efficient and better learning during training.

Lastly, this model consists of 3 fully connected layers. The input to the first FC layer is the output from the second convolutional layer; the output of the first FC layer is

the input to the second FC layer; and lastly, the input to the third FC layer is the output of the second FC layer, and its output corresponds to the number of output classes (the 10 behavior labels). As mentioned in the introduction of this chapter, fully connected layers are usually used in CNNs as the final layer, so that the network's output is based on all the features from previous layers. In this model, the third FC layer is the output layer.

This model was trained on 79,200 static images from 22 animals (see Chapter 2, Experiment 1). Feature information was extracted from raw pixels from each image during training. After training, the model was then tested on 7,200 images from 2 different animals from the same dataset and test accuracy (percentage) was calculated.



**Figure 3.3.** Basic CNN model

The network architecture of the basic CNN model. Input is a single image (640x480 JPEG) that undergoes convolutions with ReLu functions followed by pooling (maxpooling2d). The final layer is the predicted probabilities for each of the 10 behavior labels.

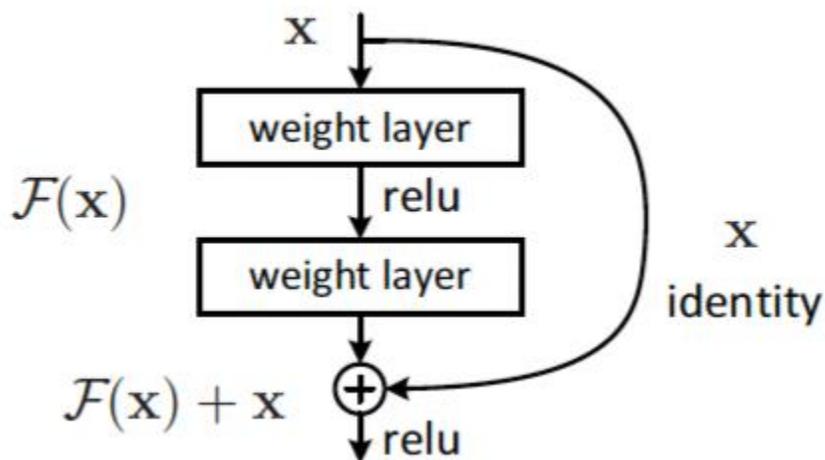
### 3.2.3 Feature information from ResNet18

We also created and trained a deep residual neural network containing 18 layers called ResNet18 (see He et al., 2016). Briefly, the convolutional layers in this model undergo *residual learning*, which essentially performs mathematical computations that “skip” layers, so that the input of one layer is the sum of the input to the residual block plus the residual from “skipping” the subsequent layers (also known as the residual mapping) (Figure 3.4).

$H(x)$  is the underlying mapping where  $x$  is the input

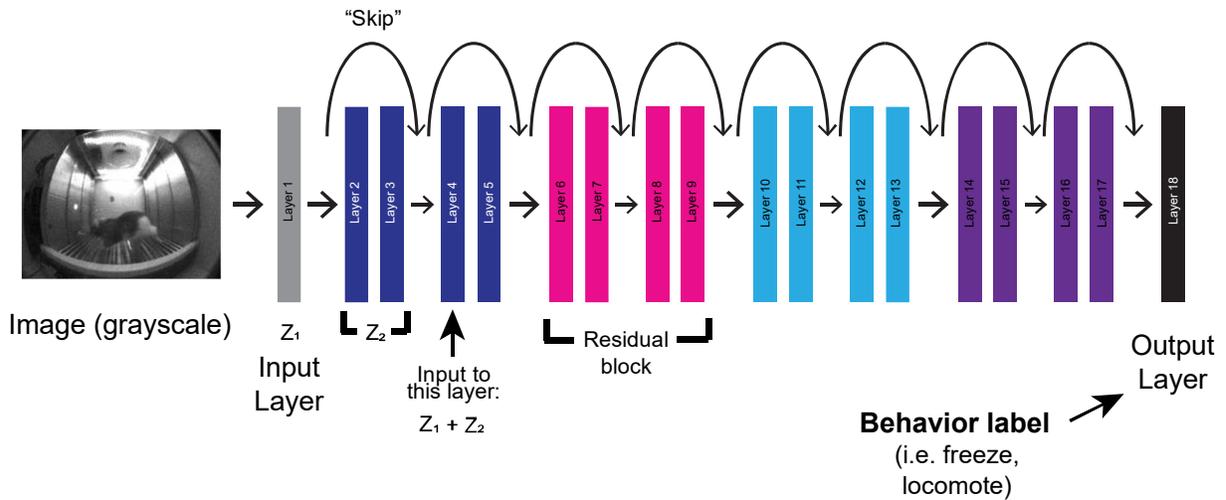
Residual mapping or “skip”:  $F(x) = H(x) - x$

Output of the residual block:  $F(x) + x$  where  $x$  is the input to the block



**Figure 3.4.** Residual block

A schematic of a residual block is shown. Source: He et al. (2016).



**Figure 3.5.** ResNet18 model

Input to the first layer is one grayscale image. Residual learning occurs in the subsequent layers, made up of residual blocks. The output layer consists of the most probable behavior label predicted by the model.

Residual learning addresses two common issues in DNNs: 1) *the degradation problem*, which is when the performance of a deep neural network worsens as more layers are added, and 2) *the vanishing gradient problem*, which is when the gradients become increasingly small as they backpropagate through the layers of DNNs during training and thus, the layers stop learning. Gradients are related to the loss function (they are the vector of partial derivatives of the loss function); so while minimizing losses and gradients is the ultimate goal of neural networks, it becomes a problem if the gradient becomes too small as it makes the network harder to train. Residual learning helps mitigate these issues and improve learning and thus, overall performance. This model was trained on 79,200 static images from 22 animals, then tested on 7,200 images from 2 different animals from the same dataset (see above).

### 3.2.4 Motion information from optical flow estimation

We trained an optical flow neural network using TinyMotionNet architecture (Zhu et al., 2018). Briefly, this convolutional neural network extracts motion information by measuring the displacement, pixel-by-pixel, between two sequential frames for all the frames within a single trial – this is the *optical flow estimation* that is measured between two frames. To calculate the optical flow, the model uses a reconstruction approach, where taking in the displacement of all pixels between two adjacent images ( $I_1$  and  $I_2$ ), it constructs a flow of vectors called  $V$ . Using  $I_2$  and  $V$  as its input, it reconstructs an image of  $I_1$ , called  $I_1'$ . The idea is that if the model can reconstruct the first image  $I_1$  from the second image ( $I_2$ ) and the flow ( $V$ ), then the model has learned useful motion information.

TinyMotionNet consists of a contracting part and an expanding part (Figure 3.6). The contracting part is made up of convolutional layers (reduces the dimensions of the inputs), while the expanding part is made up of convolutional layers and deconvolutional layers (increases the dimensions of the inputs). TinyMotionNet first focuses on small displacement motion using a 3x3 kernel. TinyMotionNet computes three types of losses: first, a standard pixel wise reconstruction error function ( $L_{pixel}$ ); second, a smoothness loss function ( $L_{smooth}$ ); and third, a structural similarity (SSIM) loss function ( $L_{SSIM}$ ). The pixel wise reconstruction error function compares the changes in the horizontal (x) and vertical (y) directions of the pixels between two adjacent images. The smoothness loss function compares the changes of the pixels and accounts for the instances when pixels are the same color. Finally, the structural similarity loss function helps the model learn about the structure of the images by comparing the means and variances of image

patches. Note, image patches refer to when the model partitions a whole image into local patches using a sliding window. Thus, it calculates and compares the means and variances between image patches. After the three losses are calculated, the total loss is calculated: the sum of the products of each loss multiplied by their respective weights of importance during training. Details on the architecture of TinyMotionNet and the calculations performed are described in Zhu et al. (2018).

After training, the weights from this motion network serve as additional information to our feature extractor network (ResNet18), in order to improve the prediction accuracy of behaviors that occur over multiple frames (e.g. locomotion, freezing).

Name	Kernel	Str	Ch I/O	In Res	Out Res	Input
conv1	$7 \times 7$	1	33/64	$224 \times 224$	$224 \times 224$	Frames
conv2	$5 \times 5$	2	64/128	$224 \times 224$	$112 \times 112$	conv1
conv3	$3 \times 3$	2	128/256	$112 \times 112$	$56 \times 56$	conv2
conv4	$3 \times 3$	2	256/128	$56 \times 56$	$28 \times 28$	conv3
flow4 (loss4)	$3 \times 3$	1	128/20	$28 \times 28$	$28 \times 28$	conv4
deconv3	$4 \times 4$	2	128/128	$28 \times 28$	$56 \times 56$	conv4
xconv3	$3 \times 3$	1	404/128	$56 \times 56$	$56 \times 56$	deconv3+flow4+conv3
flow3 (loss3)	$3 \times 3$	1	128/20	$56 \times 56$	$56 \times 56$	xconv3
deconv2	$4 \times 4$	2	128/64	$56 \times 56$	$112 \times 112$	xconv3
xconv2	$3 \times 3$	1	212/64	$112 \times 112$	$112 \times 112$	deconv2+flow3+conv2
flow2 (loss2)	$3 \times 3$	1	64/20	$112 \times 112$	$112 \times 112$	xconv2

**Figure 3.6.** TinyMotionNet architecture

The layers of the TinyMotionNet is shown. The convolutional layers comprise the “contracting” part while the deconvolutional layers comprise the “expanding part.”  
*Source: Zhu et al. (2018).*

### 3.3 Results

#### 3.3.1 ResNet18 model achieves proficient overall accuracy but lacks accuracy for specific behaviors of interest

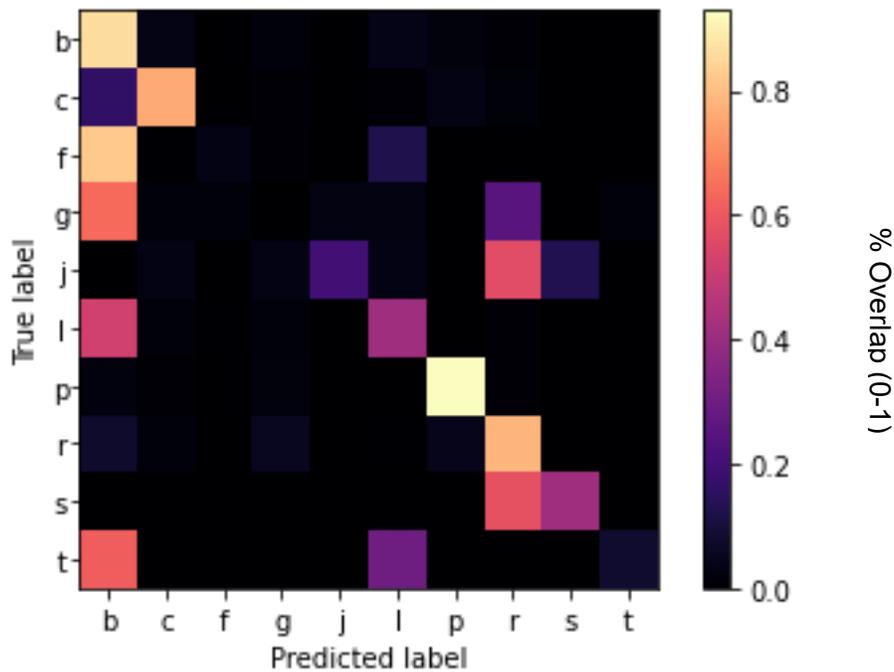
Our initial basic CNN model achieved a test accuracy of 49.8% (Table 3.1). Our subsequent model (ResNet18) achieved a higher test accuracy (between 74 - 75%) with either 10 or 50 epochs and varying weight decay parameters (0.01 or 0.001). Importantly, the loss calculated after each epoch plateaued near or at zero after the 10th epoch (Table 3.1). Since the calculated loss after each epoch plateaued quickly, this implied that the ResNet18 model was ‘overfitting’ our dataset, and thus not generalizing well to new data.

Net Architecture	# of Epochs	Weight Decay	Test Accuracy (%)
Basic CNN	10	0.1	49.8
ResNet18	10	0.1	74.78
ResNet18	50	0.1	74.33
ResNet18	50	0.001	74.25

**Table 3.1.** Test accuracies for CNNs

In order to determine how well the CNN predicted each individual behavior, I constructed a confusion matrix for the ResNet18 model. The confusion matrix (Figure 3.7) compares the percentage overlap between the predicted label, as predicted by the ResNet18 model (x-axis) and the true label, as labeled by human raters (y-axis). There are a handful of behaviors that the ResNet18 model is proficient at predicting (>75% overlap). The model predicts background (86%), cup (76%), poking (93%), and rearing (79%) behaviors very well. However, there are some behaviors that the model predicts

with very low accuracy. For example, freezing (3%), jumping (20%), locomotion (41%), and scaling (42%) behaviors were not well-predicted by the model. In the case of freezing and locomotion, the ResNet18 model often mislabeled these behaviors as “background” (83% of freezing frames and 52% of locomotion frames).



**Figure 3.7.** Confusion matrix for ResNet18 model

The confusion matrix shows the amount of overlap of scored frames between the label predicted by the model (x-axis) and the human rater label, or true label (y-axis). Tick marks on the x- and y-axes are labeled with the 10 behavior labels (b for background, c for cup, f for freeze, g for groom, j for jump, l for locomote, p for poke, r for rear, s for scale, and t for stretch). Light yellow squares indicate 100% overlap or a perfect prediction by the ResNet18 model, while black squares indicates 0% overlap.

### 3.3.2 *Optical flow estimation*

To address the overfitting issue with the ResNet18 model, we sought to devise an optical flow estimation neural network, using the TinyMotionNet architecture (Zhu et al., 2018). Thus far, the data pre-processing is completed and the network has been trained. We are currently in the process of optimizing and implementing the optical flow network into the ResNet18 model in order to improve performance.

### 3.4 Discussion

The goal of this project was to develop and train CNNs that match our human rater accuracy levels from the experiments in Chapter 2. Generally, the models we trained thus far achieved near-human performance. Our ResNet18 model was proficient at predicting certain behaviors like background, cup, poking, and rearing, but lacked proficiency at predicting behaviors like locomotion, freezing, and jumping. Locomotion, freezing, and jumping occur over multiple sequences of frames, as stated in their definitions (see Chapter 2, Table 2.2). Thus, an optical flow estimation network was developed and trained on the same dataset, in order to improve the model's performance of predicting the behaviors that occur over sequential frames. The final step will be to incorporate motion information from the trained optical flow network to the ResNet18 model in order to improve the accuracy of motion-based behaviors (freezing, locomotion, and jumping).

With the implementation of automatic behavior quantification in the laboratory, future projects may efficiently pair comprehensive behavioral ethograms with neuroscience techniques that our laboratory uses such as electrophysiology, optogenetics, and multi-channel recording with NeuroPixels. A huge advantage of using a machine learning approach is that if future studies include more behaviors beyond the scope of the current projects, the neural networks can be trained on new data at any point in time. Thus, a machine learning approach is not only efficient and accurate, but also adaptable for any future behavioral settings, aiding in rigorous and comprehensive behavioral neuroscience research.

## **Chapter 4: Investigating a role for ventral tegmental area dopaminergic neurons in fear expression**

## 4.1 Forward

Before discussing in detail the experiment I conducted in this chapter, there is critical information that needs to be addressed regarding the experimental design and outcomes. The original study I designed included the use of transgenic Th-cre rats, who express cre recombinase specifically in Th<sup>+</sup> (tyrosine hydroxylase, a dopamine marker) cells. The original aim was to infuse cre-dependent viruses encoding an enzyme that produces cell death (cre-Caspase3) or a control yellow fluorophore (cre-YFP) into the VTA of Th-cre rats. The original design consisted of 32 total rats with 16 rats (8 females) in the deletion Casp group and 16 rats (8 females) in the control YFP group. To achieve this, I sought to breed homozygous Th-cre males with wild type females, with the goal of breeding Th-cre pups for use in the experiment. We ran the experiment with the knowledge and assumption that the pups were Th-cre<sup>+</sup> and thus, would express cre recombinase in Th<sup>+</sup> cells.

However, it was found out much later after the conclusion of the experiment, that there was miscommunication about the transgenic status of the male breeders. In reality, the male breeders were hemizygous Th-cre rats (not homozygous) which means that when bred with wild type females, the breeding pairs would produce both wild type and Th-cre rats. This meant that the experimental rats I had used were a mixture of both wild type and transgenic rats. Upon the receipt of this news, we sent extra brain tissue samples I had stored and saved to a genotyping company and confirmed the genetic status of every experimental animal. In the Materials and Methods section of this chapter, you will notice the varied, unbalanced sample sizes of each experimental group and the unintended experimental groups added to the study. The effects of this

circumstance impacted the statistical analyses performed and behavior data. The consequences of this issue will be greatly considered in the discussion of the results later in this chapter.

Additionally, through a software error in the frame collection software, the lighting for the first 4 boxes was noticeably darker than the lighting for the last 4 boxes. This lighting difference was not noticeable to the rats – the chambers were dark through testing. The lighting difference was the result of different gain and exposure settings being used on the cameras in boxes 1-4. To account for possible differences in camera lighting conditions it is used as an MANOVA/ANOVA factor. Collectively, the two issues meant that two unplanned variables needed to be considered during analyses: strain and lighting.

## 4.2 Introduction

Given the suite of behaviors expressed during fear discrimination, a natural question that followed was: what are the neural substrates underlying the expression of these defensive behaviors? Because the brain regions and circuitry involved in freezing have been widely studied in fear conditioning (LeDoux et al., 1988; Kim et al., 1993; Fanselow, 1994; and see Chapter 1), it was especially intriguing to determine the neural underpinnings of the activity-promoting defensive behaviors we observed: locomotion, jumping, and rearing.

Although the neural circuit involved in the expression of conditioned, activity-promoting defensive behaviors in fear conditioning settings remains unclear, there are a few studies that have investigated the brain regions underlying locomotion, jumping, rearing, and avoidance in a variety of fear settings (Schmitt et al., 1986; 1990; Morgan et al., 1998; Kemble, et al., 1990; Muñoz, et al., 2010; Deng et al., 2016; Borkar et al., 2024).

### 4.2.1 *The role of the periaqueductal gray in flight and avoidance*

The periaqueductal gray (PAG) is involved in the expression of “flight” behaviors, such as locomotion and jumping, in different behavioral settings. In a pharmacological study, Schmitt et al. (1986) found that blocking GABA receptors via injection of a GABA receptor antagonist in the medial hypothalamus increased locomotion, rearing, and oriented jumping in rats. Further, the researchers demonstrated that injecting a GABA receptor antagonist into the PAG promoted locomotion as well as non-oriented, erratic-like jumping. It is important to note that in the Schmitt et al. (1986) study, rats did not receive any aversive stimuli, and were observed in a closed chamber after injections of

GABA-manipulating drugs. In the absence of external aversive cues or stimuli, pharmacological blocking of GABA receptors in the PAG was still sufficient to produce locomotion, jumping, and rearing in rats. Although the Schmitt et al. study manipulated GABA receptors within the entirety of the PAG, it has been shown that ventral and dorsal subdivisions of the PAG (vPAG and dPAG) have distinct roles in immobility (freezing) and flight, respectively (Morgan et al., 1998). In another pharmacological study, Morgan and colleagues found that the dPAG is important for active anxiety-like behavior, testing rats in an open-field test. In this study, they activated the dPAG via injection of kainic acid, a glutamatergic receptor agonist, which produced an increase in activity during an open field test, a common test for measuring anxiety in rodents. In contrast, activating the vPAG did not produce the same effect – instead it reduced the overall activity of the rats in the open-field test. In support of Morgan and colleagues, optogenetic stimulation of dPAG neurons promotes flight behavior in mice during a naturalistic threat setting (Deng et al., 2016). In this setting, a mouse (test subject) was placed in an arena with a rat (predator); the two animals were separated by a mesh wall, while the researchers recorded instances of freezing and avoidance in the mouse subjects. Interestingly, in this experiment, Deng and researchers also observed an increase of freezing and avoidance behaviors while optogenetically activating dPAG neurons. They also recorded from dPAG neurons and saw that neuronal firing not only increased in the face of threat (rat predator), but also decreased as a function of the distance away from the predator, suggesting that the dPAG is responsive not merely to motor activity, but to threat proximity. Collectively, these studies suggest that the ventral

PAG likely promotes freezing and immobility, while the dorsal PAG seems to promote flight, or more generally, activity in response to threat in unconditioned fear settings.

#### *4.2.2 The role of cortical and amygdalar structures in flight and avoidance*

There is a body of literature that suggests the amygdala is also a part of the brain circuitry responsible for activity-promoting fear behaviors. It is well established that the amygdala is the center for fear learning and plays an important role in producing conditioned freezing (LeDoux et al., 1988; Fanselow, 1994; LeDoux, 2000); however, there is some evidence that it is likely involved in other fear behaviors. In an unconditioned fear experiment, Kemble and colleagues (1990) revealed that electrolytic lesions to the corticomedial amygdala and basolateral amygdala structures reduced flight behaviors in rats during a naturalistic threat setting. In this naturalistic setting, human experimenters (threat) approached the rat subjects, while researchers measured flight and attack behaviors in the rats. Corticomedial amygdala and basolateral amygdala lesion groups showed significantly less flight behavior, as well as a trend towards reduced attack behaviors compared to controls.

In a Sidman instrumental active avoidance task, Muñoz and colleagues (2010) explored the role of the central, lateral, and basal amygdala in active avoidance. Briefly, a Sidman active avoidance task, also known as two-way active avoidance task, is an instrumental procedure where an animal is placed in a chamber divided into two compartments. Unsignaled shocks are delivered through the floors on one side of the chamber, and animals can successfully avoid the shock if they cross the chamber into the other compartment before the shock is delivered again. Thus, the active avoidance

behavior is reinforced by the absence of the foot shock. In their study, rats received electrolytic lesions to the central, lateral, or basal amygdala before active avoidance training while another group received the lesions after overtraining. The researchers measured the amount rats successfully avoided the foot shock in this instrumental procedure. The main findings of this study were that lesioning the lateral and basal amygdala prior to training significantly impaired active avoidance, while lesioning the central amygdala only had an effect after overtraining – central amygdala lesions abolished freezing and rescued active avoidance in poor performing rats. These results suggest that the lateral and basal structures of the amygdala are involved in the acquisition of instrumental fear learning, while the central amygdala (which is involved in Pavlovian fear learning) is not involved in active avoidance. In support of a role for the amygdala as well as the cortex in flight behaviors, Borkar et al. (2024) recently identified a neural pathway for avoidance and flight behavior in mice. They found that the dorsal peduncular (DP)-to-medial subdivision of the central amygdala (CeM) pathway is both necessary and sufficient for avoidance during an open-field test and flight behavior during a high-threat fear conditioning paradigm. In sum, these findings suggest that the lateral, basal, and central amygdala are a part of the neural pathway that promotes active fear behaviors via cortical inputs.

Overall, there is strong evidence that the dPAG and the amygdala (which likely receive inputs from cortical structures) are involved in the expression of activity-promoting fear behaviors like flight and active avoidance in rodents. It is well known that inputs from the amygdala to the vPAG are critical for freezing, while amygdala's projections to the dPAG are likely promoting flight and avoidance behaviors. However, a

complete pathway for conditioned, activity-promoting fear behaviors remains elusive. In the next section, I will propose the VTA as a region that is part of the larger brain circuit responsible for activity-promoting fear behaviors.

#### *4.2.3 A potential role for dopamine and the VTA in fear behaviors*

I was interested in whether dopaminergic neurons in the VTA contributed to activity-promoting behaviors during fear discrimination for two reasons: first, some studies have shown that manipulating dopamine in the VTA affects the expression of fear behaviors such as fear-potentiated startle and active avoidance (Borowski and Kikkinidis, 1996; Reis et al, 2004; Ribeiro de Oliveira et al., 2008); and second, there is evidence showing that dopaminergic VTA inputs to the CeA and BLA are involved in fear learning (Fadok et al., 2009; Fellingner et al., 2021; Tang et al., 2020).

Before elaborating on midbrain dopamine's role in fear settings, it is important to briefly address its well-established role in other important processes. Since its discovery in the early 1900s, midbrain dopamine has been established in essential processes such as locomotive movement and reward processing and learning. Midbrain dopamine is critical for self-initiated movement, and a deficiency in dopamine levels is a hallmark of Parkinson's disease in humans (Carlsson, 2002). Additionally, VTA dopamine plays an important role in addiction, and more broadly, incentive salience or "wanting" (Vanderschuren and Kalivas, 2000; Berridge, 2007). Because of this, dopamine has been widely studied in reward settings (Schultz, 1998; Stuber et al., 2008; Roesch et al., 2007; Stalnaker et al., 2019). The dopaminergic projections from the VTA to the nucleus accumbens (NAc) is commonly referred to as the "reward circuit" and disruptions to this

circuit is studied in animal models for depression (Russo and Nestler, 2013). In addition, VTA dopaminergic projections to the medial prefrontal cortex (mPFC), specifically the infralimbic cortex (IL), also promote positive reinforcement learning (Han et al., 2017). It is well known that dopaminergic projections from the VTA to the NAc and mPFC comprises the “reward circuit” and is studied primarily in positive reinforcement learning, mood disorders, and addiction settings.

Although the mesolimbic dopamine pathway has been widely studied in reward contexts, there is a growing body of evidence for the involvement of VTA dopaminergic cells in active-like fear behaviors (Borowski and Kikkinidis, 1996; Reis et al, 2004; Ribeiro de Oliveira et al., 2008). For example, VTA dopamine is involved in producing the startle response in the fear-potentiated startle paradigm (FPS) (Borowski and Kikkinidis, 1996). The startle response in rodents is a reflexive-like response characterized by swift jumping in response to a fear conditioned light cue paired with an unconditioned auditory tone. In another FPS experiment, Ribeiro de Oliveira et al. (2008) show that VTA dopaminergic neurons are involved in the expression of FPS, but not in the acquisition of FPS. Specifically, they saw that pharmacological manipulations in the VTA, using a dopamine receptor subtype 2 (D2) agonist, reduced FPS in mice by decreasing dopamine levels at the terminals.

Additionally, dopamine receptors (subtypes D1 and D2) are involved in the acquisition of active avoidance response (Reis et al., 2004). Reis and colleagues demonstrated that systemic injections of a D2 receptor agonist prior to two-way active avoidance training increased avoidance, while D1 and D2 receptor antagonists decreased the level of avoidance responding. These studies indicate that dopamine in

the VTA contributes to the expression of activity-promoting defensive behaviors such as the startle response and active avoidance.

#### *4.2.4 The role of dopamine and the VTA in fear learning*

Another reason why VTA dopaminergic cells were of interest is because this region and cell type have also been implicated in fear learning in various fear paradigms (Nader and LeDoux, 1999; Fadok et al., 2009; Fellingner et al., 2021; Tang et al., 2020; Jo et al., 2018). More specifically, there is evidence that some dopaminergic cells in the VTA project to the CeA and BLA, and that this pathway modulates fear learning.

In a second-order fear conditioning paradigm, Nader and LeDoux (1999) investigated the effects of inhibiting the dopaminergic pathway from the VTA to the lateral and basal amygdala. They found that activating D2 receptors within the VTA or inhibiting D1 receptors within the lateral and basal amygdala decreased freezing to a second-order cue. This finding implicates a role for dopaminergic projections from the VTA to the lateral and basal amygdala in modulating the retrieval of cue-shock association. Fellingner et al. (2021) found that dynorphin signaling in the VTA during fear conditioning impaired discrimination and also increased anxiety-like behaviors in an elevated-plus maze test. Although the Fellingner et al. study looked at dynorphin instead of dopamine in the VTA, this study still highlights the VTA as a region involved in fear learning and behavioral expression.

Studies manipulating dopamine cell populations in the VTA and its projections to the BLA reveal that this pathway is important in fear learning. For example, during FPS in mice, Fadok and colleagues (2009) demonstrate that activation of D1 receptors are

necessary for the acquisition of FPS. Moreover, their findings reveal that dopaminergic projections from the VTA to the BLA facilitate short-term fear memory. Further validating the VTA→BLA pathway's involvement in fear memory, BLA-projecting dopaminergic neurons in the VTA reflect the saliency of foot shock in order to form the fear memory during fear conditioning (Tang et al., 2020). Additionally, it has been shown that distinct dopamine populations in the VTA respond differently to threat-predicting cues and also modulate fear discrimination via projections to the CeA during fear conditioning (Jo et al., 2018). Jo and colleagues found that there are two distinct populations of dopaminergic neurons in VTA: one population that is inhibited by a threat-predicting cue and the other population that is activated by a threat-predicting cue. Further, the researchers demonstrate that optogenetically inhibiting dopaminergic projections from the VTA to the CeA impaired fear discrimination by increasing freezing responses to a safety cue (CS-). The findings of these studies heavily implicate VTA dopaminergic neurons to contribute to fear learning, most likely via their projections to the CeA and/or BLA.

#### *4.2.5 The role of dopamine and the VTA in fear extinction and safety learning*

Lastly, VTA dopaminergic neurons are also involved in fear extinction and safety learning. Dopaminergic projections from the VTA to the medial shell of the NAc facilitate fear extinction during shock omission (Luo et al., 2018). A study by Salinas-Hernández and colleagues (2023) supports the findings from Luo et al. and further proposes a neural circuit responsible for fear extinction. They suggest that the prediction error signal is generated in the dorsal raphe nucleus, which sends projections to dopaminergic cell populations in the VTA, and the VTA-to-anteromedial NAc pathway

encodes the prediction error, all to facilitate fear extinction. Similarly, activity in medial VTA dopaminergic neurons during shock omission predicts better safety learning (Yau and McNally, 2022). fMRI studies performed in humans reflect the findings from rodent studies (Esser et al., 2021; Frick et al., 2022). Researchers show that administering L-DOPA (the precursor to dopamine) in humans strengthens the connectivity between the VTA, NAc, and the BLA during fear extinction and facilitates extinction learning (Esser et al., 2021). Moreover, Esser et al. (2021) supports evidence from rodent studies that the VTA and BLA are involved in fear and extinction learning. Relatedly, Frick and colleagues found that dopamine-induced long-term potentiation (LTP) in the amygdala promoted better fear learning in humans. Altogether, these findings confirm a role for the VTA and dopamine (likely by way of their projections to the amygdala) in fear extinction and safety learning in both humans and animals.

#### 4.2.6 Chapter aim

Generally, the findings of the studies I described above implicate the VTA and dopamine in fear learning and behavioral responding in a variety of fear settings, including both conditioned and unconditioned settings. However, there is no clear role of VTA dopaminergic neurons in *conditioned* defensive behaviors during fear learning. In this chapter, I will elaborate on the experiment I conducted to determine whether dopaminergic neurons within the VTA contribute to the expression of the defensive behaviors observed in our fear discrimination procedure. In order to assess the contribution of VTA dopamine in the expression of defensive behaviors, I deleted dopaminergic neurons in the VTA prior to fear discrimination and measured nine discrete behaviors during discrimination and extinction.

## 4.3 Materials and Methods

### 4.3.1 Subjects

Long Evans male Th-cre rats, Hemizygous LE-Tg(TH-Cre)<sup>3.1</sup>Deis rats (Th-cre), were bred with Long Evans wild type (WT) females. From these litters, thirty-two rats (16 female) were selected for experimentation. Rats were single-housed on a 12-hr light cycle (lights off at 6:00pm) and maintained at their initial body weight with standard laboratory chow (18% Protein Rodent Diet #2018, Harlan Teklad Global Diets, Madison, WI). Water was available *ad libitum* in the home cage. All protocols were approved by the Boston College Animal Care and Use Committee and all experiments were carried out in accordance with the NIH guidelines regarding the care and use of rats for experimental procedures.

### 4.3.2 Stereotaxic surgery

Once rats reached adulthood (P55), animals underwent stereotaxic surgery. Animals were anesthetized with oxygen isoflurane (1-5%). For the WT x Casp group (n=8, 4 females) and Th-cre x Casp group (n=8, 4 females), rats received bilateral infusions of rAAV5-Flex-taCasp3-TEVp in the VTA at the following coordinates: (from Bregma) AP - 5.3 mm, ML +/- 0.7 mm, DV -8.6 mm and -8.1 mm (0.5 ul at each DV level, 1.0 ul total for one hemisphere). For the WT x YFP group (n=10, 3 females) and Th-cre x YFP group (n=6, 5 females), rats received bilateral infusions of a cre-dependent adeno-associated virus encoding enhanced yellow fluorophore (rAAV-EF1a-DIO-eYFP) in the VTA, with the same volume and coordinates as the deletion group.

#### 4.3.3 *Behavior apparatus*

The apparatus for Pavlovian fear discrimination consisted of four individual chambers with aluminum front and back walls, clear acrylic sides and top, and a grid floor. LED strips emitting 940 nm light were affixed to the acrylic top to illuminate the behavioral chamber for frame capture. 940 nm illumination was chosen because rats do not detect light wavelengths exceeding 930 nm (Nikbakht and Diamond, 2021). Each grid floor bar was electrically connected to an aversive shock generator (Med Associates, St. Albans, VT). An external food cup, and a central port equipped with infrared photocells were present on one wall. Auditory stimuli were generated with an Arduino-based device and presented through two speakers mounted on the ceiling.

#### 4.3.4 *Pellet exposure and nose poke shaping*

Rats were food-restricted and specifically fed to maintain their body weight throughout behavioral testing. Each rat was given four grams of experimental pellets in their home cage in order to overcome neophobia. Next, the central port was removed from the experimental chamber, and rats received a 30-minute session in which one pellet was delivered every minute. The central port was returned to the experimental chamber for the remainder of behavioral testing. Each rat was then shaped to nose poke in the central port for experimental pellet delivery using a fixed ratio schedule in which one nose poke into the port yielded one pellet. Shaping sessions lasted 30 min or until approximately 50 nose pokes were completed. Each rat then received 6 sessions during which nose pokes into the port were reinforced on a variable interval schedule. Session 1 used a variable interval 30 s schedule (poking into the port was reinforced every 30 s on average). All remaining sessions used a variable interval 60 s schedule. For the

remainder of behavioral testing, nose pokes were reinforced on a variable interval 60 s schedule independent of cue and shock presentation.

#### *4.3.5 Cue pre-exposure*

Each rat was pre-exposed to the two cues to be used in Pavlovian discrimination in one session. Auditory cues consisted of repeating motifs of broadband click and phaser. This 48 min session consisted of four presentations of each cue (8 total presentations) with a mean inter-trial interval (ITI) of 3.5 min. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

#### *4.3.6 Pavlovian fear discrimination*

Each rat received twelve, 48-minute sessions of fear discrimination. Each session consisted of 8 trials, with a mean ITI of 3.5 min. Auditory cues were 10 s in duration. Each cue was associated with a unique foot shock probability (0.5 mA, 0.5 s): danger,  $p=1.00$  and safety,  $p=0.00$ . Foot shock was administered 2 s following the termination of the auditory cue on danger trials. Auditory identity was counterbalanced across rats. Each session consisted of four danger trials and four safety trials. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

#### *4.3.7 Fear extinction*

Each rat received an extinction test. Extinction sessions were 48 minutes in duration, and consisted of 8 trials, with a mean ITI of 3.5 min. Auditory cues were 10 s in duration. Foot shocks were not delivered after danger cue termination. Auditory identity was counterbalanced across rats. Each session consisted of four danger trials and four

safety trials. Trial type order was randomly determined by the behavioral program and differed for each rat, each session.

#### *4.3.8 Calculating suppression ratio*

Time stamps for cue presentations, shock delivery, and nose pokes (photobeam break) were automatically recorded by the Med Associates program. Baseline nose poke rate was calculated for each trial by counting the number of pokes during the 20-s pre-cue period and multiplying by 3. Cue nose poke rate was calculated for each trial by counting the number of pokes during the 10-s cue period and multiplying by 6. Nose poke suppression was calculated as a ratio:  $(\text{baseline poke rate} - \text{cue poke rate}) / (\text{baseline poke rate} + \text{cue poke rate})$ . A suppression ratio of '1' indicated complete suppression of nose poking during cue presentation relative to baseline. A suppression ratio of indicated '0' indicates equivalent nose poke rates during baseline and cue presentation. Gradations in suppression ratio between 1 and 0 indicated intermediate levels of nose poke suppression during cue presentation relative to baseline. Negative suppression ratios indicated increased nose poke rates during cue presentation relative to baseline.

#### *4.3.9 Frame capture system*

Behavior frames were captured using Imaging Source monochrome cameras (DMK 37BUX28; USB 3.1, 1/2.9" Sony Pregius IMX287, global shutter, resolution 720x540, trigger in, digital out, C/CS-mount). Frame capture was triggered by the Med Associates behavior program. The 28V Med Associates pulse was converted to a 5V TTL pulse via Adapter (SG-231, Med Associates, St. Albans, VT). The TTL adapter was wired to the camera's trigger input. Captured frames were saved to a PC (OptiPlex 7470 All-in-One)

running IC Capture software (Imaging Source). Frame capture began 5 s before cue onset and continued throughout 10-s cue presentation and 5 s after cue termination. Frames were captured at a rate of 5 per second, with a target of capturing 100 frames per trial (5 frame/s x 20s = 100 frames), and 800 frames per session (100 frames/trial x 8 trials = 800 frames).

#### *4.3.10 Post-acquisition frame processing*

We aimed to capture 800 frames per session, and selected discrimination session 12 and extinction session 1 to hand score. A MatLab script sorted the 800 frames into 8 folders, one for each trial, each containing 100 frames. Each 100-frame trial was made into a 100-slide PowerPoint presentation to be used for hand scoring.

#### *4.3.11 Anonymizing trial information*

A total of 512 trials of behavior were scored from the 32 rats during session 12 of discrimination and 1 extinction session (8 trials per session). We anonymized trial information in order to score behavior without bias. The numerical information from each trial (session #, rat # and trial #) was encrypted as a unique number sequence. A unique word was then added to the front of this sequence. The result was that each of the 512 trials was converted into a unique word+number sequence. For example, trial dx01\_12\_07 (rat #1, session #12 and trial #7) would be encrypted as: abundant28515581. The 512 trials were randomly assigned to 1 of the 7 observers. The result of trial anonymization was that observers were completely blind to subject, trial type, and session number. Further, random assignment meant that the 8 trials composing a single session were scored by different observers.

#### 4.3.12 Behavior categories and definitions

Frames were scored as one of ten mutually exclusive behavior categories, defined as follows:

Background. Specific behavior cannot be discerned because the rat is turned away from the camera or position of forepaws is not clear, or because the rat is not engaged in any of the other behaviors.

Cup. Any part of the nose above the food cup but below the nose port.

Freeze. Arched back and stiff, rigid posture in the absence of movement, all four limbs on the floor (often accompanied by hyperventilation and piloerection). Side to side head movements and up and down head movements that do not disturb rigid posture are permitted. Activity such as sniffing or investigation of the bars is not freezing. Freezing, as opposed to pausing, is likely to be 3 or more frames (600+ ms) long.

Groom. Any scratching, licking, or washing of the body.

Jump. All four limbs off the floor. Includes hanging which is distinguished when hind legs are hanging freely.

Locomote. Propelling body across chamber on all four feet, as defined by movement of back feet. Movement of back feet with front feet off the floor is rearing. Locomote considers the current frame (t) and the next two frames (t+1 and t+2). By frame t+2 the body and both back feet must be displaced forward relative to frame t. The rat can move the body and both feet by t+1, t+2 or move in combination of both frames; all count as locomote for trial t.

Port. Any part of the nose in the port. Often standing still in front of the port but sometimes tilting head sideways with the body off to the side of the port.

Rear. One or two hind legs on the grid floor with both forepaws off the grid floor and not on the food cup. Usually (not always) stretching to full extent, forepaws usually (not always) on top of side walls of the chamber, often pawing walls; may be accompanied by sniffing or slow side-to-side movement of head. Does not include grooming movements or eating, even if performed while standing on hind legs.

Scale. All four limbs off the floor but at least two limbs on the side of the chamber. Standing on the food cup counts as scaling.

Stretch. Body is elongated with the back posture 'flatter' than normal. Stretching is often accompanied by immobility, like freezing, but is distinguished by the shape of the back.

#### *4.3.13 Frame scoring system*

Frames were scored using a specific procedure. Frames were first watched in real time in Microsoft PowerPoint by setting the slide duration and transition to 0.19 s, then playing as a slideshow. Behaviors clearly observed were noted. Next, the observer went through the 75 all the frames scoring one behavior at a time. A standard scoring sequence was used: port, cup, rear, scale, jump, groom, freeze, locomote, and stretch. When the specific behavior was observed in a frame, that frame was labeled. Once all behaviors had been scored, the video was re-watched for freezing. The unlabeled frames were then labeled 'background'. Finally, all background frames were checked to ensure they did not contain a defined behavior.

#### *4.3.14 Inter-observer reliability*

To assess inter-observer reliability, we selected 10 trials from outside session 12 of discrimination and extinction session, five from females and five from males. Half of the comparison trials (5) were from light trials and half were from dark trials. Each of our seven observers scored these 10 trials, interweaving the 10 comparison trials with the primary data trials. As a result, each observer scored 1,000 comparison frames which were then used to assess inter-observer reliability.

#### *4.3.15 Statistical analyses*

Analysis of variance (ANOVA) was performed for baseline nose poke rate, suppression ratios, and specific behaviors. Sex was used as a factor for all analyses. Cue, session, and time were used as factors when relevant. Univariate ANOVA following MANOVA used a Bonferroni-corrected p value significance of 0.0055 (0.05/9) to account for the nine quantified behaviors. Multiple analysis of variance (MANOVA) was performed for the nine quantified behaviors with factors of sex, cue, and time. Pearson's correlation coefficient was used to examine the relationship between baseline nose poke rate and body weight, baseline nose poke rate and cue discrimination, as well as the relationship between danger cue-elicited behaviors during early and late cue presentation in session. Within-subject comparisons were made using 95% bootstrap confidence intervals with the Matlab bootci function. Comparisons were said to differ when the 95% bootstrap confidence interval did not contain zero. Between subject's comparisons were made using independent samples t-test.

#### *4.3.16 Perfusion and tissue collection*

After the final behavior session, animals were overdosed on isoflurane and transcardially perfused with 0.9% biological saline and 10% buffered formalin. Brains were post-fixed in 10% buffered formalin with 2% sucrose. Brains were flash frozen with dry ice and stored in -80°C freezer prior to slicing. All brains were subsequently sliced coronally (40 microns) and stored long-term in cryoprotectant at 4°C.

#### *4.3.17 Histology*

Anterior and posterior VTA sections (from Bregma: -5.30 and -6.28 AP) from each animal were selected for anti-Th fluorescent immunohistochemical staining. Briefly, coronal sections were washed in 0.05 M kPBS then incubated overnight at 4°C in primary blocking solution: Rabbit Anti-TH 1° (Millipore AB152), 0.05 M kPBS-T with 0.1% triton and 3% normal donkey serum. After 0.05 M kPBS washes, tissue were incubated for 1 hour in the dark at room temperature in secondary blocking solution: Donkey, Anti-Rabbit 2° Alexa 594 (AffiniPure Jackson Immuno 711-585-152), 0.05 kPBS-T with 0.1% triton and 2% normal donkey serum. After final 0.05 M kPBS washes, brain sections were mounted on slides and coverslipped with VECTASHIELD HardSet Antifade Mounting Medium (H-1400) for fluorescent imaging.

#### *4.3.18 Cell quantification*

All cell counting was performed with ImageJ software. Slides containing brain sections were imaged with Texas Red channel for visualization of Th+ cells or with YFP channel for visualization of YFP+ cells on the Zeiss AxioImager Z2 (Boston College Imaging Facility). On ImageJ, each image was first converted to an 8-bit image (black and white) prior to modifications. Then, thresholds for each image were set appropriately to limit

background fluorescence. Gaussian Blur was set for each image and Sigma Radius was set to 2.00. Each image was made binary, set to 'fill holes', and set to 'watershed.' Finally, using the freehand selection tool, each subregion (e.g. left substantia nigra, right VTA) were hand-selected before analyzing particles with the following settings: Size, 50-infinity and circularity, 0.5-1.00.

#### *4.3.19 Genotyping*

Genotyping of animals was performed by Transnetyx, Inc. Brain tissue samples were washed in 0.05 M kPBS then dipped in 95% ethanol for 5 seconds and let dry under the fume hood overnight. Dried tissue samples were sent in for automated genotyping analysis by Transnetyx, Inc. the following day.

## 4.4 Results

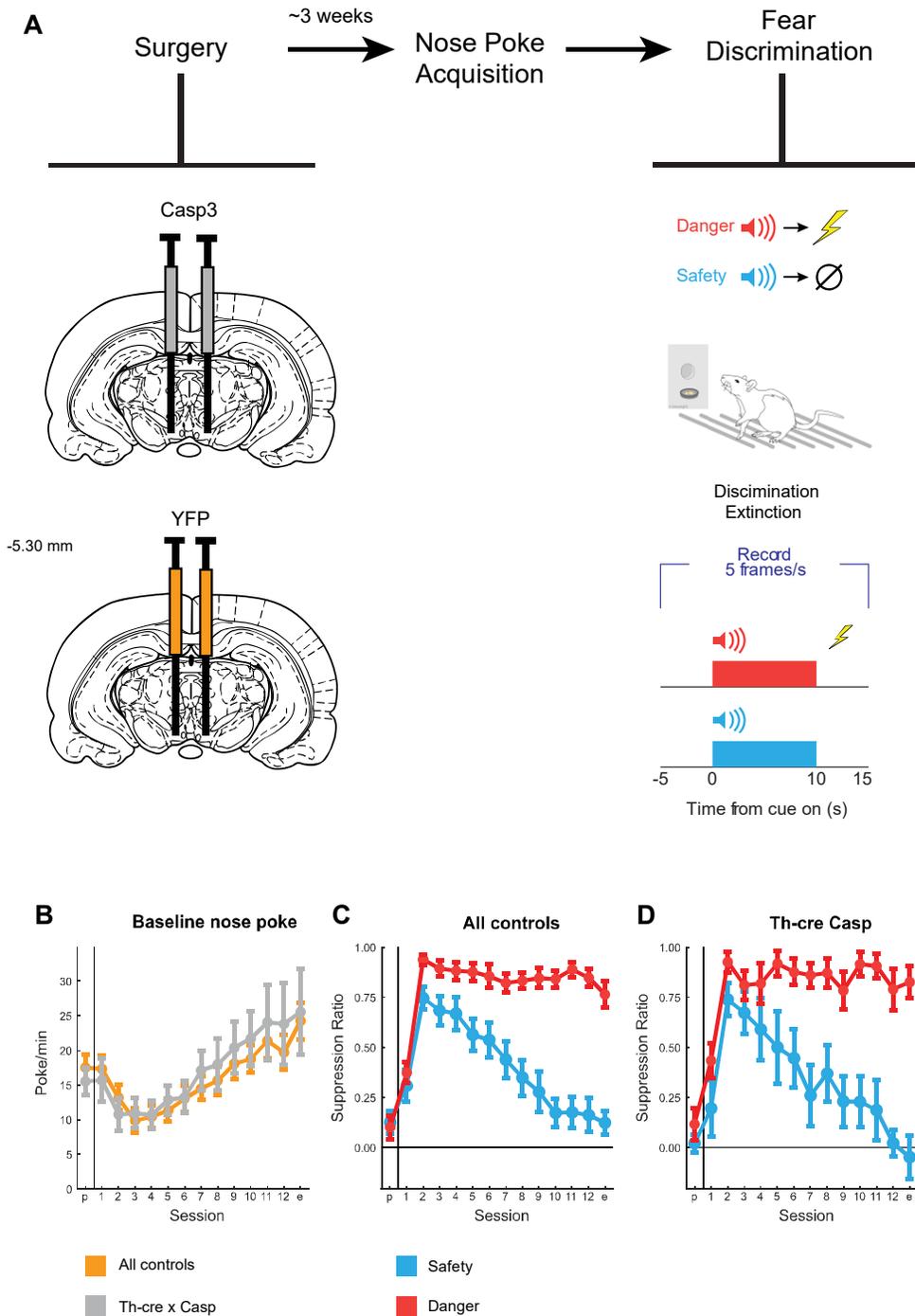
### 4.4.1 *Caspase-mediated deletion of ventral tegmental area dopaminergic neurons leaves discriminative conditioned suppression intact*

Thirty-two Long Evans rats (16 females) underwent stereotaxic surgery. Following recovery, rats were trained to nose poke in a central port for food reward. Nose poking was reinforced on a 60-s variable interval schedule throughout behavioral testing. Independent of the poke-food contingency, auditory cues were played through overhead speakers, and foot shock delivered through the grid floor. The experimental design consisted of two cues deterministically predicting foot shock: danger ( $p=1.00$ ) and safety ( $p=0.00$ ) (Figure 4.1A, right).

It is important to note that all statistical analyses were performed with rat strain (WT vs. Th-cre) and AVV (cre-YFP vs. cre-Casp) as factors; but for visualization purposes, we collapsed three groups: Th-cre x YFP, WT x YFP, and WT x Casp into one group (all controls).

Suppression ratios were calculated as described in Chapter 2. To determine if we observed complete behavioral discrimination, we performed ANOVA for suppression ratios [factors: cue (danger vs. safety), session (13 total: 1 pre-exposure and 12 discrimination), and sex (female vs. male)]. As measured by suppression ratios, there were no differences in discrimination patterns between Th-cre x YFP, WT x YFP, WT x Casp, Th-cre x Casp groups (between-subjects effects for strain and AVV,  $F_s > 0$ ,  $p_s > 0.4$ ). There was no significant cue x session x strain x AVV interaction (within-subjects effects,  $F_{12,192} = 0.95$ ,  $p = 0.49$ ), and no significant cue x strain x AVV interaction (within-subjects effects,  $F_{1,16} = 0.43$ ,  $p = 0.52$ ). Complete behavioral

discrimination emerged over testing for all groups (Figure 4.1C and D). ANOVA found a significant main effect of cue and a significant cue x session interaction ( $F_s > 17$ ,  $p_s < 0.0001$ ; see Table 4.1 for specific values). Additionally, there were no significant cue x strain x AAV ( $F_{1,16} = 0.43$ ,  $p = 0.52$ ) or cue x session x strain x AAV interactions ( $F_{12,192} = 0.95$ ,  $p = 0.49$ ). Following the 12 discrimination sessions, each rat received an extinction test, where each cue was presented 4 times. For each group, t-tests between safety vs. danger reveal high suppression ratios for danger and low suppression ratios for safety (WT x YFP:  $p = 8.05 \times 10^{-5}$ ; WT x Casp:  $p = 0.002$ ; Th-cre x YFP:  $p = 0.0004$ ; Th-cre x Casp:  $p = 0.0006$ ). Together, these results demonstrate that there were no differences in suppression ratios between experimental groups and thus, caspase-mediated deletion of VTA dopaminergic neurons did not affect fear discrimination as measured by suppression of reward seeking.



**Figure 4.1.** Experimental design and nose poke suppression

**(A)** Experimental design. **(B)** Baseline nose poke (poke/min) for all controls,  $n = 24$  (orange) and Th-cre x Casp,  $n = 8$  (gray). **(C)** Suppression ratios show complete discrimination by session 12 for all controls, safety (blue) vs. danger (red). **(D)** Suppression ratios show complete discrimination by session 12 for Th-cre x Casp group, safety (blue) vs. danger (red).

	Suppression ratio	
<b>Cue</b>	<b>F(1,16) = 123.986</b>	<b>p = 2.696e-10</b>
<b>Cue * Sex</b>	<b>F(1,16) = 15.172</b>	<b>p = 0.001</b>
<b>Session</b>	<b>F(12,192) = 31.460</b>	<b>p = 2.002e-48</b>
<b>Cue * Session</b>	<b>F(12,192) = 17.463</b>	<b>p = 4.050e-30</b>
<b>Cue * Session * Strain * Lighting</b>	<b>F(12,192) = 3.260</b>	<b>p = 4.014e-05</b>

**Table 4.1.** Discrimination ANOVA results for suppression ratio. Significant main effects and interactions are bolded.

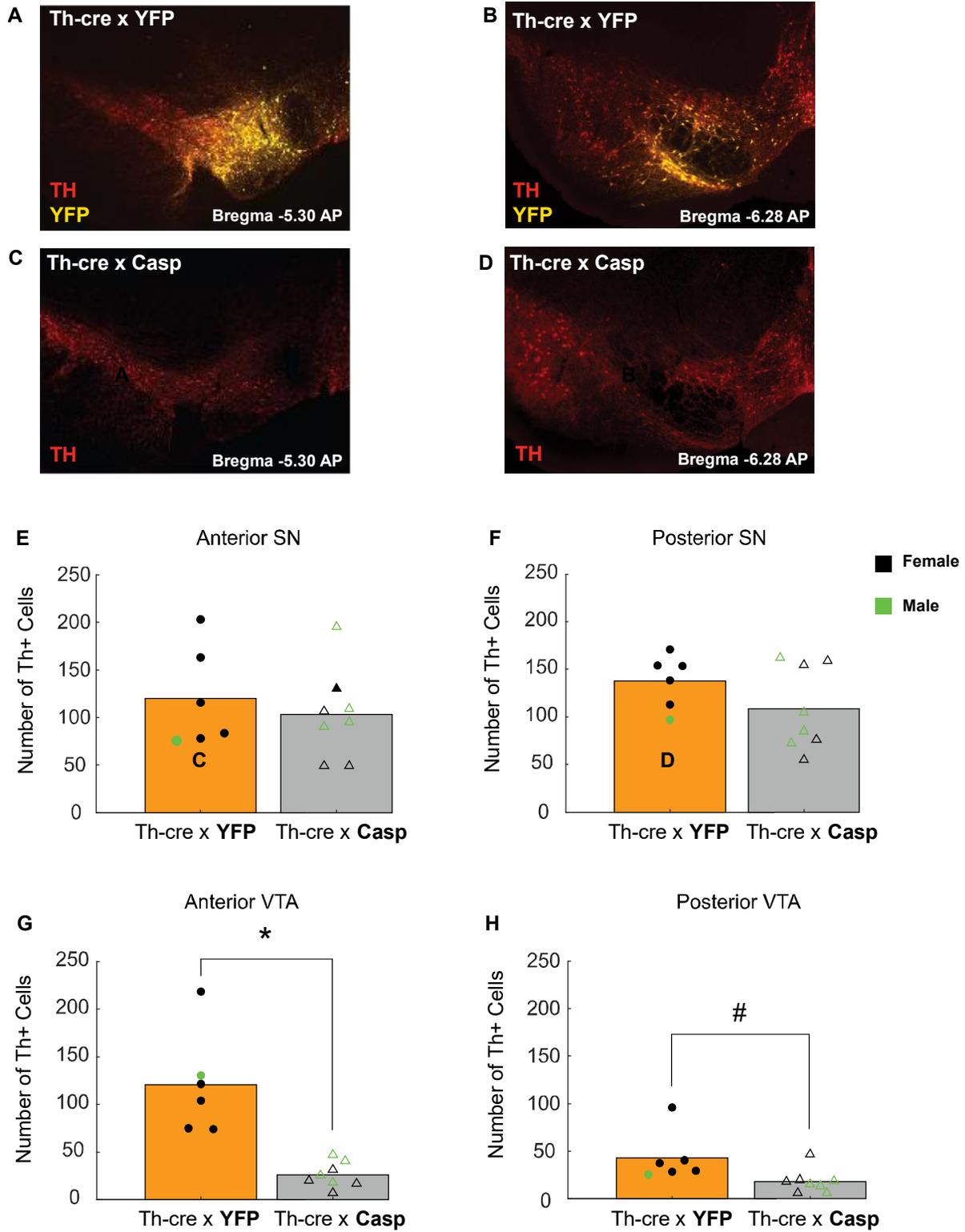
	Suppression ratio	
<b>Cue</b>	<b>F<sub>(1,24)</sub> = 73.500</b>	<b>p = 7.35e-12</b>
Cue * Sex	F <sub>(1,24)</sub> = 0.048	p = 0.828
Cue * Strain	F <sub>(1,24)</sub> = 1.323	p = 0.261
Cue * AAV	F <sub>(1,24)</sub> = 0.050	p = 0.826
Cue * Sex * Strain	F <sub>(1,24)</sub> = 0.226	p = 0.639
Cue * Sex * AAV	F <sub>(1,24)</sub> = 0.372	p = 0.548
Cue * Strain * AAV	F <sub>(1,24)</sub> = 0.152	p = 0.700
Cue * Sex * Strain * AAV	F <sub>(1,24)</sub> = 0.249	p = 0.622
<b>Sex</b>	<b>F<sub>(1,24)</sub> = 6.674</b>	<b>p = 0.016</b>

**Table 4.2.** Extinction ANOVA results for suppression ratio. Significant main effects and interactions are bolded.

We captured 51,200 total frames (800 frames/test x 32 rats x 2 tests) during discrimination session 12 and extinction. Frames were hand scored for nine discrete behaviors: cup, freezing, grooming, jumping, locomotion, port, rearing, scaling, and stretching, plus “background” as in Chapter 2, but with a slightly modified definition of “locomotion” (see Materials and Methods, Behavior categories and definitions).

#### 4.4.2 *Histological verification of viral infusions*

For practical purposes, cell counting and statistical analyses were performed exclusively on rats with confirmed Th-cre<sup>+</sup> status (n = 14). No evidence of YFP<sup>+</sup> cells was observed in wild type rats that received infusions of rAAV-EF1a-DIO-eYFP, and no evidence of Th<sup>+</sup> cell deletion was observed in wild type rats that received infusions of rAAV5-Flex-taCasp3-TEVp. Cell counts were measured for the substantia nigra (SN) and the VTA. Fluorescent immunohistochemical staining for tyrosine hydroxylase (Th) and cell counting of Th<sup>+</sup> cells (red) revealed no significant difference in the average amount of Th<sup>+</sup> cells in the anterior SN between groups (unpaired t-test,  $p = 0.51$ , Figure 4.2E). Similarly, there was no difference in the Th<sup>+</sup> cell counts in the posterior SN between groups (unpaired t-test,  $p = 0.14$ , Figure 4.2F). In the anterior VTA, there was a significantly lower average amount of the Th<sup>+</sup> cells in the Casp group compared to the YFP group (unpaired t-test,  $p = 0.0001$ , Figure 4.2G). Additionally, there was a considerable amount of Th<sup>+</sup> deletion in the posterior VTA of the Casp group compared to the YFP group; however, it is not as extensive as the Th<sup>+</sup> cell deletion observed in the anterior VTA (unpaired t-test,  $p = 0.004$ , Figure 4.2H). These results confirm a successful deletion of Th<sup>+</sup> cells in the anterior VTA, with little deletion in the SN. Further verifying surgical coordinates for the VTA, in Th-cre x YFP animals, YFP<sup>+</sup> cells (yellow) were mostly concentrated in the VTA with little expression in the SN throughout the anterior-posterior axis, as intended (Figure 4.2A, B). Together, these results confirm that viral infusions were successfully targeted in the VTA.

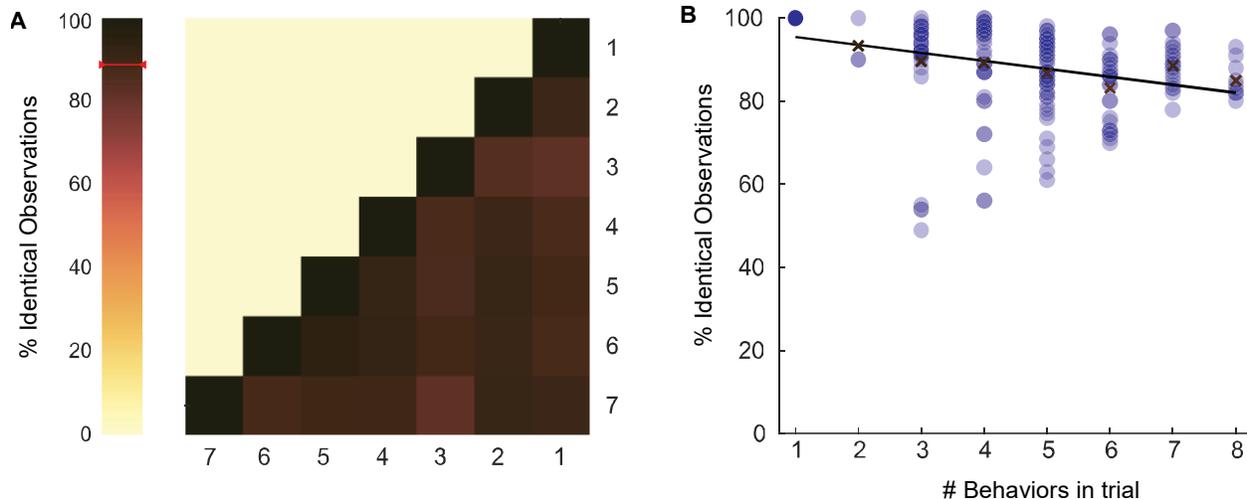


**Figure 4.2.** Th+ cell counts in SN and VTA across anterior and posterior levels

**(A)** Th-cre x YFP rat with anti-TH IHC staining and Texas red + YFP channels, TH+ cells (red) and YFP+ cells (yellow) overlap show YFP expression in the anterior VTA but not SN. **(B)** Th-cre x YFP rat with anti-TH IHC staining and Texas red channel only, TH+ cells (red) are intact in posterior VTA and SN. **(C)** Th-cre x Casp rat with anti-TH IHC staining, show less TH+ cells in the anterior VTA but intact TH+ cells in anterior SN. **(D)** Th-cre x Casp rat with anti-TH IHC staining, show intact TH+ cells in the posterior VTA and posterior SN. **(E)** No significant difference in average number of Th+ cells in the anterior SN between YFP (orange, n=6) and Casp (grey, n=8) groups (unpaired t-test,  $p=0.5094$ ). **(F)** No significant difference in average number of Th+ cells in the posterior SN between YFP (orange) and Casp (grey) groups (unpaired t-test,  $p=0.1417$ ). **(G)** Casp group shows less Th+ cells in anterior VTA compared to YFP controls (unpaired t-test,  $p=0.0001$ ). **(H)** Casp group shows a trend towards less Th+ cells in posterior VTA compared to YFP controls (unpaired t-test,  $p=0.0043$ ,  $p>0.05$  after Bonferroni correction).

#### 4.4.3 *Hand scoring with high inter-rater reliability*

Similar to the experiments in Chapter 2, frames were hand scored for nine discrete behaviors. Frames were systematically hand scored by seven observers blind to rat identity, session number, and trial type (as in Chapter 2). A comparison data set consisting of 10 trials (1,000 frames) was also scored by each observer. A correlation matrix compared % identical observations for the 1,000 comparison frames for each observer-observer pair (Figure 4.3A). Mean % identical observation was 88.67%, with a minimum observer-observer pair agreement of 81.70% and a maximum of 93.90%. Mean percent identical observation was 93.33% when two behavior categories were present, and 89.57% when three behavior categories were present. Even when eight behavior categories were present, a mean percent identical observation of 85% was achieved (Figure 4.3B). Our approach yielded very high inter-observer reliability across trials with few and many behavior categories present.



**Figure 4.3.** Inter-rater reliability

**(A)** Percentage of identical observations between observer-observer pairs. **(B)** Percentage of identical observations as a function of the number of behaviors present in a trial.

#### 4.4.4 Controls did not exhibit expected danger-specific behavioral patterns

##### 4.4.4.1 Discrimination

As discussed previously, ANOVA for suppression ratio demonstrated excellent discrimination across all rats. For the last discrimination session, MANOVA [factors: cue (danger vs. safety), time (20 s total, 1-s bins), sex (female vs. male), strain (WT vs. Th-cre), and AAV (cre- YFP and cre-Casp)] for all nine quantified behaviors returned a significant cue x time interaction ( $F_{171,2907} = 3.16$ ,  $p = 2.06 \times 10^{-35}$ ). This demonstrates that rats behaved differently to the two cues over presentation. However, this does not inform us of the specific behavioral pattern rats exhibited during the safety trials versus danger trials. Further, we also observed a cue x time x lighting interaction ( $F_{171,2907} = 1.41$ ,  $p = 0.001$ ). Note, the lighting differences in the boxes were not visible to the rats,

only to the human raters. The presence of an interaction indicates that the different lighting conditions led to different behavioral judgments. The lighting effects will be addressed later in the discussion.

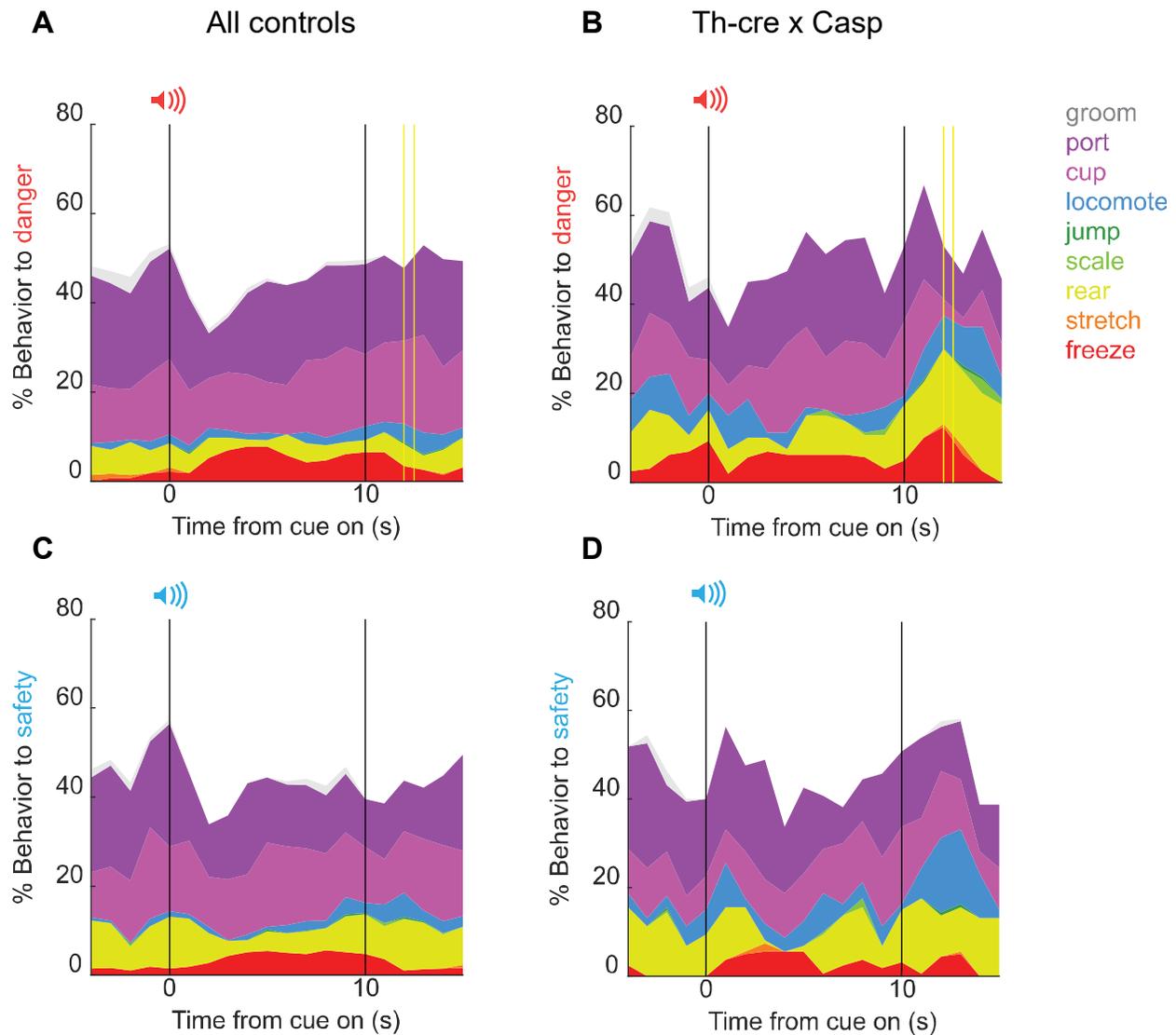
MANOVA further found no cue x time x strain x AAV interaction, suggesting that the multi-behavior patterns observed during discrimination were not altered by VTA dopamine deletion. However, MANOVA did reveal a five-way interaction for cue x time x strain x AAV x lighting ( $F_{171,2907}$ ,  $p = 1.10 \times 10^{-7}$ ) suggesting that the behaviors that were altered by VTA dopamine deletion may have been the same behaviors whose scoring was affected by the lighting condition.

To reveal specific, danger-elicited behaviors during discrimination, we performed univariate ANOVA on each behavior [factors: cue (danger vs. safety), time (20 s total, 1-s bins), sex (female vs. male), strain (WT vs. Th-cre), and AAV (cre- YFP and cre-Casp)]. Surprisingly, some of the behaviors observed to be danger-elicited in Chapter 2 were not observed in this study. Specifically, during discrimination, a significant cue x time interaction was not observed for locomotion. Univariate ANOVA returned a significant cue x time interaction for freezing ( $F_{4612,243} = 3.79$ ,  $p = 3.35 \times 10^{-7}$ ); however, discrimination ethograms (Figure 4.4, Figure 4.4.1) show overall levels of freezing that are lower than the levels observed from Experiments 1 and 2 in Chapter 2 (<10% of overall behavior), and the percent difference in freezing levels between the danger and safety cues was smaller (<4%) across both groups (Figure 4.4.1A and E). Although there were cue x time interactions for jump and rear ( $F_{25,1} = 2.91$ ,  $p = 5.9 \times 10^{-5}$ ;  $F_{6938,365} = 2.72$ ,  $p = 1.74 \times 10^{-4}$ ), looking at the discrimination ethograms for all controls, they were in the opposite direction from what was expected (Figure 4.4A and C, Figure

4.4.1C and D). Strangely, in the all controls group, the safety cue increased jumping and rearing relative to the danger cue during discrimination (Figure 4.4.1C and D).

Univariate ANOVA returned no significant cue x time x strain x AVV interactions for freezing ( $F_{595,31} = 0.5$ ,  $p = 0.97$ ), locomotion ( $F_{450,24} = 0.73$ ,  $p = 0.79$ ), or rearing ( $F_{2163} = 0.85$ ,  $p = 0.65$ ). However, we did observe danger-elicited jumping during discrimination in the Th-cre x Casp group (but not in the all controls group), which I will discuss in the following section (Figure 4.4.1F). These results contrast with the experiments I conducted and discussed in Chapter 2, where we observed robust danger-elicitation of freezing, locomotion, jumping, and rearing during the last session of discrimination.

## Discrimination Session 12

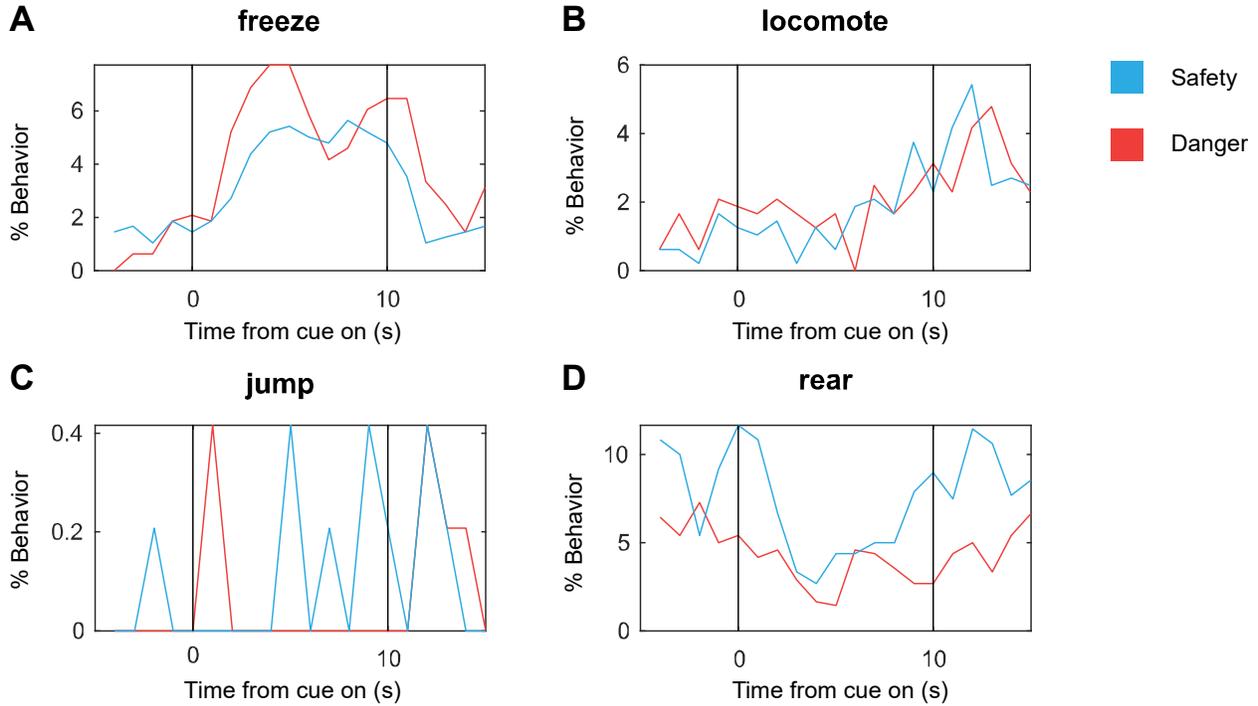


**Figure 4.4.** Temporal ethograms during discrimination

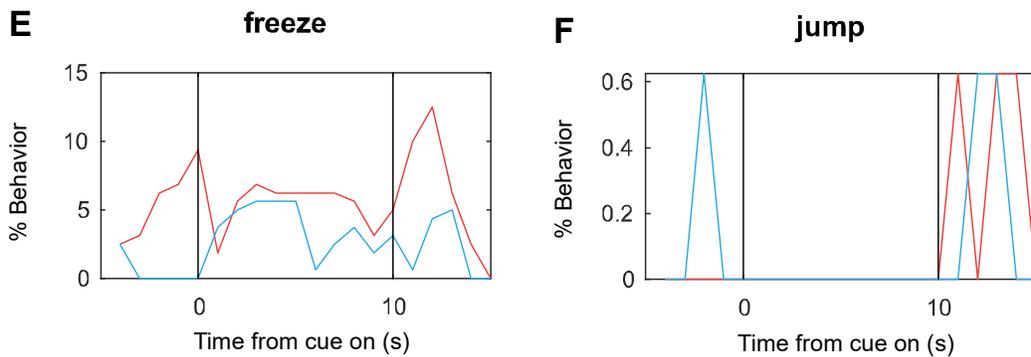
Mean percent behavior from 5s prior through 5s following cue offset is shown for the danger cue during discrimination for **(A)** all controls, and **(B)** Th-cre x Casp; and the safety cue during discrimination for **(C)** all controls, and **(D)** Th-cre x Casp. Behaviors are groom (gray), port (dark purple), cup (light purple), locomote (blue), jump (dark green), scale (light green), rear (yellow), stretch (orange), and freeze (red).

## Discrimination Session 12

### All Controls



### Th-cre x Casp



**Figure 4.4.1.** Behavior comparisons for danger vs. safety during discrimination

% Behavior shown for all controls group for **(A)** freeze, **(B)** locomote, **(C)** jump, and **(D)** rear during discrimination. % Behavior shown for Th-cre x Casp group for **(E)** freeze and **(F)** jump during discrimination.

#### 4.4.4.2 Extinction

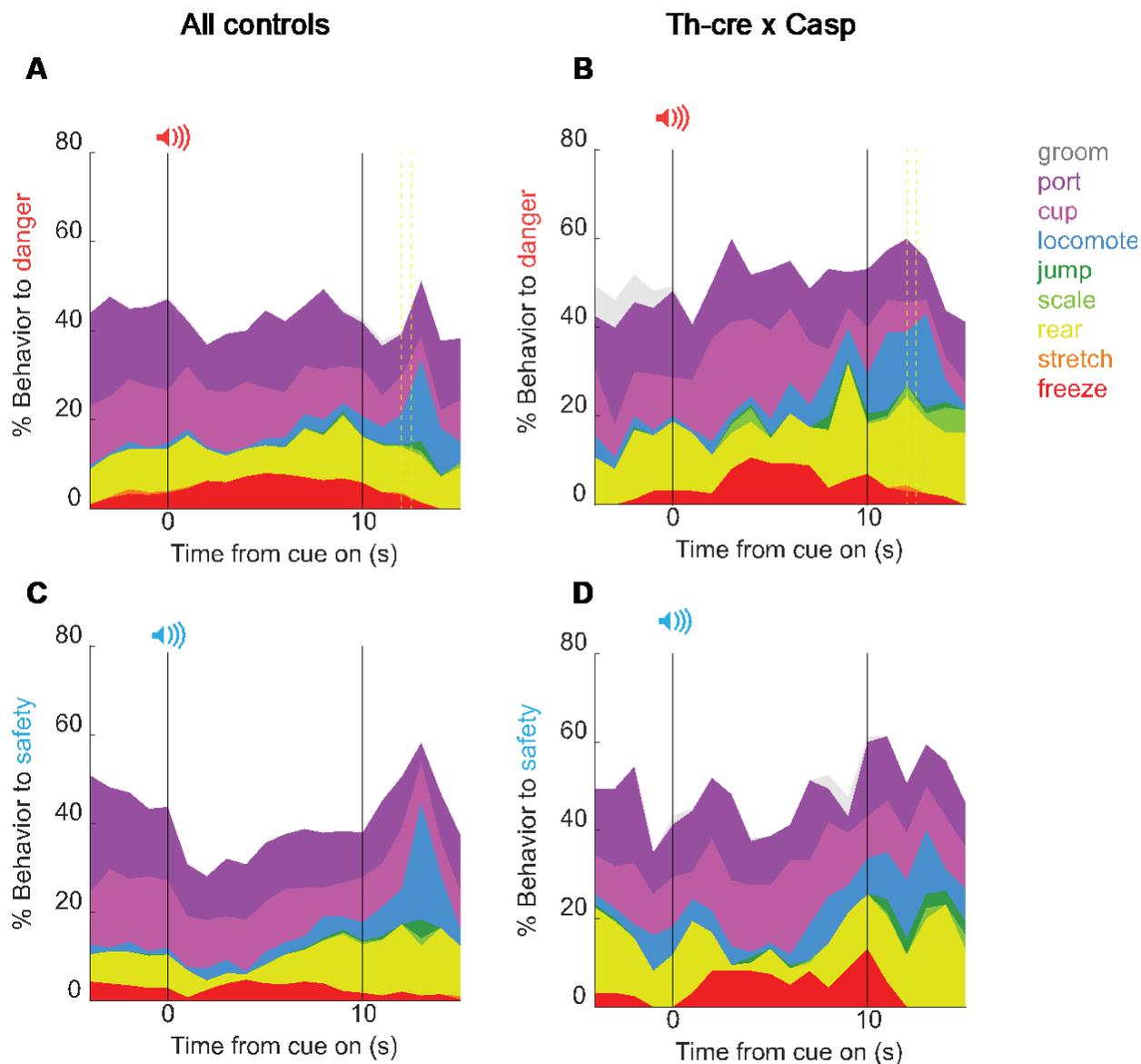
For the extinction session, MANOVA [factors: cue (danger vs. safety), time (20 s total, 1-s bins), sex (female vs. male), strain (WT vs. Th-cre), and AAV (cre- YFP and cre-Casp)] for all nine quantified behaviors returned a significant cue x time interaction ( $F_{171,2907} = 2.06$ ,  $p = 2.02 \times 10^{-13}$ ). This demonstrates that rats behaved differently to the two cues over cue presentation. Yet again, this does not inform us of the specific behavioral pattern the rats exhibited during the safety trials versus danger trials. Similar to the discrimination session, we also observed a cue x time x lighting interaction during extinction ( $F_{171,2907} = 1.52$ ,  $p = 2.5 \times 10^{-5}$ ). The presence of an interaction indicates that the different lighting conditions led to different behavioral judgments in extinction, as in discrimination.

Next, univariate ANOVA [factors: cue (danger vs. safety), time (20 s total, 1-s bins), sex (female vs. male), strain (WT vs. Th-cre), and AAV (cre- YFP and cre-Casp)] were performed for each behavior. Recall the main findings from Experiment 2 in Chapter 2: we observed danger-elicited freezing and locomotion during extinction with reward present. In the present study, significant cue x time interactions were not observed for locomotion during extinction. Similarly, we did not observe a cue x time interaction for jump during extinction. Univariate ANOVA returned a cue x time interaction for freezing ( $F_{3831,202} = 2.85$ ,  $p = 8.2 \times 10^{-5}$ ). However, similar to the freezing patterns during discrimination, extinction ethograms (Figure 4.5, Figure 4.5.1) show overall levels of freezing that are lower than the levels observed from Experiment 2 in Chapter 2 (<10% for all controls group) and the percent difference in freezing levels between the danger and safety cues was smaller (<4% difference) across both groups

(Figure 4.5.1A and E). Although there was cue x time interaction for scaling ( $F_{165,9} = 3.12$ ,  $p = 1.7 \times 10^{-5}$ ), looking at the extinction ethograms for all controls, it was in the opposite direction from what was expected (Figure 4.5A and C, Figure 4.5.1C). Strangely, in the all controls group, the safety cue increased scaling relative to the danger cue during extinction (Figure 4.5.1C). Lastly, univariate ANOVA returned a cue x time interaction for rearing ( $F_{10566,556} = 3.51$ ,  $p = 2 \times 10^{-6}$ ). The extinction ethograms for the all controls group shows that there is some increase in rearing to the danger cue relative to the safety cue, but most of the difference is accounted for by the early-cue period (first 5 s) (Figure 4.5.1D).

Given the general lack of expected behavioral patterns during discrimination and extinction in the all controls group, the interpretation of further results from the deletion group, the Th-cre x Casp group, needs to be made with many considerations, which I will elaborate on in the discussion section of this chapter.

## Extinction

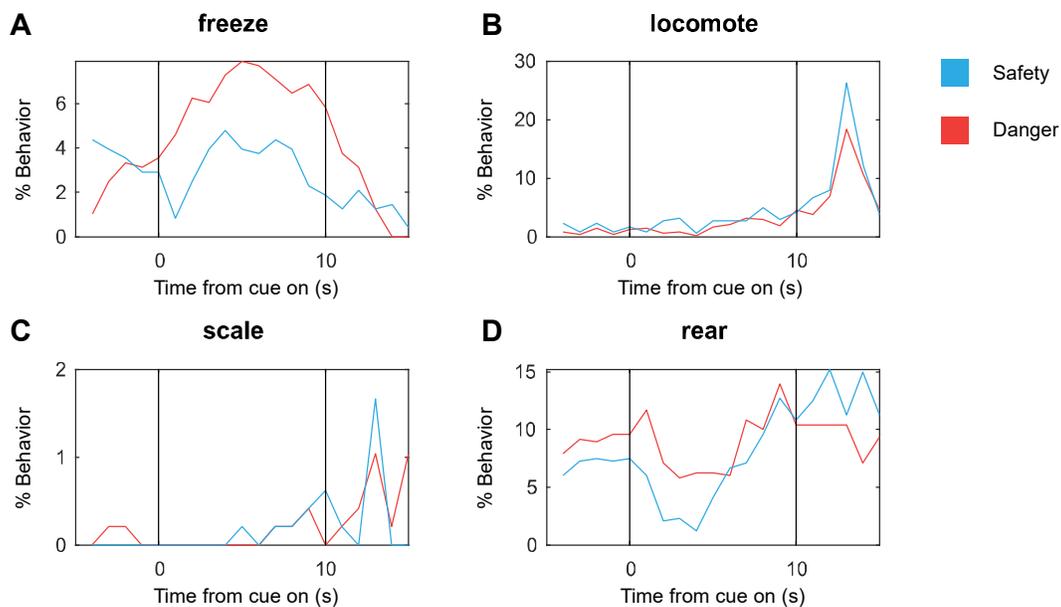


**Figure 4.5.** Temporal ethograms during extinction

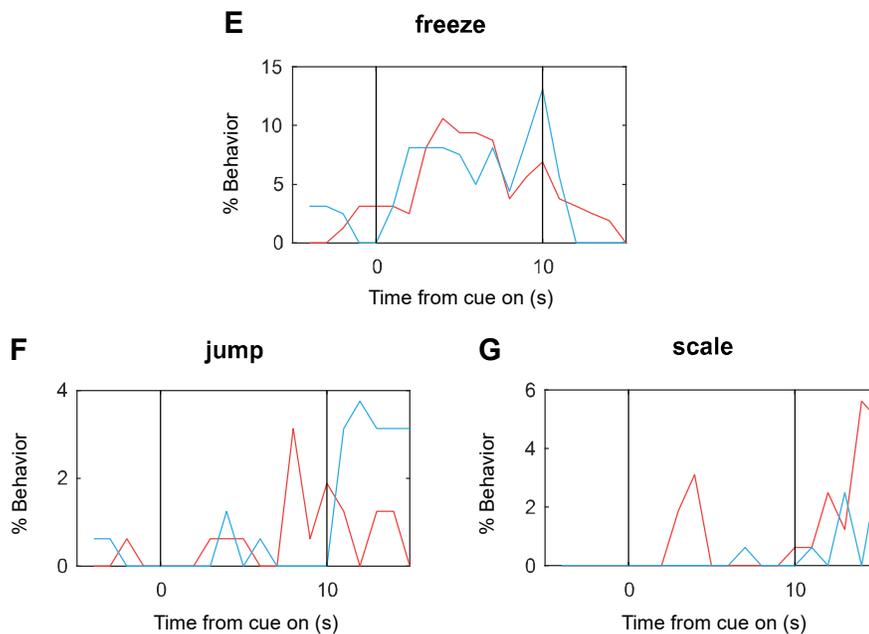
Mean percent behavior from 5s prior through 5s following cue offset is shown for the danger cue during extinction for **(A)** all controls, and **(B)** Th-cre x Casp; and the safety cue during extinction for **(C)** all controls, and **(D)** Th-cre x Casp. Behaviors are groom (gray), port (dark purple), cup (light purple), locomote (blue), jump (dark green), scale (light green), rear (yellow), stretch (orange), and freeze (red).

## Extinction

### All Controls



### Th-cre x Casp



**Figure 4.5.1.** Behavior comparisons for danger vs. safety during extinction

% Behavior shown for all controls group for **(A)** freeze, **(B)** locomote, **(C)** scale, and **(D)** rear during extinction. % Behavior shown for Th-cre x Casp group for **(E)** freeze, **(F)** jump, and **(G)** scale during extinction.

#### *4.4.5 Deletion of VTA dopaminergic neurons potentially enhances jumping during extinction*

The central question driving this study was does VTA dopamine contribute to the expression of defensive behaviors during fear discrimination and/or extinction? To answer this, we constructed complete temporal ethograms for nine discrete behaviors in intact controls and VTA dopamine depleted rats during the last session of discrimination (session 12) when discrimination was complete, as well as during extinction when foot shocks were absent (Figures 4.4 and 4.5).

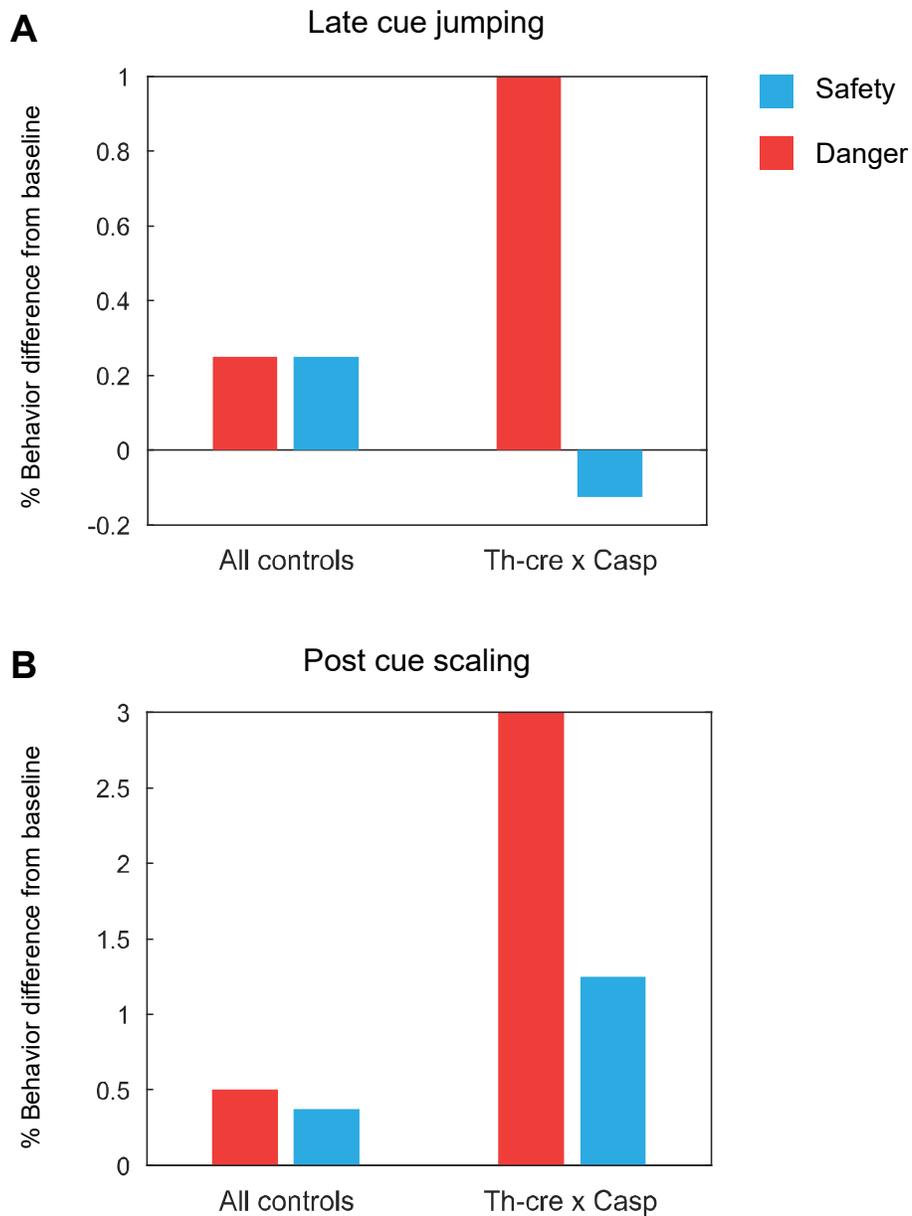
##### *4.4.5.1 Discrimination*

To reveal specific, danger-elicited behaviors during discrimination, we performed univariate ANOVA on each behavior [factors: cue (danger vs. safety), time (20 s total, 1-s bins), sex (female vs. male), strain (WT vs. Th-cre), and AAV (cre- YFP and cre-Casp)]. Univariate ANOVA results were subjected to Bonferroni correction ( $p < 0.0055$ ,  $0.05/9 = 0.0055$ ) to account for the nine separate analyses. To determine any effects of VTA dopamine deletion on each behavior, we address specifically cue x time x strain x AAV interactions. Univariate ANOVA found a significant cue x time x strain x AAV interaction for jumping ( $F_{21,1} = 2.49$ ,  $p = 0.001$ ), where jumping peaked during foot shock delivery (2 s after cue termination) for the Th-cre x Casp group, but not for any other groups (Figure 4.4B, Figure 4.4.1F).

##### *4.4.5.2 Extinction*

To reveal specific, danger-elicited behaviors during extinction, we performed univariate ANOVA on each behavior [factors: cue (danger vs. safety), time (20 s total, 1-s bins), sex (female vs. male), strain (WT vs. Th-cre), and AAV (cre- YFP and cre-

Casp)]. Univariate ANOVA results were subjected to Bonferroni correction ( $p < 0.0055$ ,  $0.05/9 = 0.0055$ ) to account for the nine separate analyses. To determine any effects of VTA dopamine deletion on each behavior, we assessed specifically cue x time x strain x AAV interactions. Univariate analyses found a cue x time x strain x AAV interaction for jumping and scaling ( $F_{212,11} = 2.66$ ,  $p = 2.5 \times 10^{-4}$ ;  $F_{197,10} = 3.73$ ,  $p = 4.59 \times 10^{-7}$ ). Extinction ethograms show that jumping increased during the late cue period (last 5 s of cue) in the Th-cre x Casp group (Figure 4.5B, Figure 4.5.1F). Also of note is the increase in scaling during the post-cue period (when foot shock would have occurred) for the Th-cre x Casp group (Figure 4.5.1G). These jumping and scaling patterns did not occur in the other groups. Importantly, for scaling, univariate ANOVA found a cue x time x lighting interaction ( $F_{234,12} = 4.45$ ,  $p = 6.01 \times 10^{-9}$ ) and a cue x time x strain x AAV x lighting interaction ( $F_{263,14} = 4.99$ ,  $p = 2.32 \times 10^{-10}$ ), indicating that the scoring of scaling was impacted by the lighting condition. Together, these results suggest that caspase-mediated deletion of VTA dopaminergic neurons potentially promote danger-cue specific jumping and scaling during extinction (Figure 4.6A and B).



**Figure 4.6.** Caspase-mediated deletion of VTA DA increases jumping and scaling during extinction

**(A)** % jumping during the late cue period for safety (blue) and danger (red) between all controls and Th-cre x Casp. **(B)** % scaling during the post cue period for safety (blue) and danger (red) between all controls and Th-cre x Casp. There was a 4-way ANOVA interaction (time x cue x strain x AAV)  $p < 0.001$ , for late cue jumping and post cue scaling during extinction.

## 4.5 Discussion

### 4.5.1 Summary

In summary, the findings from this study suggest a potential role for VTA dopamine in the suppression of jumping and scaling to danger cues during extinction. Specifically, caspase-mediated deletion of VTA dopaminergic neurons seems to promote appropriately-timed jumping during the late-cue period (before foot shock delivery) during extinction, as well as scaling during the post-cue period, when foot shock would have occurred during extinction. During extinction, jumping and scaling were not observed during the danger cue in either the controls from the present study, nor the intact rats from Experiment 2 of Chapter 2. The present findings need to be interpreted in the context of the behavioral patterns we observed in the all controls group which includes Th-cre rats receiving cre-YFP (Th-cre x YFP), wild type rats receiving cre-YFP (WT X YFP), and wild type rats receiving cre-Casp (WT x Casp).

### 4.5.2 Limitations and caveats

As mentioned previously, the original design of this study included only transgenic rats with Th-cre<sup>+</sup> status. This original design ensured that there would be two groups: the control group that received cre-YFP infusions into the VTA (n=16, 8 females) and the deletion group that received cre-Casp infusions into the VTA (n=16, 8 females). However, due to the unforeseen circumstances regarding the transgenic status of the male breeders we used to breed rats for the intended experiment, we had to genotype all animals after the conclusion of the experiment. Because of this unanticipated circumstance, a factor lacking from the original experimental design was added: rat strain. We then were left with four groups: Th-cre x YFP (n=6, 5 females),

WT x YFP (n= 10, 3 females), Th-cre x Casp (n=8, 4 females), and WT x Casp (n= 8, 4 females). This additional factor added another variable into statistical analyses.

Moreover, the sample size of the deletion group was reduced by half (n = 16 to n = 8).

The added factor of rat strain and reduced sample size decreased statistical power and added another layer to the interpretation of the results.

Another unanticipated factor we had to consider in the analyses was the lighting of the boxes. Half of the behavior boxes were dimmer than the other half. This affected all human raters' judgments during behavior scoring as there were significant interactions with lighting in the MANOVA tests. However, it seemed to have affected the raters equally, as the inter-rater reliability was exceptionally high for this study (IRR = 88.67%), surpassing the inter-rater reliability of Experiments 1 and 2 from Chapter 2 (IRR = 79.25% - 82.83%). Since MANOVA did not return a significant cue x time x strain x AAV interaction for either discrimination nor extinction, but returned a cue x time x strain x AAV x lighting interaction and cue x time x lighting interaction during both discrimination and extinction, this may indicate that lighting influenced the observations of the specific behaviors that were altered by VTA dopamine deletion.

A major difference between the present study and the previous studies from Chapter 2 was that background accounted for the majority of the 51,200 hand-scored frames (~60%). One possible reason for this increase in frames labeled as background is that perhaps the human raters were more "restrictive" with scoring the frames as a discrete behavior. But since the IRR for the present study was extremely high (>88%), this implies that all human raters scored background frames similarly, and were equally as "restrictive." Since the majority of the frames were labeled as background and

excluded from analyses, it reduces the number of frames labeled as one of the nine discrete behaviors. This fact combined with the reduced sample size of the deletion group, negatively impacts the statistical power of the findings from the deletion group.

Another consideration of the present study is that caspase-mediated deletion of dopaminergic neurons was primarily concentrated in the anterior portions of the VTA. It is possible that anterior and posterior populations of dopaminergic neurons in the VTA have different functions. The VTA extends quite extensively throughout the anterior-posterior axis of the brain (-5.20 to -6.30 AP from Bregma). Many of the fear learning studies manipulate dopaminergic cells in the entirety of the VTA, and do not distinguish between possible sub-populations of these cells in the VTA. Future studies that target the anterior and posterior VTA separately, and also the entirety of the VTA would address this possibility.

Altogether, the caveats of this study likely affected the behavior data and is the most probable explanation for why we did not observe the expected behavioral patterns in all the control animals. A replication of the original study design is necessary to more accurately and completely answer the research question: does VTA dopamine contribute to the expression of defensive behaviors during fear discrimination and/or extinction?

#### *4.5.3 What our findings mean in the context of the existing literature*

If we assume that the findings from this study are replicable, what does it mean for the brain circuitry responsible for producing activity-promoting defensive behaviors? Since there have been few studies on the role of dopaminergic cells in the VTA in the context of fear, I aimed to delete dopaminergic neurons in the VTA to see if there were

any effects on the defensive behaviors we observe in our laboratory. We observed that caspase-mediated deletion of dopaminergic neurons in the VTA potentially promotes danger-evoked, timed jumping during extinction.

#### *4.5.3.1 Revisiting the role of the dPAG and amygdala in the expression of activity-promoting defensive behaviors*

Recently, studies have demonstrated that flight and darting are conditioned, defensive behaviors that readily occur in fear settings (Fadok et al., 2009; Borkar et al., 2024; Gruene et al., 2015; Mitchell et al., 2022). Moreover, some studies have investigated the neural underpinnings of such flight behaviors. There is evidence that supports a role for the dPAG to increase flight and overall locomotor activity in fear settings (Morgan et al., 1998; Deng et al., 2016). In addition, the dorsolateral PAG (dIPAG) has been shown to mediate circa-strike behaviors (i.e. locomotion right after foot shock) (Fanselow, 1995). Fanselow and colleagues proposed that the dIPAG inhibits the vIPAG in fear settings and demonstrate that electrolytic lesions to the dIPAG enhances freezing since it “releases” the vPAG from inhibition. This competitive inhibitory circuit determines whether or not defensive freezing is expressed. In other words, neural circuits for freezing and flight compete with each other during threat encounter.

There is evidence that implicates the subregions of the amygdala in the expression of conditioned flight and active avoidance (Fadok et al., 2017; Borkar et al., 2024; Munoz et al., 2010). Fadok et al. (2017) demonstrates that a competitive inhibitory circuit within the CeA between corticotropin-releasing factor (CRF)+ neurons and somatostatin (SOM)+ neurons produces conditioned flight and freezing,

respectively. Borkar et al. (2024) reveals that the DP→CeM pathway is necessary and sufficient for producing flight in mice during a high-threat conditioning paradigm. Further, the authors in this study demonstrate that the neurons in the CeM receiving inputs from the DP also project to the dIPAG and lateral PAG (lPAG). This CeM→dIPAG/lPAG pathway likely conveys fear associations in order to produce flight during fear conditioning. Together, these studies form the basis of a potential brain circuit for conditioned, activity-promoting defensive behaviors.

#### *4.5.3.2 Linking VTA dopamine to the amygdala and PAG in the expression of activity-promoting defensive behaviors*

Anatomical studies show that there are both dopaminergic and non-dopaminergic projections from the VTA and SN to the amygdala (Fallon et al., 1978; Loughlin and Fallon, 1983). Furthermore, studies have revealed a role of VTA dopamine in altering fear learning via its projections to the CeA and BLA (Nader and LeDoux, 1999; Fadok et al., 2009; Jo et al., 2018; Tang et al., 2020). Specifically, it seems that dopaminergic projections from the VTA to the amygdala play an important role in promoting and facilitating fear learning. Synthesizing these findings from various studies, one possible hypothesis is that VTA dopamine modulates fear learning and the expression of flight and other activity-promoting fear behaviors primarily via projections to the amygdala. Subsequently, the amygdala conveys this information to the dPAG to produce an appropriate active response during fear conditioning.

In the present study, we found that caspase-mediated deletion of dopaminergic neurons in the VTA did not affect fear learning, as VTA dopamine depleted rats successfully discriminated between safety and danger cues as measured by nose poke

suppression. These findings are quite contradictory to the existing literature. Since studies have shown that dopaminergic neurons in the VTA are critical for fear learning, it would be a reasonable hypothesis for our study that VTA dopamine-depleted rats would show impaired discrimination. As mentioned previously, caspase-mediated deletion of dopaminergic neurons was primarily in the anterior VTA, and did not kill dopaminergic neurons all throughout the anterior-posterior axis of the VTA. So, it is possible that in order to observe an impairment in fear discrimination, a complete deletion of dopaminergic neurons in the VTA is necessary. It is important to highlight that we measured fear discrimination via conditioned suppression. In doing so, we also measured baseline nose poke rates and observed no differences in baseline nose poke rates between controls and VTA dopamine-depleted rats (Figure 4.1B). Dopaminergic neurons in the VTA play an important role in reward seeking (Schultz, 1998; Stuber et al., 2008; Roesch et al., 2007; Stalnaker et al., 2019). So a reasonable hypothesis for the present study is that dopamine-depleted rats would have lower baseline nose poke rates. Studies lesioning the VTA with a dopamine-selective neurotoxin (6-OHDA) in animals show reduced food seeking in hungry rats (Papp and Bal, 1986), but these lesions were more extensive than the ones performed in this study. Therefore, differences in baseline nose poke rates may have been observed if the dopaminergic cell deletion were more extensive throughout the VTA.

Before discussing jumping and scaling patterns during extinction, it is important to highlight that unlike jumping, scaling was affected by the lighting conditions of the boxes as reported by the univariate ANOVA interactions, and therefore, is likely not a very reliable result. So the following discussion will focus on the late-cue period jumping

patterns we found. We observed that the VTA dopamine deletion group exhibited increased, timed jumping to the danger cue during extinction. Based on the literature that implicates a role of dopaminergic neurons in the VTA to facilitate the expression of various conditioned and unconditioned activity-promoting behaviors (such as fear potentiated startle or flight), a reasonable hypothesis would be that abolishing the VTA's dopaminergic neuronal population would actually decrease the levels of similar activity-promoting behaviors, such as locomotion and jumping in our behavioral setting. Instead, we observed an increase in jumping during the late cue period in the VTA dopamine depleted rats during extinction. Further, we did not observe increased jumping during extinction in the control group of the present study nor in the wild type, intact animals in Experiment 2 of Chapter 2. We did, however, observe increased, timed jumping during the danger cue during discrimination, where foot shock was present in Experiment 1 of Chapter 2. Together, the findings from these experiments implicate that potentially, a lack of VTA dopaminergic cells facilitates the persistence of avoidant-like behaviors (like jumping) even in the absence of the aversive stimulus, foot shock. It is possible that dopaminergic projections from the VTA to the amygdala modulate the expression of jumping. Specifically, VTA dopamine may suppress the amygdala's output normally, so that jumping ceases once a fear conditioned cue is no longer paired with an aversive foot shock. And when this VTA DA→amygdala projection is inhibited, the amygdala produces jumping to a fear conditioned cue in a setting where the foot shock is no longer present.

In summary, the present study indicates a potential role for VTA dopamine in suppressing jumping to a danger cue in the absence of foot shock. A replication of the

original experiment design is needed to confirm the present findings. Further, future studies that manipulate specific dopaminergic projections from the VTA to other regions, such as the BLA or CeA, in our laboratory's discrimination procedure would help our understanding of its potential contributions to the expression of defensive behaviors. These future studies not only can potentially uncover a role for VTA dopamine in defensive behaviors, but can also paint a more complete picture of the neural circuitry underlying the expression of activity-promoting defensive behaviors.

## **Chapter 5: Summary of results and discussion**

## 5.1 Summary of results

In Chapter 2, I conducted two experiments to investigate the complete breadth of behaviors exhibited during fear discrimination with auditory cues that predicted unique foot shock probabilities. The findings from Experiment 1 demonstrate that a suite of behaviors are expressed to a fear conditioned cue during discrimination including freezing, locomotion, jumping, and rearing. The findings from Experiment 2 demonstrate that danger-evoked freezing patterns persist during extinction only when the reward apparatus is present, while danger-evoked locomotion patterns persist during extinction when the reward apparatus is present or absent. Together, the findings from Chapter 2 support reports from other laboratories that observe conditioned flight behavior and darting in rodents during fear conditioning (Gruene et al., 2015; Mitchell et al., 2022; Borkar et al., 2024).

In Chapter 3, I devised a machine learning pipeline for automatic behavior scoring using CNNs. For this project, I used the hand-scored behavior frames from Experiment 1 of Chapter 2 to train a neural network model to predict behavior labels. With the optimization and implementation of this neural network pipeline, future behavior-focused projects in the laboratory will be able to produce and process behavioral data quickly and efficiently.

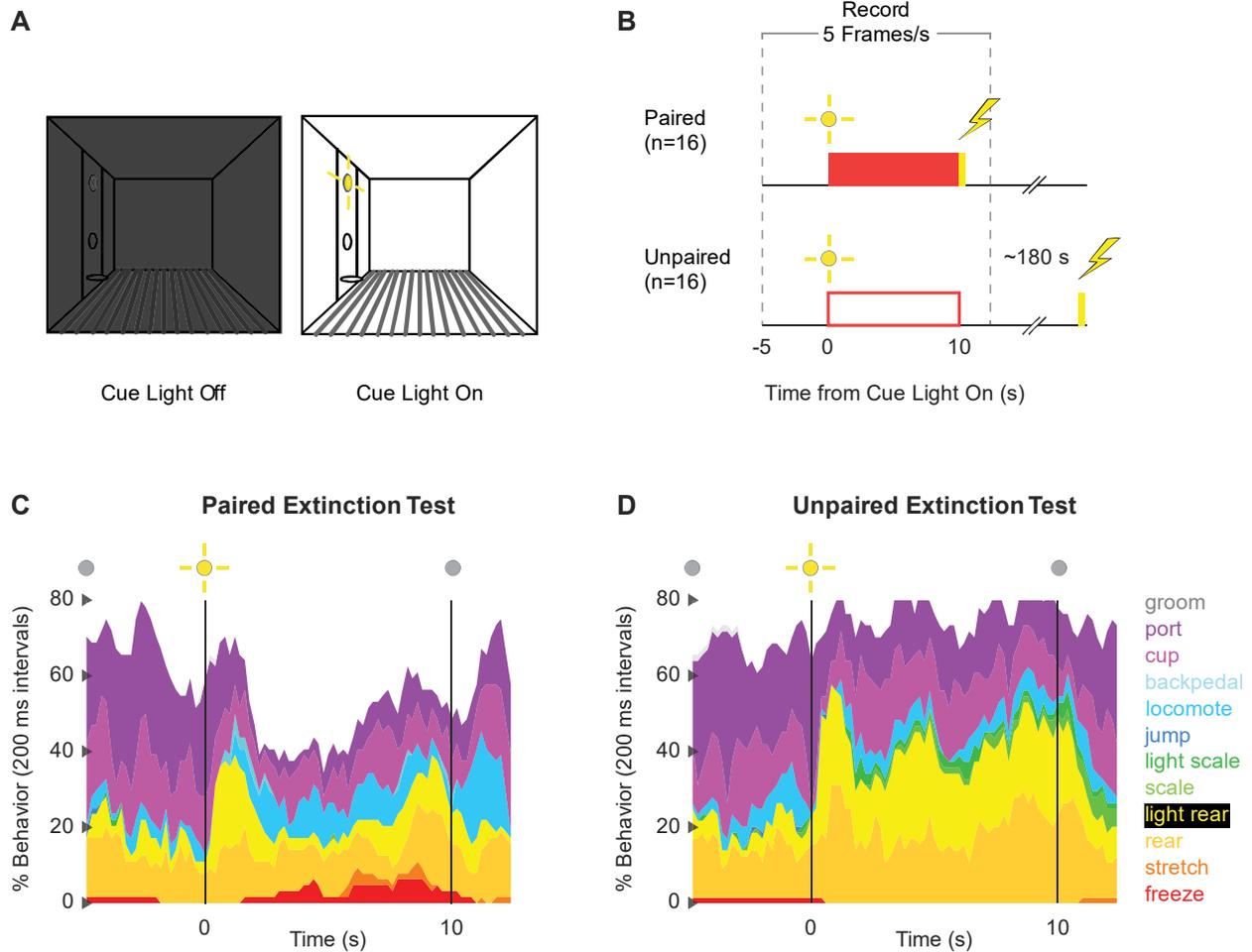
In Chapter 4, I conducted an experiment to uncover a potential role for dopaminergic neurons in the VTA in the expression of the conditioned, defensive behaviors we observed from Chapter 2. Using a combined transgenic and cre-dependent viral approach, I was able to successfully delete dopaminergic neurons in the anterior VTA and compare behavioral patterns between the deletion rats and intact

rats (control). The findings from this study indicate that VTA dopamine potentially suppresses the expression of an avoidant-like behavior, jumping to danger during extinction, when foot shocks are absent.

## **5.2 Discussion**

### *5.2.1 Re: Limitations of Chapter 4*

In Chapter 4, I discussed in detail the several considerations of the experimental design and how it affected the behavioral data. Upon seeing the unexpected behavioral patterns from the controls in this study, one may call into question the reliability of the behaviors observed in Chapter 2 (primarily danger-elicited freezing and locomotion). However, researchers in our laboratory conducted an experiment after the conclusion of the studies presented here, using the same behavior software program and camera system that were used in Chapters 2 and 4, with light cues instead of auditory cues (Figure 5.1A and B). Experimenters quantified the nine discrete behaviors that were measured in Chapters 2 and 4, in addition to behaviors unique to the light cue (i.e. back pedal and light rear). In this experiment, there were expected patterns of danger-elicited freezing and locomotion as well as danger-suppressed poke and cup in the paired groups during extinction (Figure 5.1C). The behavioral patterns of the paired groups in this study confirm that danger-evoked freezing and locomotion are routinely expressed in our laboratory's behavioral setting.



**Figure 5.1.** Danger-elicited freezing and locomotion are routinely expressed in our laboratory

**(A)** Behavior box with cue light off (left) and cue light on (right). Cue light illuminated the boxes. **(B)** Experimental design: paired group received foot shock immediately after light cue termination while unpaired group did not receive foot shock immediately after light cue presentation (after ~180 s). **(C)** Temporal ethogram for paired group during extinction test shows increased freezing (red) and locomotion (sky blue) during light cue period. **(D)** Temporal ethogram for unpaired group during extinction test do not show increased levels of freezing (red) or locomotion (sky blue) during light cue period. However, there is an increase in rearing (dark yellow) and light rearing (light yellow).

### *5.2.2 Temporal organization of defensive behaviors: revisiting the brain circuit for activity-suppressing and activity-promoting defensive behaviors*

The findings from Chapter 2 support the hypothesis that a mixture of freezing and activity-promoting defensive behaviors are expressed in a temporally organized manner during a fear conditioned cue. We also report that the behaviors measured in this study were expressed independently from one another as we observed no strong negative or positive relationships between any pair of behaviors. However, activity-suppressing behavior (freezing) and activity-promoting behavior (flight and locomotion) have been shown to “compete” with each other in fear settings (Fanselow, 1995; Fadok et al., 2017). What is going on here? Even though we do not observe strong negative correlations between freezing and activity-promoting behaviors (locomotion, jumping, and rearing), we do observe temporal organization of these behaviors during the 10 s danger cue presentation. Temporal ethograms from experiments in Chapter 2 showed that generally, early-cue freezing gave way to late-cue mixture of locomotion, jumping, and rearing. This suggests that in our behavioral setting, the immobile defensive response is not directly “competing” with activity-promoting defensive behaviors, but rather the contrasting behaviors are expressed in temporally-specific manner throughout the danger cue. This implies that temporal information plays a crucial role in defensive responding, and must originate somewhere in the brain.

Knowing that the dPAG plays a critical role in expressing flight, avoidance, and locomotion, a question that still remains unanswered is how? Where is the dPAG receiving inputs from to produce these activity-promoting defensive behaviors? dPAG neurons increase their firing rate as a function of threat proximity in a naturalistic threat

setting (Deng et al., 2016), which means that the dPAG may not be completely dependent on amygdalar or other regions' inputs for threat processing. It is possible that since dPAG neurons increase their firing as threat approaches (i.e. as a function of distance from threat) in a naturalistic threat setting, this could be analogous to temporal information that is conveyed in a fear conditioning setting. Since foot shock is consistently delivered after the termination of a 10 s danger cue, a population of dPAG neurons may integrate this timing information over the course of learning, and once trained, appropriately increase or decrease their firing depending on the length of cue presentation (increase firing rate as the cue plays and foot shock is imminent). In support of the dPAG's involvement in assessing information about threat, there is evidence that the dPAG projects to the BLA, conveying information about unconditioned stimuli in both fear conditioning and naturalistic threat settings (Kim et al., 2013). Moreover, electrical stimulation of the dPAG, independent from amygdalar inputs, produces unconditioned freezing and escape in rats (Oliveira et al., 2004). Together, these studies implicate a role for dPAG neurons in many functions including threat and stimuli processing and defensive behavior expression.

If the dPAG and dIPAG are involved in the production of conditioned defensive behaviors, I do not suspect that the dPAG alone is sufficient to produce such behaviors like flight and locomotion during fear conditioning. It is likely that although some neuronal populations in the dPAG may respond to the salience of threat, additional information about the environment and relationship between a conditioned cue and aversive stimulus must be relayed to the dPAG for the dPAG to appropriately integrate information and produce defensive activity-promoting responses. Since the amygdala is

critical for integrating information during fear conditioning and is critical for producing conditioned freezing, it is likely that the amygdala conveys this information to the dPAG to produce conditioned, activity-promoting responses as well. A retrograde tracing study (in mice) done in the dorsal and ventral subdivisions of the PAG demonstrates that CeM neurons and CeL neurons project to the dIPAG (Li and Sheets, 2018). Additionally, the DP→CeM pathway has been shown to be necessary and sufficient to produce conditioned flight in mice, and optogenetic stimulation of CeM-projecting neurons in the DP produce excitatory neuronal firing in the lateral (IPAG) and dorsolateral portions (dIPAG) of the PAG innervated by CeM projections (Borkar et al., 2024). Therefore, a CeM→dIPAG pathway is a likely candidate for producing conditioned, activity-promoting defensive behaviors. Studies directly manipulating the CeM→dIPAG pathway and observing its effects on behavior would address whether this specific pathway is necessary and/or sufficient for the expression of flight.

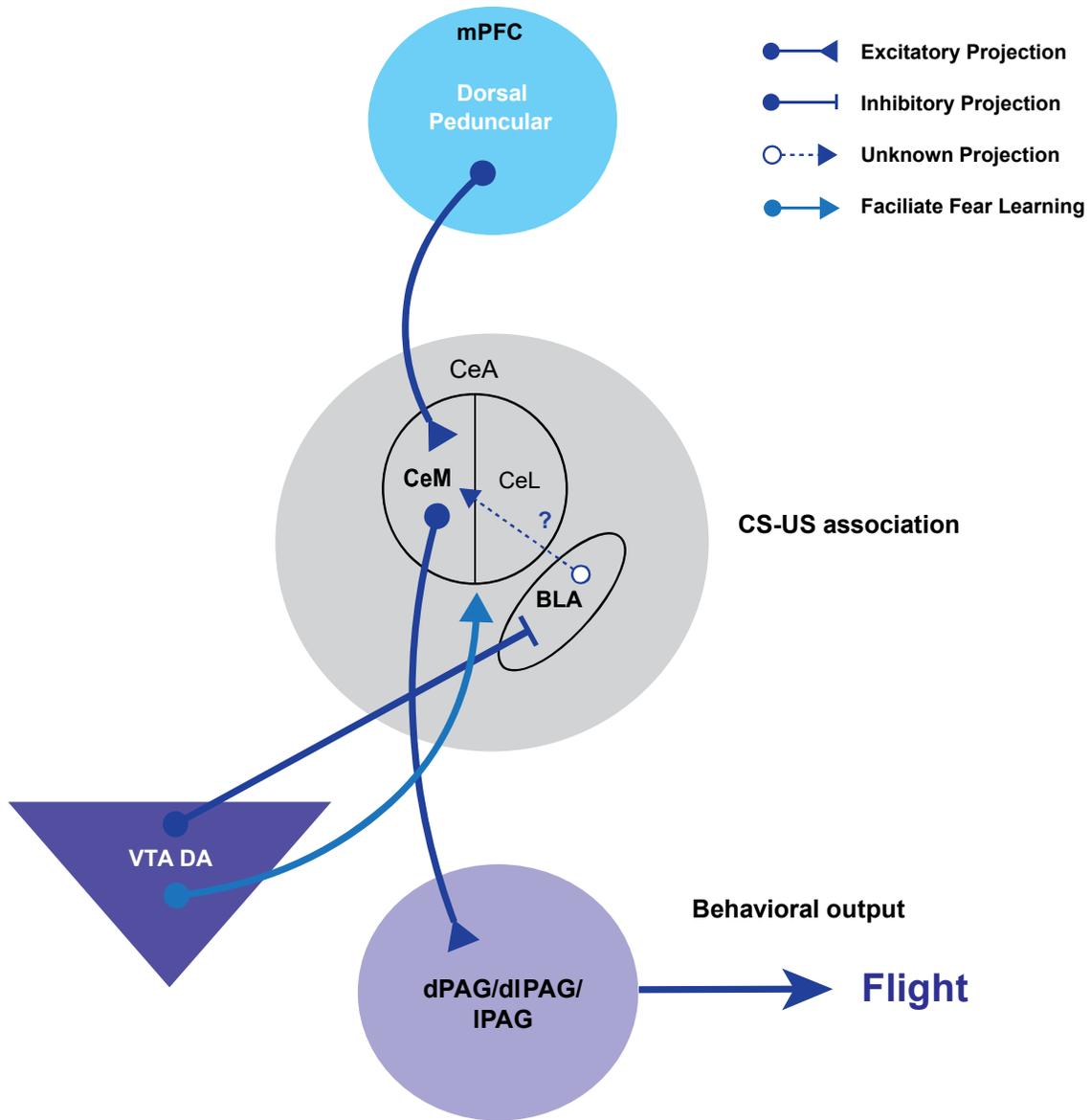
### *5.2.3 Where does the VTA fit in the fear circuit?*

The VTA projects to the CeA, and optogenetic stimulation of VTA dopaminergic terminals in the CeA enhances fear discrimination under a strong US intensity (Jo et al., 2018). Additionally, optogenetic inhibition of VTA dopaminergic projections to the CeA reduces fear discrimination by increasing freezing levels to a safety cue (Jo et al., 2018). Thus, the VTA may modulate fear learning and subsequent behavioral responding via its projections to the CeA.

Another way the VTA may exert its influence on fear learning and responding is through projections to the BLA. In an appetitive setting, cellular recording data show that VTA dopaminergic activity normally inhibits BLA neuronal activity (Esber et al., 2012).

The BLA projects to the CeA, including the CeM and CeL (Pitkänen et al., 1997). Hence, VTA dopaminergic projections may tonically inhibit the BLA, so once dopaminergic projections from the VTA are removed, it releases the BLA from inhibition. Once released from inhibition, the BLA may convey additional information about stimuli to the CeA (including CeL and CeM), to produce the appropriate output (defensive behavior) via its projection to downstream targets, like the PAG.

As mentioned previously, the CeM→dIPAG pathway is a likely candidate for the expression of flight (Li and Sheets, 2018; Borkar et al., 2024). Therefore, there are two potential ways the VTA can influence fear processes and responding: first, through its projections to the CeA, and second, through its projections to the BLA. The VTA→CeA (CeL and CeM)→dIPAG pathway and VTA→BLA→CeM→dIPAG pathway with the CeM receiving additional information from the dorsal peduncular (Figure 5.2), are potential pathways that are part of a larger fear circuit for the integration of sensory and associative information and subsequent behavioral output (defensive behaviors). This proposed brain network would support the hypothesis that VTA dopamine plays a modulatory role in fear learning and expression, and more specifically, VTA dopamine suppresses activity-promoting behaviors like jumping to a danger cue during extinction. In sum, these studies indicate that VTA dopamine alters fear learning and behavior through its projections to the CeA and BLA. Future studies probing the projections from the VTA to the BLA and CeA, as well as the CeA to the dPAG (including dIPAG) and the effects on behavior would provide a more complete understanding of the full circuitry underlying the expression of activity-promoting defensive behaviors during fear discrimination.



**Figure 5.2.** Linking the VTA to the cortico-amygdalar-PAG pathway for fear learning and expression of flight and other activity-promoting defensive behaviors

A recently identified pathway for flight from DP→CeM→dl/IPAG is shown (Borkar et al., 2024). Known dopaminergic VTA projections to the amygdala include inhibitory inputs to the BLA in an appetitive setting (Esber et al., 2012) and excitatory input to the CeM to facilitate fear learning (Jo et al., 2018). The BLA projects to the CeA, including CeM and CeL, and the BLA→CeM pathway is one way VTA dopaminergic projections may influence flight (dashed lines).

#### 5.2.4 *Relevance to clinical research*

Understanding the neural underpinnings of fear responses is important for our general knowledge of anxiety and panic disorders. Disrupted threat processing and inappropriate fear responses are a hallmark of Post-Traumatic Stress Disorder (PTSD) and anxiety disorders (American Psychiatric Association, DSM-5 Task Force, 2013). Like humans, animals express fear in many ways – from immobile defensive responses such as freezing to activity-promoting responses such as locomotion, jumping, and more. Despite this fact, the majority of fear research in animal models have used conditioned freezing as a marker for fear in the past several years. As more and more research laboratories observe and investigate other defensive behaviors in their behavioral settings, it is becoming increasingly apparent that fear research is lacking in comprehensive behavioral measures. The work presented here emphasizes the need for more thorough investigations of animal defensive behaviors. Importantly, this work shows that non-freezing behaviors, specifically locomotion, are robustly expressed in fear conditioning settings, highlighting the feasibility of measuring activity-promoting behaviors in fear research.

In all of the experiments presented here, both females and males were included. The findings of the first experiment in Chapter 2 reveal that there are sex differences in the temporal organization of freezing, where males maintain high freezing levels throughout the 10 s danger cue presentation, while females exhibit high levels of freezing during the early cue period and low levels of freezing during the late cue period. This finding is particularly significant, since the majority of past fear conditioning studies looking at freezing were conducted exclusively in males, and only more recently,

has there been a shift in neuroscience research to include both sexes (Shansky, 2018). This suggests that freezing may not be as a reliable measure of fear for females as it is for males. In support, Shansky and lab observe sex differences in the expression of darting, where females tend to dart more in comparison to males (Gruene et al., 2015; Mitchell et al., 2022). These findings suggest that females may express conditioned fear differently from males. Importantly, in the U.S., anxiety disorders are more prevalent in females than in males (McLean et al., 2011). This underscores the value of including females in animal studies for fear in order to understand potential sex differences in the expression of fear and also in the neural mechanisms by which fear processing and responding manifests. This research is crucial to understanding anxiety disorders in humans and properly treating those affected by anxiety disorders.

### **5.3 Conclusion**

The work presented here aimed to reveal the full range of behavioral responding during fear discrimination and investigate the contributions of VTA dopamine in defensive behaviors. The results support other findings that activity-promoting behaviors occur in fear conditioning settings, and the results are the first to present comprehensive temporal ethograms of behavior during fear discrimination. Additionally, the findings uncover a possible role for VTA dopamine in the brain circuit underlying the expression of activity-promoting defensive behaviors. Future work is needed to fully illustrate the VTA's involvement in the pathway responsible for conditioned defensive responses, but the present results provide a strong foundation for future behavior-focused experiments probing the neural circuits for defensive behaviors.

## References

- Amorapanth P, Nader K, LeDoux JE (1999) Lesions of Periaqueductal Gray Dissociate-Conditioned Freezing From Conditioned Suppression Behavior in Rats. *Learn Mem* 6:491–499.
- Anagnostaras SG, Maren S, Sage JR, Goodrich S, Fanselow MS (1999) Scopolamine and Pavlovian fear conditioning in rats: dose-effect analysis. *Neuropsychopharmacology* 21:731–744.
- Anon (2013) *Diagnostic and statistical manual of mental disorders: DSM-5™*, 5th ed. Arlington, VA, US: American Psychiatric Publishing, Inc.
- Arico C, Bagley EE, Carrive P, Assareh N, McNally GP (2017) Effects of chemogenetic excitation or inhibition of the ventrolateral periaqueductal gray on the acquisition and extinction of Pavlovian fear conditioning. *Neurobiol Learn Mem* 144:186–197.
- Badrinarayan A, Wescott SA, Vander Weele CM, Saunders BT, Couturier BE, Maren S, Aragona BJ (2012) Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J Neurosci* 32:15779–15790.
- Blanchard RJ, Blanchard DC (1969) Crouching as an index of fear. *J Comp Physiol Psychol* 67:370–375.
- Blanchard RJ, Flannelly KJ, Blanchard DC (1986) Defensive behavior of laboratory and wild *Rattus norvegicus*. *J Comp Psychol* 100:101–107.

Bohnslav JP, Wimalasena NK, Clausing KJ, Dai YY, Yarmolinsky DA, Cruz T, Kashlan AD, Chiappe ME, Orefice LL, Woolf CJ, Harvey CD (2021) DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. *eLife* 10:e63377.

Bolles RC, Collier AC (1976) The effect of predictive cues on freezing in rats. *Animal Learning & Behavior* 4:6–8.

Borkar CD, Stelly CE, Fu X, Dorofeikova M, Le Q-SE, Vutukuri R, Vo C, Walker A, Basavanhalli S, Duong A, Bean E, Resendez A, Parker JG, Tasker JG, Fadok JP (2024) Top-down control of flight by a non-canonical cortico-amygdala pathway. *Nature* 625:743–749.

Borowski TB, Kokkinidis L (n.d.) Contribution of Ventral Tegmental Area Dopamine Neurons to Expression of Conditional Fear: Effects of Electrical Stimulation, Excitotoxin Lesions, and Quinpirole Infusion on Potentiated Startle in Rats.

Bouton ME, Bolles RC (1980) Conditioned fear assessed by freezing and by the suppression of three different baselines. *Animal Learning & Behavior* 8:429–434.

Carrive P, Leung P, Harris J, Paxinos G (1997) Conditioned fear to context is associated with increased Fos expression in the caudal ventrolateral region of the midbrain periaqueductal gray. *Neuroscience* 78:165–177.

Chu A, Gordon NT, DuBois AM, Michel CB, Hanrahan KE, Williams DC, Anzellotti S, McDannald MA (2024) A fear conditioned cue orchestrates a suite of behaviors in rats. *Elife* 13:e82497.

- Costa KM, Schoenbaum G (2022) Dopamine. *Current Biology* 32:R817–R824.
- Deng H, Xiao X, Wang Z (2016) Periaqueductal Gray Neuronal Activities Underlie Different Aspects of Defensive Behaviors. *Journal of Neuroscience* 36:7580–7588.
- Dielenberg RA, McGregor IS (2001) Defensive behavior in rats towards predatory odors: a review. *Neurosci Biobehav Rev* 25:597–609.
- DiLeo A, Wright KM, McDannald MA (2016) Subsecond fear discrimination in rats: adult impairment in adolescent heavy alcohol drinkers. *Learn Mem* 23:618–622.
- Esber GR, Roesch MR, Bali S, Trageser J, Bissonette GB, Puche AC, Holland PC, Schoenbaum G (2012) Attention-related Pearce-Kaye-Hall signals in basolateral amygdala require the midbrain dopaminergic system. *Biol Psychiatry* 72:1012–1019.
- Esser R, Korn CW, Ganzer F, Haaker J (2021) L-DOPA modulates activity in the vmPFC, nucleus accumbens, and VTA during threat extinction learning in humans. *eLife* 10:e65280.
- Estes WK, Skinner BF (1941) Some quantitative properties of anxiety. *Journal of Experimental Psychology* 29:390–400.
- Fadok JP, Dickerson TMK, Palmiter RD (2009) Dopamine Is Necessary for Cue-Dependent Fear Conditioning. *J Neurosci* 29:11089–11097.

- Fadok JP, Krabbe S, Markovic M, Courtin J, Xu C, Massi L, Botta P, Bylund K, Müller C, Kovacevic A, Tovote P, Lüthi A (2017) A competitive inhibitory circuit for selection of active and passive fear responses. *Nature* 542:96–100.
- Fallon JH, Koziell DA, Moore RY (1978) Catecholamine innervation of the basal forebrain II. Amygdala, suprarhinal cortex and entorhinal cortex. *Journal of Comparative Neurology* 180:509–531.
- Fanselow MS (1982) The postshock activity burst. *Animal Learning & Behavior* 10:448–454.
- Fanselow MS (1994) Neural organization of the defensive behavior system responsible for fear. *Psychon Bull Rev* 1:429–438.
- Fanselow MS, Decola JP, De Oca BM, Landeira-Fernandez J (1995) Ventral and dorsolateral regions of the midbrain periaqueductal gray (PAG) control different stages of defensive behavior: Dorsolateral PAG lesions enhance the defensive freezing produced by massed and immediate shock. *Aggr Behav* 21:63–77.
- Fanselow MS, Hoffman AN, Zhuravka I (2019) Timing and the transition between modes in the defensive behavior system. *Behav Processes* 166:103890.
- Fanselow MS, Lester LS (1988) A functional behavioristic approach to aversively motivated behavior: Predatory imminence as a determinant of the topography of defensive behavior. In: *Evolution and learning*, pp 185–212. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

- Fellinger L, Jo YS, Hunker AC, Soden ME, Elum J, Juarez B, Zweifel LS (2021) A midbrain dynorphin circuit promotes threat generalization. *Current Biology* 31:4388-4396.e5.
- Fibiger H, LePiane F, Jakubovic A, Phillips A (1987) The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *J Neurosci* 7:3888–3896.
- Foilib AR, Flyer-Adams JG, Maier SF, Christianson JP (2016) Posterior insular cortex is necessary for conditioned inhibition of fear. *Neurobiol Learn Mem* 134:317–327.
- Frick A, Björkstrand J, Lubberink M, Eriksson A, Fredrikson M, Åhs F (2022) Dopamine and fear memory formation in the human amygdala. *Mol Psychiatry* 27:1704–1711.
- Furlong TM, Richardson R, McNally GP (2016) Habituation and extinction of fear recruit overlapping forebrain structures. *Neurobiol Learn Mem* 128:7–16.
- Gabriel CJ, Zeidler Z, Jin B, Guo C, Goodpaster CM, Kashay AQ, Wu A, Delaney M, Cheung J, DiFazio LE, Sharpe MJ, Aharoni D, Wilke SA, DeNardo LA (2022) BehaviorDEPOT is a simple, flexible tool for automated behavioral detection based on markerless pose tracking. *eLife* 11:e74314.
- Goosens KA, Maren S (2001) Contextual and Auditory Fear Conditioning are Mediated by the Lateral, Basal, and Central Amygdaloid Nuclei in Rats. *Learn Mem* 8:148–155.
- Gruene TM, Flick K, Stefano A, Shea SD, Shansky RM (2015) Sexually divergent expression of active and passive conditioned fear responses in rats. *Elife* 4:e11352.

- He K, Zhang X, Ren S, Sun J (2016) Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 770–778. Las Vegas, NV, USA: IEEE. Available at: <http://ieeexplore.ieee.org/document/7780459/> [Accessed April 17, 2024].
- Holland PC (1977) Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *J Exp Psychol Anim Behav Process* 3:77–104.
- Holland PC (1979) The effects of qualitative and quantitative variation in the US on individual components of Pavlovian appetitive conditioned behavior in rats. *Animal Learning & Behavior* 7:424–432.
- Hsu AI, Yttri EA (2021) B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat Commun* 12:5188.
- Iordanova MD, Westbrook RF, Killcross AS (2006) Dopamine activity in the nucleus accumbens modulates blocking in fear conditioning. *Eur J Neurosci* 24:3265–3270.
- Jo YS, Heymann G, Zweifel LS (2018) Dopamine Neurons Reflect the Uncertainty in Fear Generalization. *Neuron* 100:916-925.e3.
- Jones CE, Monfils M-H (2016) Dominance status predicts social fear transmission in laboratory rats. *Anim Cogn* 19:1051–1069.
- Kamin LJ (1967) Predictability, surprise, attention, and conditioning. Available at: <https://ntrs.nasa.gov/citations/19680014821> [Accessed May 17, 2024].

- Kemble ED, Blanchard DC, Blanchard RJ (1990) Effects of regional amygdaloid lesions on flight and defensive behaviors of wild black rats (*Rattus rattus*). *Physiology & Behavior* 48:1–5.
- Killcross S, Robbins TW, Everitt BJ (1997) Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature* 388:377–380.
- Kim EJ, Horovitz O, Pellman BA, Tan LM, Li Q, Richter-Levin G, Kim JJ (2013) Dorsal periaqueductal gray-amygdala pathway conveys both innate and learned fear responses in rats. *Proceedings of the National Academy of Sciences* 110:14795–14800.
- Kim JJ, Rison RA, Fanselow MS (1993) Effects of amygdala, hippocampus, and periaqueductal gray lesions on short- and long-term contextual fear. *Behav Neurosci* 107:1093–1098.
- Kong M-S, Zweifel LS (2021) Central amygdala circuits in valence and salience processing. *Behav Brain Res* 410:113355.
- Koo JW, Han J-S, Kim JJ (2004) Selective Neurotoxic Lesions of Basolateral and Central Nuclei of the Amygdala Produce Differential Effects on Fear Conditioning. *J Neurosci* 24:7654–7662.
- LaBarbera JD, Caul WF (1976) An opponent-process interpretation of postshock bursts in appetitive responding. *Animal Learning & Behavior* 4:386–390.

Lázaro-Muñoz G, LeDoux JE, Cain CK (2010) Sidman Instrumental Avoidance Initially Depends on Lateral and Basal Amygdala and Is Constrained by Central Amygdala-Mediated Pavlovian Processes. *Biological Psychiatry* 67:1120–1127.

LeDoux JE (n.d.) Emotion Circuits in the Brain.

Lever C, Burton S, O'Keefe J (2006) Rearing on hind legs, environmental novelty, and the hippocampal formation. *Rev Neurosci* 17:111–133.

Li J-N, Sheets PL (2018) The central amygdala to periaqueductal gray pathway comprises intrinsically distinct neurons differentially affected in a model of inflammatory pain. *The Journal of Physiology* 596:6289–6305.

Loughlin SE, Fallon JH (1983) Dopaminergic and non-dopaminergic projections to amygdala from substantia nigra and ventral tegmental area. *Brain Research* 262:334–338.

Luo R, Uematsu A, Weitemier A, Aquili L, Koivumaa J, McHugh TJ, Johansen JP (2018) A dopaminergic switch for fear to safety transitions. *Nat Commun* 9:2483.

Maren S, Aharonov G, Fanselow MS (1997) Neurotoxic lesions of the dorsal hippocampus and Pavlovian fear conditioning in rats. *Behav Brain Res* 88:261–274.

Maren S, Fanselow MS (1996) The Amygdala and Fear Conditioning: Has the Nut Been Cracked? *Neuron* 16:237–240.

Mast M, Blanchard RJ, Blanchard DC (1982) The relationship of freezing and response suppression in a CER situation. *The Psychological Record* 32:151–167.

Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* 21:1281–1289.

McDannald MA (2010) Contributions of the amygdala central nucleus and ventrolateral periaqueductal grey to freezing and instrumental suppression in Pavlovian fear conditioning. *Behav Brain Res* 211:111–117.

McDannald MA (2023) Pavlovian Fear Conditioning Is More than You Think It Is. *J Neurosci* 43:8079–8087.

McDannald MA, Galarce EM (2011) Measuring Pavlovian fear with conditioned freezing and conditioned suppression reveals different roles for the basolateral amygdala. *Brain Res* 1374:82–89.

McLean CP, Asnaani A, Litz BT, Hofmann SG (2011) Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. *J Psychiatr Res* 45:1027–1035.

McNally GP, Johansen JP, Blair HT (2011) Placing prediction into the fear circuit. *Trends in Neurosciences* 34:283–292.

Menegas W, Akiti K, Amo R, Uchida N, Watabe-Uchida M (2018) Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat Neurosci* 21:1421–1430.

Mitchell JR, Trettel SG, Li AJ, Wasielewski S, Huckleberry KA, Fanikos M, Golden E, Laine MA, Shansky RM (2022) Darting across space and time: parametric modulators of sex-biased conditioned fear responses. *Learn Mem* 29:171–180.

Mobbs D, Petrovic P, Marchant JL, Hassabis D, Weiskopf N, Seymour B, Dolan RJ, Frith CD (2007) When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science* 317:1079–1083.

Morgan MM, Whitney PK, Gold MS (1998) Immobility and flight associated with antinociception produced by activation of the ventral and lateral/dorsal regions of the rat periaqueductal gray.

Nader K, LeDoux JE (n.d.) Inhibition of the Mesoamygdala Dopaminergic Pathway Impairs the Retrieval of Conditioned Fear Associations.

Nikbakht N, Diamond ME (n.d.) Conserved visual capacity of rats under red light. *eLife* 10:e66429.

Oliveira ARD, Reimer AE, Brandão ML (2009) Role of dopamine receptors in the ventral tegmental area in conditioned fear. *Behavioural Brain Research* 199:271–277.

Oliveira LC, Nobre MJ, Brandão ML, Landeira-Fernandez J (2004) Role of amygdala in conditioned and unconditioned fear generated in the periaqueductal gray. *NeuroReport* 15:2281.

Ozawa T, Ycu EA, Kumar A, Yeh L-F, Ahmed T, Koivumaa J, Johansen JP (2017) A feedback neural circuit for calibrating aversive memory strength. *Nat Neurosci* 20:90–97.

- Papp M, Bal A (1986) Motivational versus motor impairment after haloperidol injection or 6-OHDA lesions in the ventral tegmental area or substantia nigra in rats. *Physiology & Behavior* 38:773–779.
- Parnas AS, Weber M, Richardson R (2005) Effects of multiple exposures to D-cycloserine on extinction of conditioned fear in rats. *Neurobiol Learn Mem* 83:224–231.
- Phillips RG, LeDoux JE (1992) Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav Neurosci* 106:274–285.
- Phung VH, Rhee EJ (2019) A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. *Applied Sciences* 9:4500.
- Pickens CL, Navarre BM, Nair SG (2010) Incubation of conditioned fear in the conditioned suppression model in rats: role of food-restriction conditions, length of conditioned stimulus, and generality to conditioned freezing. *Neuroscience* 169:1501–1510.
- Pitkänen A, Savander V, LeDoux JE (1997) Organization of intra-amygdaloid circuitries in the rat: an emerging framework for understanding functions of the amygdala. *Trends in Neurosciences* 20:517–523.
- Quirk GJ (2002) Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. *Learn Mem* 9:402–407.

- Ray MH, Moaddab M, McDannald MA (2022) Threat and Bidirectional Valence Signaling in the Nucleus Accumbens Core. *J Neurosci* 42:817–833.
- Ray MH, Russ AN, Walker RA, McDannald MA (2020) The Nucleus Accumbens Core is Necessary to Scale Fear to Degree of Threat. *J Neurosci* 40:4750–4760.
- Reis FLV, Masson S, De Oliveira AR, Brandão ML (2004) Dopaminergic mechanisms in the conditioned and unconditioned fear as assessed by the two-way avoidance and light switch-off tests. *Pharmacology Biochemistry and Behavior* 79:359–365.
- Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical Conditioning II: Current Research and Theory*.
- Rescorla RA (1967) Pavlovian conditioning and its proper control procedures. *Psychol Rev* 74:71–80.
- Rescorla RA (1968) Probability of shock in the presence and absence of CS in fear conditioning. *J Comp Physiol Psychol* 66:1–5.
- Rogers JL, Kesner RP (2004) Cholinergic modulation of the hippocampus during encoding and retrieval of tone/shock-induced fear conditioning. *Learn Mem* 11:102–107.
- Rosas JM, Alonso G (1996) Temporal Discrimination and Forgetting of CS Duration in Conditioned Suppression. *Learning and Motivation* 27:43–57.

- Salinas-Hernández XI, Zafiri D, Sigurdsson T, Duvarci S (2023) Functional architecture of dopamine neurons driving fear extinction learning. *Neuron* 111:3854-3870.e5.
- Schmitt P, Carrive P, Di Scala G, Jenck F, Brandao M, Bagri A, Moreau J-L, Sandner G (1986) A neuropharmacological study of the periventricular neural substrate involved in flight. *Behavioural Brain Research* 22:181–190.
- Shansky RM (2018) Sex differences in behavioral strategies: avoiding interpretational pitfalls. *Curr Opin Neurobiol* 49:95–98.
- Shumake J, Furgeson-Moreira S, Monfils MH (2014) Predictability and heritability of individual differences in fear learning. *Anim Cogn* 17:1207–1221.
- Spiacci A, Sergio T de O, da Silva GSF, Glass ML, Schenberg LC, Garcia-Cairasco N, Zangrossi H (2015) Serotonin in the dorsal periaqueductal gray inhibits panic-like defensive behaviors in rats exposed to acute hypoxia. *Neuroscience* 307:191–198.
- Strickland J, Dileo A, Moaddab M, Ray M, Walker R, Wright K, McDannald M (2021) Foot shock facilitates reward seeking in an experience-dependent manner. *Behav Brain Res* 399:112974.
- Totty MS, Warren N, Huddleston I, Ramanathan KR, Ressler RL, Oleksiak CR, Maren S (2021) Behavioral and brain mechanisms mediating conditioned flight behavior in rats. *Sci Rep* 11:8215.
- Tovote P, Esposito MS, Botta P, Chaudun F, Fadok JP, Markovic M, Wolff SBE, Ramakrishnan C, Fenno L, Deisseroth K, Herry C, Arber S, Lüthi A (2016) Midbrain circuits for defensive behaviour. *Nature* 534:206–212.

- Trott JM, Hoffman AN, Zhuravka I, Fanselow MS (2022) Conditional and unconditional components of aversively motivated freezing, flight and darting in mice. *eLife* 11:e75663.
- Verharen JPH, Luijendijk MCM, Vanderschuren LJMJ, Adan RAH (2020) Dopaminergic contributions to behavioral control under threat of punishment in rats. *Psychopharmacology* 237:1769–1782.
- Walker RA, Andreansky C, Ray MH, McDannald MA (2018) Early adolescent adversity inflates threat estimation in females and promotes alcohol use initiation in both sexes. *Behav Neurosci* 132:171–182.
- Wilensky AE, Schafe GE, LeDoux JE (1999) Functional inactivation of the amygdala before but not after auditory fear conditioning prevents memory formation. *J Neurosci* 19:RC48.
- Wright KM, DiLeo A, McDannald MA (2015) Early adversity disrupts the adult use of aversive prediction errors to reduce fear in uncertainty. *Front Behav Neurosci* 9:227.
- Wright KM, Jhou TC, Pimpinelli D, McDannald MA (n.d.) Cue-inhibited ventrolateral periaqueductal gray neurons signal fear output and threat probability in male rats. *eLife* 8:e50054.
- Wright KM, McDannald MA (2019) Ventrolateral periaqueductal gray neurons prioritize threat probability over fear output. *Elife* 8:e45013.

Yau JO-Y, McNally GP (2022) The activity of ventral tegmental area dopamine neurons during shock omission predicts safety learning. *Behavioral Neuroscience* 136:276–284.

Zambetti PR, Schuessler BP, Lecamp BE, Shin A, Kim EJ, Kim JJ (2021) Pavlovian fear conditioning does not readily occur in rats in naturalistic environments.

:2021.10.20.465116 Available at:

<https://www.biorxiv.org/content/10.1101/2021.10.20.465116v1> [Accessed May 18, 2024].

Zhu Y, Lan Z, Newsam S, Hauptmann A (2019) Hidden Two-Stream Convolutional Networks for Action Recognition. In: *Computer Vision – ACCV 2018* (Jawahar CV, Li H, Mori G, Schindler K, eds), pp 363–378. Cham: Springer International Publishing.